1

**PM-1 3½"x4" PHOTOGRAPHIC MICROCOPY TARGET
NBS 1010a ANSI/ISO #2 EQUIVALENT**

PRECISIONSM RESOLUTION TARGETS

PIONEERS IN METHYLENE BLUE TESTING SINCE 1954

1000 COUNTY ROAD 1, DUNWOODY, GA 30027, USA
TEL. 404 435 7800 FAX 404 435 7807 TLA 04000000

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

**WEAK CONVERGENCE OF STOCHASTIC PROCESSES
IN WEIGHTED METRICS AND THEIR APPLICATIONS TO
CONTIGUOUS CHANGEPOINT ANALYSIS**

by

Barbara Szyszkowicz M.Sc.

A thesis submitted to
the Faculty of Graduate Studies and Research
in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

Department of Mathematics and Statistics
Carleton University
Ottawa, Ontario
June 1992

© Copyright, 1992
Barbara Szyszkowicz

National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-79869-6

Canada

The undersigned hereby recommend to
The Faculty of Graduate Studies and Research
acceptance of the thesis

WEAK CONVERGENCE OF STOCHASTIC PROCESSES
IN WEIGHTED METRICS AND THEIR APPLICATIONS TO
CONTIGUOUS CHANGEPOINT ANALYSIS

submitted by
BARBARA SZYSZKOWICZ M.Sc.

in partial fulfilment of the requirements
for the degree of Doctor of Philosophy

Chair, Department of Mathematics and Statistics

Thesis Supervisor

External Examiner

Carleton University

June 1992

To Mietek and Sebastian

"Although this may seem a paradox, all science
is dominated by the idea of approximation."

Bertrand Russell

ABSTRACT

This dissertation is concerned with asymptotic methods in probability and statistics. In particular, we deal with invariance principles in terms of studying the weak convergence of various stochastic processes in weighted metrics based on independent identically distributed random variables (i.i.d.r.v.'s), as well as on independent observations whose distributions are governed by contiguous measures. The latter type of results will enable us to describe the possibility of a change in the distribution of large sets of chronologically ordered data.

The approach we use to prove invariance principles is not the "traditional" way of proving tightness and identifying final dimensional distributions. Rather, we use the technique of approximating stochastic processes of interest by their ideal versions in the limit, usually Gaussian processes.

Partial sums of independent r.v.'s, and the standard Wiener process are the main building blocks for the stochastic processes considered in the thesis. In Chapter 1 we approximate partial sums of i.i.d.r.v.'s by a Wiener process in weighted supremum norm and L_p distances. The obtained results are optimal in terms of the weight functions used, as well as in the assumed moment conditions.

Chapter 2 is expository, describing the theory of contiguity as created by Le Cam. We show also that the optimal $\|\cdot/q\|$ -metric weak convergence results of Chapter 1 remain true under contiguous measures.

Chapter 3 summarizes various notions of changepoint analysis, describing also how the results that are obtained later on relate to some of these.

We study partial sum processes of sequential ranks in Chapter 4, while Chapters 5 and 6 are devoted to "bridge type partial sums" and to their generalized U -statistics type processes respectively.

Chapter 7 is on extensions of the so-called Kendall-Kendall pontogram via studying the asymptotic behaviour of weighted nonhomogeneous Poisson bridges.

We consider the classical uniform empirical process, as well as various types of two-time parameter empirical processes in Chapter 8.

The mentioned processes of Chapters 4-8 are studied in weighted metrics under the null assumption of i.i.d. sampling versus contiguous alternatives.

The many repetitions in describing the hypotheses considered, as well as the class of weight functions, are done on purpose for the convenience of the reader. Due to the variability of the topics, we prefer to give as much as possible of the details in the formulation of our results and of their proofs instead of sending the reader back to previous chapters.

The following is a list of the author's own results which are believed to be new:

Theorem 1.2.1	Theorem 8.7.1
Theorem 1.4.1	Theorem 8.7.2
Theorem 2.5.1	Theorem 8.9.1
Theorem 4.1.1	Proposition 7.2.1
Theorem 4.1.2	Proposition 7.2.2
Theorem 4.2.1	Corollary 1.2.1
Theorem 5.2.1	Corollary 4.1.1
Theorem 5.3.1	Corollary 4.2.1
Theorem 5.4.1	Corollary 5.2.1
Theorem 6.1.1	Corollary 5.4.1
Theorem 6.1.2	Corollary 6.3.1
Theorem 6.2.1	Corollary 6.3.2
Theorem 6.2.2	Corollary 6.3.3
Theorem 6.3.1	Corollary 6.3.4
Theorem 6.3.2	Corollary 7.3.1
Theorem 6.3.3	Corollary 7.3.2
Theorem 6.3.4	Corollary 8.2.1
Theorem 7.3.1	Corollary 8.3.1
Theorem 8.2.1	Corollary 8.4.1
Theorem 8.3.1	Corollary 8.5.1
Theorem 8.4.1	Corollary 8.8.1
Theorem 8.5.1	Corollary 8.8.2
Theorem 8.6.1	

Lemmas 8.2.1 and 8.4.1 amount to somewhat new, useful observations.

In the next list the results themselves or their proofs are new in some of their parts only.

Lemma 2.3.1	Proposition 7.1.2
Theorem 2.4.1	Corollary 7.1.1
Lemma 2.4.1	Lemma 8.7.1
Proposition 7.1.1	Lemma 8.7.2

ACKNOWLEDGEMENTS

My deep gratitude is gratefully expressed to Professor Miklós Csörgő for his continuous guidance, encouragement, helpful advice and suggestions freely given during my stay at Carleton University as a Ph.D. student. I thank him for his time, the numerous discussions we have had and, in final analysis, for teaching me what doing mathematics is all about.

I also thank Miklós Csörgő and Lajos Horváth for sharing their forthcoming book with me, prior to its publication.

I am grateful to the members of my Advisory Committee (Professors Miklós Csörgő, Donald Dawson, Gail Ivanoff) for their interest in my studies and research, and for giving me the opportunity to speak on my results on several occasions in the last three years.

My heartfelt thanks go to my parents and my brothers in Poland for their moral support and appreciation of my work.

Thanks to my husband, Mietek, for his strong encouragement in my studies and to my wonderful son, Sebastian, for being so proud of his mother and of her mathematics; to both of them for sharing with me all my highs and lows.

Diane Berezowski is to be thanked for her excellent typing of this manuscript.

This research was supported by a Scholarship from the Faculty of Graduate Studies and Research and by an NSERC Canada Postgraduate Scholarship.

Table of Contents

Abstract	iv
Acknowledgements	vii
Introduction	xii
Chapter 1 Weighted approximations of partial sums of i.i.d.r.v.'s	
1.0. Introduction	1
1.1. Local functions of a Wiener process	3
1.2. Sup-metric approximation of partial sums	11
1.3. Random integrals	22
1.4. L_p -approximations of partial sums	23
Chapter 2 Weighted partial sum processes under contiguous measures	
2.0. Introduction	28
2.1. The notion of contiguity	30
2.2. Contiguity of product measures	43
2.3. A parametrization of contiguity	47
2.4. Partial sums under contiguous measures	49
2.5. Weighted partial sums under contiguous measures	51

Chapter 3 Changepoint Problem

3.0. Introduction.....	54
3.1. Change in the mean.....	58
3.2. L^2 -statistic type processes.....	63
3.3. Sequential ranks.....	65
3.4. Kendall-Kendall pontograms.....	67
3.5. Empirical processes.....	71

Chapter 4 Weighted partial sum processes of sequential ranks

4.0. Introduction.....	77
4.1. Null hypothesis.....	80
4.2. Contiguous alternatives.....	87

Chapter 5 Tied Down Partial Sum Processes

5.0. Introduction.....	95
5.1. Local functions of a Brownian bridge.....	96
5.2. Supremum norm approximation.....	98
5.3. L_p -approximations.....	107
5.4. Contiguous alternatives.....	114

Chapter 6 U-statistic Type Processes

6.0. Introduction	119
6.1. Symmetric kernel.....	123
6.2. Antisymmetric kernel	138
6.3. Contiguous alternatives.....	142

Chapter 7 Pontograms (Poisson bridges) with changepoints in their intensity function

7.0. Introduction	153
7.1. Weighted pontograms in sup-metrics.....	157
7.2. L_p-norms of weighted pontograms.....	165
7.3. Contiguous alternatives and a changepoint problem	174

Chapter 8 Weighted Two-Time Parameter Empirical Processes Under Contiguous Measures

8.0. Introduction	183
8.1. Uniform empirical processes.....	185
8.2. Weighted two-time parameter empirical processes based on observations.....	189
8.3. Two-time parameter empirical processes tied down at $t = 1$.	200
8.4. Weighted empirical processes based on ranks.....	209

8.5. Weighted empirical processes based on sequential ranks.....	216
8.6. Two-time parameter empirical processes of sequential ranks tied down at $t = 1$	221
8.7. Contiguous alternatives.....	226
8.8. Changepoint alternatives	233
8.9. Weighted uniform empirical processes under contiguous measures.....	234
 Appendix	
A.1. Some measure theoretical results	237
A.2. Skorohod-Dudley-Wichura theorem.....	238
A.3. The Zero-One law for Brownian motion.....	239
References.....	241

INTRODUCTION

Invariance principles constitute a significantly new chapter in the development of the theory of limit theorems in probability and statistics. They have evolved from two major sources: partial sum processes and empirical processes. Both will play a dominant role in this thesis, where we deal with invariance principles in terms of studying the weak convergence, in weighted metrics, of various stochastic processes based on independent identically distributed random variables (i.i.d.r.v.'s), as well as on independent observations whose distributions are governed by contiguous measures.

Concerning partial sums, Erdős and Kac (1946) had the idea of deriving the asymptotic distributions of the functionals

$$n^{-1/2} \max_{1 \leq k \leq n} S_k, \quad n^{-1/2} \max_{1 \leq k \leq n} |S_k|, \quad n^{-2} \sum_{k=1}^n S_k^2, \quad n^{-3/2} \sum_{k=1}^n |S_k|$$

of partial sums $S_k = \sum_{i=1}^k X_i$ ($k = 1, 2, \dots$) of mean zero and variance 1 i.i.d.r.v.'s X_i ($i = 1, 2, \dots$) via first calculating their limiting distributions based on convenient forms of their underlying distribution function F , and then showing that in the limit ($n \rightarrow \infty$), the above functionals did not remember the initially assumed forms for F . They called this method of proof the *invariance principle*.

As to the role played by empirical processes in the origin of the invariance principle, Doob (1949) considered the empirical distribution function

$$E_n(s) = n^{-1} \sum_{i=1}^n 1\{U_i \leq s\}, \quad 0 \leq s \leq 1,$$

of the first $n \geq 1$ of independent uniform $-(0, 1)$ r.v.'s $U_1, U_2, \dots$. On account of the uniform empirical processes

$$\alpha_n(s) = n^{1/2}(E_n(s) - s), \quad 0 \leq s \leq 1,$$

having, for each $n \geq 1$, the same covariance structure as that of a Brownian bridge Gaussian process $\{B(s), 0 \leq s \leq 1\}$, i.e. $EB(s) = E\alpha_n(s) = 0$ and $EB(s)B(t) = E\alpha_n(s)\alpha_n(t) = s(1-t)$, $0 \leq s \leq t \leq 1$ and, by the multivariate central limit theorem, as $n \rightarrow \infty$, the limit of the joint distribution of $(\alpha_n(s_1), \alpha_n(s_2), \dots, \alpha_n(s_k))$ being the corresponding finite dimensional distribution of $(B(s_1), B(s_2), \dots, B(s_k))$, $0 \leq s_1 \leq s_2 \leq \dots \leq s_k \leq 1$, he argued the *invariance principle* that in calculating asymptotic $\alpha_n(s)$ process distributions when $n \rightarrow \infty$ we should simply replace the $\alpha_n(s)$ process by the Gaussian $B(s)$ process. In particular, since the direct evaluation of the limiting distribution of $\sup_{0 \leq s \leq 1} \alpha_n(s)$ (resp. $\sup_{0 \leq s \leq 1} |\alpha_n(s)|$) is rather complicated (cf. Smirnov 1939 (resp. Kolmogorov 1933) and Feller 1948), Doob (1949) proceeded to evaluate the distribution of $\sup_{0 \leq s \leq 1} B(s)$ and also that of $\sup_{0 \leq s \leq 1} |B(s)|$, leaving the justification of his heuristic approach open.

Donsker (1951) succeeded in extending the results of Erdős and Kac (1946) via proving his famous functional central limit theorem. His idea was that from the partial sums $\{S_n\}_{n=0}^{\infty}$ ($S_0 \equiv 0$) of mean zero and variance 1 i.i.d.r.v.'s $X_1, X_2, \dots$ one should construct a sequence of stochastic processes $\{S_n(t), 0 \leq t \leq 1\}$ on $C[0, 1]$ so that $S_n(k/n) = S_k/n^{1/2}$, with $S_n(t)$ being the linear interpolation of the latter for $k/n < t < (k+1)/n$, $k = 0, 1, \dots, n$, and study the limiting behaviour of $S_k/n^{1/2}$ via that of $S_n(t)$ on $C[0, 1]$. He showed that the distributional properties of $\{S_n(t), 0 \leq t \leq 1\}$ as $n \rightarrow \infty$, coincide with those of a standard Wiener process (Brownian motion) $\{W(t), 0 \leq t \leq 1\}$, i.e. a Gaussian process with $EW(t) = 0$ and $EW(s)W(t) = s$, $0 \leq s \leq t \leq 1$. One possible way of saying this precisely is via Donsker's *functional central limit theorem* (1951): $h(S_n(t)) \xrightarrow{\mathcal{D}} h(W(t))$ as $n \rightarrow \infty$ for all $h : C[0, 1] \rightarrow \mathbb{R}$ that are supremum norm $\|\cdot\|$ continuous almost surely (a.s.) with respect to Wiener measure W . Here $\xrightarrow{\mathcal{D}}$ means convergence in distribution,

i.e. convergence of the sequence of distribution functions generated by $\{h(S_n(t))\}$ to that of $h(W(t))$. Donsker (1952) was first also in proving a similar functional central limit theorem for the empirical processes $\{\alpha_n(s), 0 \leq s \leq 1\}$. In both cases we will simply write

$$n^{-1/2}S_{[nt]} \xrightarrow{\mathcal{D}} W(t) \text{ in } D[0, 1]$$

and

$$\alpha_n(s) \xrightarrow{\mathcal{D}} B(s) \text{ in } D[0, 1],$$

without describing the space $D[0, 1]$ and without mentioning that, due to the a.s. continuity of the Wiener and Brownian bridge processes $W(\cdot)$ and $B(\cdot)$, this is just like weak convergence on $C[0, 1]$ in the supremum norm $\|\cdot\|$ topology. Similar conventions will be used throughout, in general, when talking about weak convergence of processes of interest on $D[0, 1]$ in the so-called $\|\cdot/q\|$ -metrics. A general methodology for weak convergence of probability measures was established in the fundamental papers of Prohorov (1956) and Skorohod (1956) and developments since are summarized in the books of Billingsley (1968), and Shorack and Wellner (1986).

Rényi (1953) studied the asymptotic distributions of statistics like

$$\sup_{a \leq s \leq b} \alpha_n(s)/s, \quad 0 < a < b \leq 1, \text{ and } \sup_{a \leq s \leq b} \alpha_n(s)/(1-s), \quad 0 \leq a < b < 1,$$

as well as those of their two-sided versions. His idea of introducing these modifications of the classical Kolmogorov-Smirnov statistics was to make them more sensitive on the tails from a hypothesized continuous distribution function $s = F(x)$. In the classical Kolmogorov-Smirnov formulation the difference $|E_n(F(x)) - F(x)| = |F_n(x) - F(x)|$, where $E_n(F(x)) = F_n(x)$ is the empirical distribution function of

a random sample $X_1, \dots, X_n$ with distribution function F , is considered without taking into account the value of $F(x)$ at x . For example, the difference $|F_n(x) - F(x)| = 0.01$ at some point x where $F(x) = 0.5$ (a difference that is only 2% of the value of $F(x)$) is considered just as significant as the same difference at the point x where $F(x) = 0.01$, though the difference now is 100% of the value of $F(x)$. Rényi (1953) therefore proposed that, instead of the absolute deviations $|F_n(x) - F(x)|$, the relative values $|F_n(x) - F(x)|/F(x)$ should be considered. In the same spirit, Anderson and Darling (1952) studied the asymptotics of statistics like

$$\sup_{a \leq s \leq b} \alpha_n(s)/(s(1-s))^{1/2}, \quad 0 < a < b < 1,$$

and their two-sided versions, measuring distances in terms of their standard deviation.

Considerations like these weighted statistics have lead to posing the problem of weak convergence of the empirical processes α_n in $D[0, 1]$ in the so-called $\|\cdot/q\|$ -metrics (weighted supremum metrics) to a Brownian bridge process B . Namely, if we let $\mathcal{Q}^*$ denote the class of positive functions on $(0, 1)$ that are bounded away from zero on $(\delta, 1 - \delta)$ for all $0 < \delta < 1/2$, nondecreasing in a neighbourhood of 0 and nonincreasing in a neighbourhood of 1, then one would want to characterize those functions $q \in \mathcal{Q}^*$ for which we have

$$(0.1) \quad \alpha_n(s)/q(s) \xrightarrow{D} B(s)/q(s).$$

Assuming $q \in \mathcal{Q}^*$ continuous, such a class of functions was first characterized by Chibisov (1964) for $q(s)$ resp. $q(1-s)$ regularly varying at 0 of order $0 < \alpha \leq 1/2$, and then by O'Reilly (1974), assuming only the continuity of $q \in \mathcal{Q}^*$. From their

theorem one can conclude that we have (0.1) if and only if for a standard Wiener process W we have

$$(0.2) \quad \limsup_{t \downarrow 0} |W(t)|/q(t) = \limsup_{t \uparrow 1} |W(t)|/q(1-t) = 0 \quad a.s.$$

This characterization of the weighted weak convergence in (0.1) is somewhat surprising at first sight, and there have been unsuccessful attempts in the literature to reprove this Chibisov-O'Reilly theorem. For comments on these we refer to Section 2, in Csörgő, Csörgő and Horváth (1986).

Csörgő, Csörgő, Horváth and Mason [CsCsHM] (1986) obtained a weighted approximation of the uniform empirical and quantile processes by a sequence of Brownian bridges. For the sake of stating their results, let $U_{1,n} \leq \dots \leq U_{n,n}$ denote the order statistics of the first $n \geq 1$ independent uniform-(0,1) r.v.'s $U_1, U_2, \dots$ in the definition of the uniform empirical process α_n above. For each integer $n \geq 1$, define the uniform empirical quantile function

$$U_n(s) = U_{k,n}, \quad (k-1)/n < s \leq k/n, \quad k = 1, \dots, n,$$

where $U_n(0) = U_{1,n}$, and the uniform quantile process

$$u_n(s) = n^{1/2}(s - U_n(s)), \quad 0 \leq s \leq 1.$$

The essentials, in terms of applications, of this CsCsHM (1986) approximation are contained in the following two theorems, which constitute an extension of the so-called *Hungarian construction*.

Theorem 0.A (Theorem 2.1 of CsCsHM 1986). *A sequence of independent uniform-(0, 1) r.v.'s $U_1, U_2, \dots$, and a sequence of Brownian bridges $B_1, B_2, \dots$ can be constructed on the same probability space so that for all $0 \leq \nu_1 < 1/2$*

$$(0.3) \quad \sup_{1/(n+1) \leq s \leq n/(n+1)} n^{\nu_1} |u_n(s) - B_n(s)| / (s(1-s))^{1/2-\nu_1} = O_P(1).$$

Theorem 0.B (Corollary 4.2.2 of CsCsHM 1986). *On the probability space of Theorem 0.A. for all $0 \leq \nu_2 \leq 1/4$*

$$(0.4) \quad \sup_{0 \leq s \leq 1} n^{\nu_2} |\alpha_n(s) - \bar{B}_n(s)| / (s(1-s))^{1/2-\nu_2} = O_P(1),$$

where $\bar{B}_n(s) = B_n(s)$ for $1/n \leq s \leq 1 - 1/n$ and zero elsewhere.

CsCsHM (1986) used these results to establish anew many previously well-known *best possible* limit theorems for empirical and quantile processes. In particular, the above quoted Chibisov-O'Reilly theorem follows via (0.4) with not necessarily continuous $q \in \mathcal{Q}^*$. The result in (0.4) also indicates why a characterization of (0.1) could possibly be in terms of a condition like (0.2). Other best possible results based on these formulations in (0.3) and (0.4) soon followed.

The CsCsHM (1986) proof for the weighted approximations in (0.3) and (0.4) is based on a refinement of the Csörgő and Révész (1978) inequality for the almost sure approximation of uniform quantile processes by a sequence of Brownian bridges, which in turn is based on the Komlós, Major and Tusnády [KMT] (1976) inequality for the strong approximation of the partial sum process by a Wiener process. The complete proofs of these inequalities entail a substantial amount of technical details.

Csörgő and Horváth (1986) gave mathematically equivalent short and simple proofs of Theorems 0.A and 0.B. Their proof of Theorem 0.A is based directly on the KMT (1976) strong approximation of the partial sum process. They noted also that the KMT (1976) approximation in their proof of Theorem 0.A can be substituted by the Skorohod (1961) embedding, resulting in having (0.3) for all $0 \leq \nu_1 < 1/4$, instead of for all $0 \leq \nu_2 < 1/2$. The just mentioned short proof of Theorem 0.B is based on a Poissonization and on the Csörgő and Révész (1979) strong laws for the increments of a Wiener process used in conjunction with a

Wiener process strong approximation to the Poisson process. The direct, partial sum based proof of Theorem 0.A focussed attention on weighted approximations of partial sums themselves. Studying some changepoint problems like change in the mean of chronologically ordered observations, Csörgő and Horváth (1988c) put

$$Z_n(t) = \begin{cases} (S_{[(n+1)t]} - \frac{[(n+1)t]}{n} S_n) / (n^{1/2} \sigma), & 0 \leq t < 1 \\ 0, & t = 1, \end{cases}$$

where $\sigma^2 = E(X_1 - EX_1)^2$, $S_{[(n+1)t]} = \sum_{i=1}^{[(n+1)t]} X_i$, and with $X_1, X_2, \dots$ i.i.d.r.v.'s which are assumed to have r moments

$$(0.5) \quad E|X_1|^r < \infty \quad \text{for some } r > 2,$$

they concluded that there exists a sequence of Brownian bridges $\{B_n(t), 0 \leq t \leq 1\}$ such that, as $n \rightarrow \infty$

$$(0.6) \quad n^\nu \sup_{1/(n+1) \leq t \leq n/(n+1)} |Z_n(t) - B_n(t)| / (t(1-t))^{1/2-\nu} = O_P(1),$$

where $0 \leq \nu < 1/2 - 1/r$.

This theorem can be deduced from the KMT (1976) and Major (1976a) strong approximations to the partial sum process via using the method of proof of Theorem of Csörgő and Horváth (1986), namely that of Theorem 0.A. Its proof can be extended so that, under the condition (0.5), the following $\|\cdot/q\|$ -metric version of Donsker's theorem holds:

Under the condition (0.5) one can construct a standard Wiener process $\{W(t), 0 \leq t < \infty\}$ so that with $q \in \mathcal{Q}$, as $n \rightarrow \infty$, we have

$$(0.7) \quad \sup_{0 < t \leq 1} |(n\sigma^2)^{-1/2} (S_{[nt]} - W(nt))| / q(t) = o_P(1)$$

if and only if

$$(0.8) \quad \limsup_{t \downarrow 0} |W(t)|/q(t) = 0 \quad \text{a.s.,}$$

where $\mathcal{Q}$ denotes the class of positive functions on $(0, 1]$ which are bounded away from zero on $(\delta, 1]$ for all $\delta \in (0, 1)$ and nondecreasing in a neighbourhood of 0.

With $q \in \mathcal{Q}$, by (0.7) under the condition (0.5), we have also

$$(0.9) \quad (n\sigma^2)^{-1/2} S_{[nt]}/q(t) \xrightarrow{D} W(t)/q(t) \quad \text{in } D[0, 1]$$

if and only if (0.8) holds, i.e., under the condition (0.5), we have the weak convergence of $(n\sigma^2)^{-1/2} S_{[nt]}$ to $W(t)$ on $D[0, 1]$ in $\|\cdot/q\|$ -metric with $q \in \mathcal{Q}$ if and only if (0.8) holds.

A similar theorem holds under the condition (0.5), for the $Z_n(t)$ processes in $\|\cdot/q\|$ -metric if and only if we have (0.2).

The characterization of (0.9) with continuous $q \in \mathcal{Q}$ via (0.8) was proved by O'Reilly (1974) with $r = 3$ in (0.5). Now if r in (0.5), and hence also in (0.6), could be taken to be equal to 2, then (0.7) and (0.9) would also follow, assuming the existence of two moments only, just like in the classical Donsker theorem. Unfortunately, a weighted embedding like that of (0.6) with $\nu = 0$ appears to be impossible with $r = 2$ in (0.5). Consequently, we could not use the method of proof of Theorem of Csörgő and Horváth (1986) to characterize (0.7) in terms of (0.8) when assuming the existence of two moments only, i.e., (0.5) with $r = 2$. Nevertheless, (0.7) and (0.8), as well as (0.8) and (0.9) of course, remain equivalent statements under the assumption of having only two moments (cf. Theorem 1.2.1).

The proof of our Theorem 1.2.1 is based on a strong approximation theorem of Major (1979), approximating the partial sum process of i.i.d.r.v.'s having only

two moments by a Gaussian, but not a Wiener process (cf. Theorem 1.2.A) so that Donsker's functional central limit theorem and Strassen's (1964) strong invariance principle for the law of the iterated logarithm both follow from the same approximation (cf. p.15). What we show in proving Theorem 1.2.1 and Corollary 1.2.1 is that Theorem 1.2.A is strong enough to imply also the optimal weighted version of Donsker's functional central limit theorem.

While at discussing the $\|\cdot/q\|$ -metric nature of Theorem 1.2.1, we would like to note that Theorem 1.2.1(a) is an example of what one means by *weak approximation (invariance principle in probability)*. This notion is conceptually simpler than the notion of *weak convergence (functional central limit theorem)*. Also, in general, the former is a stronger form of the latter as illustrated by Corollary 1.2.1(a), in particular. Of course, once we have a statement like Corollary 1.2.1(a), then we have also its Shorohod-Dudley-Wichura version (cf. Proof of Corollary 1.2.1) which is *like* Theorem 1.2.1(a). The point we are trying to make here, however, is that, using strong and weak approximations of stochastic processes based on various forms of the so-called *Hungarian construction*, one is aiming at first proving a *weak approximation* like that of Theorem 1.2.1(a), say, and then a *weak convergence* statement like Corollary 1.2.1(a) follows immediately. Conversely, the latter combined with the Shorohod-Dudley-Wichura theorem (cf. Theorem A.2.1) yields optimality of $\|\cdot/q\|$ -metric convergence in terms of the weight functions involved.

The optimal class $\mathcal{Q}$ of weight functions q for (0.8) to be true is given in terms of integral tests for the so-called *local functions of a standard Wiener process*. These tests are based on any of the integrals:

$$E(q, c) = \int_0^1 t^{-3/2} q(t) \exp(-ct^{-1} q^2(t)) dt$$

and

$$I(q, c) = \int_0^1 t^{-1} \exp(-ct^{-1}q^2(t))dt, \quad 0 < c < \infty.$$

As presented and compared to the conditions of the classical Kolmogorov test for upper- and lower-class functions of a Wiener process (cf. e.g., Itô and McKean, 1965), our summary in Section 1.1 is based on CsCsHM (1986) and Csörgö, Shao and Szyszkowicz (1991).

The optimal $\|\cdot/q\|$ -metric weak approximation of Theorem 1.2.1(a) implies some immediate L_p -approximations for partial sums of i.i.d.r.v.'s. However, the class of functions $q \in \mathcal{Q}$ that is optimal for $\|\cdot/q\|$ -metric approximations does not yield optimal results for L_p -approximations. Csörgö, Horváth and Shao (1991) established necessary and sufficient conditions for the weak approximation of the uniform empirical and quantile processes by a sequence of Brownian bridges in weighted L_p -distances. Just like in the weighted sup-metric case, their method of proof would work also for partial sum processes of i.i.d.r.v.'s but, again, only under the assumption of them having more than two moments. Assuming the existence of two moments only, in Section 1.4 we prove a result (cf. Theorem 1.4.1) for approximating such partial sums in weighted L_p -distances, $0 < p < \infty$, namely

$$\int_0^1 \left| n^{-1/2}(S_{[nt]} - W(nt)) \right|^p / q(t) dt = o_P(1)$$

for the optimal set of functions $q \in \mathcal{Q}$, characterized by

$$(0.10) \quad \int_0^1 t^{p/2} / q(t) dt < \infty, \quad 0 < p < \infty,$$

the optimal condition for the finiteness of the random integral of a weighted Wiener process $\int_0^1 |W(t)|^p / q(t) dt$, $0 < p < \infty$ (cf. Csörgö, Horváth and Shao, 1991). In addition to this dichotomy theorem and its random integral background (cf. Section 1.3), the proof of our Theorem 1.4.1 is also based on Theorem 1.2.A, due

to assuming the existence of two moments only. For further results and references along these lines we refer to the expository paper Csörgő, Horváth, Shao and Szyszkowicz (1992).

Product measures per se with identical marginals, the so-called i.i.d.-case, are of mathematical interest on their own (as $n \rightarrow \infty$), as well as because of their usefulness in producing theories which, in turn, can be used as building blocks for more complicated asymptotic structures (cf., e.g., Csörgő, Csörgő and Horváth, 1986, and Shorack and Wellner, 1986). A particular area of mathematical and statistical interest where the asymptotics of i.i.d. structures can, and do, play a significant role is that of *contiguity*, a condition that simplifies many arguments in passages to the limit in studies involving independent but not identically distributed observations.

The notion of contiguity was introduced by Le Cam (1960) as a criterion of nearness of sequences of probability distributions. From the point of view of mathematics, this notion is an asymptotic version of the idea of absolute continuity of measures. As to the role of contiguity in statistics, on page 100 of their book, Le Cam and Yang (1990) write: "One obtains a better view of what can happen if one goes slightly off the standard i.i.d. formulation by taking n independent observations from measures $\{Q_{\theta,n}; \theta \in \Theta_n\}$ where $Q_{\theta,n}$ or Θ_n or both depend on n . That gives a context ... where limit theorems "as n tends to infinity" can take many different shapes".

Given a sequence of measurable spaces $\{(\mathcal{X}_n, \mathcal{A}_n)\}$, $n \geq 1$ with probability measures P_n, Q_n on $(\mathcal{X}_n, \mathcal{A}_n)$, we will say that the sequence of probability measures $\{Q_n\}$ is *contiguous* to the sequence $\{P_n\}$ if for every sequence of sets $\{A_n\}$, $A_n \in \mathcal{A}_n$,

$$\lim_{n \rightarrow \infty} P_n(A_n) = 0 \quad \text{implies} \quad \lim_{n \rightarrow \infty} Q_n(A_n) = 0.$$

Essentially all the important results concerning contiguity are already given in Le Cam (1960), which were then used, extended, modified and clarified by many authors. The concept itself has been highly popularized in the area of rank tests through the work of Jaroslav Hájek and his associates. Using mainly the books of Hájek and Šidák (1967), Roussas (1972), Greenwood and Shiryayev (1985), Shorack and Wellner (1986), Le Cam (1986) and Le Cam and Yang (1990), in Section 2.1 we describe the mathematical notion of contiguity and its use in statistics, focusing on Le Cam's third lemma (cf. Lemma 2.1.E), which is the main tool of contiguity theory in applications. In statistical terms, the sequence $\{P_n\}$ may, for example, correspond to the null hypothesis and $\{Q_n\}$ to the alternative, and then using Le Cam's third lemma, convergence properties of statistics of interest can be determined under contiguous alternatives if there is enough information about its behaviour under the null assumption.

In our study throughout, we use the Khmaladze and Parjanadze (1986) parameterization of contiguity. Let $X_1, X_2, \dots$ be independent random variables. We consider the null hypothesis

$$H_0 : X_i, 1 \leq i \leq n, \text{ have the same distribution } F,$$

versus the alternative hypothesis

$$H_1 : X_i, 1 \leq i \leq n, \text{ have the respective distribution functions } F_{in},$$

where we assume that all F_{in} are absolutely continuous with respect to the distribution function F and

$$\left[\frac{dF_{in}}{dF}(F^{-1}(u)) \right]^{\frac{1}{n}} = 1 + \frac{1}{2n^{\frac{1}{2}}} g_n(t, u), \quad \frac{i-1}{n} < t \leq \frac{i}{n},$$

where $F^{-1}(u) = \inf\{x : F(x) \geq u\}$, $0 < u \leq 1$, $F^{-1}(0) = F^{-1}(0+)$ (cf. (2.3.1)).

We assume also that there exists a function $g \in L^2[0, 1]^2$ such that

$$\int_0^1 g(t, u) du = 0 \quad \text{for almost all } t \in [0, 1]$$

and

$$\int_0^1 \int_0^1 \left[g_n(t, u) - g\left(\frac{[nt]}{n}, u\right) \right]^2 dudt \rightarrow 0, \quad n \rightarrow \infty.$$

This parametrization renders the direct products $F_{1n} \times \dots \times F_{nn}$, $n = 1, 2, \dots$, contiguous to the sequence $F \times \dots \times F$ (cf. Section 2.3).

The aim of the game in this context is to transform optimal i.i.d. asymptotics into optimal contiguous asymptotics. For example, having proved the optimal version of Donsker's theorem in $\|\cdot/q\|$ -metrics (cf. Corollary 1.2.1), a reasonable goal then is to show that this weighted convergence continues to hold under contiguous assumptions for the same class of weight functions when assuming the existence of only two moments. We use Le Cam's third lemma via a Lindeberg-type central limit theorem of Oosterhoff and Van Zwet (1975) for the log-likelihood ratio of product probability measures to show that the optimal $\|\cdot/q\|$ -metric weak convergence results of Chapter 1 remain true under contiguous alternatives H_1 (cf. Sections 2.2–2.5). Namely, we obtain

$$n^{-1/2} S_{[nt]}/q(t) \xrightarrow{D} \left(W(t) + \int_0^t \int_0^1 g(s, u) F^{-1}(u) du ds \right) / q(t),$$

if $q \in \mathcal{Q}^*$ and (0.8) holds.

Naturally, proving invariance principles in weighted metrics under contiguous measures is of mathematical interest on its own. Indeed, we are interested in proving limit theorems “as n tends to infinity” in order to gain a better view, to discover some of the many different shapes that can be obtained in the context of

contiguity from the theorems we will have proved in the standard i.i.d. formulation. Our results may, of course, turn out to be also useful in statistics on occasions. For example, in this thesis we try to illustrate some of their use in contiguous changepoint analysis.

Studying changepoint problems, Csörgő and Horváth (1988c) concluded that the tied down partial sum processes $\{Z_n(t), 0 \leq t \leq 1\}$, for which they established (0.6) under the condition (0.5) as quoted above, frequently appear as the natural outcome of some nonparametric, as well as parametric, considerations. Suppose, for example, that we are interested in detecting the possibility of having a change in the mean of chronologically ordered observations of a random sample of size n . Then it is only natural to compare the arithmetic mean of the first k , $1 \leq k \leq n$, of these observations to those of the remaining $(n - k)$ observations for k running through the integers $1, \dots, n - 1$. Thus, we arrive at considering the sequence (in n) of stochastic processes in k

$$\frac{S(k)}{k} - \frac{S(n) - S(k)}{n - k} = n \left(S(k) - \frac{k}{n} S(n) \right) / k(n - k), \quad k = 1, \dots, n - 1,$$

and, consequently, hoping for the convergence in distribution, as $n \rightarrow \infty$, of the statistics

$$n^{1/2} \max_{1 \leq k < n} \left| \frac{S(k)}{k} - \frac{S(n) - S(k)}{n - k} \right| = n^{-1/2} \max_{1 \leq k < n} \frac{|S(k) - kS(n)/n|}{\left(\frac{k}{n} \left(1 - \frac{k}{n} \right) \right)}$$

so that we should be able to reject the null assumption of having no change in the mean if the latter were too large. Unfortunately, these statistics, and even the statistics

$$n^{1/2} \max_{1 \leq k < n} \left(\frac{k}{n} \left(1 - \frac{k}{n} \right) \right)^{1/2} \left| \frac{S(k)}{k} - \frac{S(n) - S(k)}{n - k} \right| = n^{-1/2} \max_{1 \leq k < n} \frac{|S(k) - kS(n)/n|}{\left(\frac{k}{n} \left(1 - \frac{k}{n} \right) \right)^{1/2}},$$

converge in probability to ∞ , as $n \rightarrow \infty$, even if the null assumption of having no change in the mean were true. Hence, in order to have the above test statistics converge in distribution to a nondegenerate random variable, we have to renormalize them somehow. We may, for example, want to introduce weight functions, which are less severe on the tails than $((k/n)(1 - k/n))^{1/2}$ is, and thus we are back to the problem of having to sort out $\|\cdot/q\|$ -metric convergence, say, for partial sums tied down at 0 and 1, like in the definition of $Z_n(t)$ above, this time around for the sake of trying to describe the possibility of having had a change in the mean of, or possibly a shift in, chronologically ordered observations.

In Chapter 3 we summarize various notions of changepoint problems, where we describe also how the results of the later chapters relate to some of these. For a more global view of change analysis we refer to Parzen (1992). When considering the null hypothesis H_0 versus the alternatives of H_1 as above, we note that the parameterization of the latter accommodates the possibility of change in distribution by changing H_1 into

H_1^η : There is an $\eta \in (0, 1)$ such that $P\{X_1 \leq t\} = P\{X_2 \leq t\} = \dots = P\{X_{[n\eta]} \leq t\}$ and $X_{[n\eta]+1}, \dots, X_n$ have distributions $F_{([n\eta]+1)n}, \dots, F_{nn}$ respectively.

The hypothesis H_1^η amounts to quantizing a possible change from i.i.d. sampling to small disturbances afterwards. Our asymptotics provide a view of this phenomenon in terms of various stochastic processes of interest.

When observations are available sequentially and to be kept and examined in their chronological order, then the normalized sequential ranks

$$\xi_k = k^{-1} \sum_{i=1}^k \mathbf{1}\{X_i < X_k\}, \quad k = 1, \dots, n,$$

of the first n of the random variables $X_1, X_2, \dots$ are natural tools for sequential procedures, and they are also convenient to use for sequential studies of finite size samples of such data. A pivotal result of Barndorff-Nielsen (1963) says that if $X_1, \dots, X_n$ are i.i.d.r.v.'s with a continuous distribution, then the sequential ranks $k\xi_k$ are independent, uniformly distributed random variables over $\{0, 1, \dots, k-1\}$, $1 \leq k \leq n$. Consequently, if U_k are independent uniformly distributed r.v.'s on $(0, 1)$, we have $\xi_k \stackrel{D}{=} [kU_k]/k$, $k = 1, 2, \dots$ (cf. Csörgő and Horváth, 1987b). Using this representation, in Chapter 4 we study the asymptotic behaviour of partial sum processes of sequential ranks.

$$V_n(t) = n^{-1/2} \sum_{i=1}^{[nt]} a(\xi_i), \quad 0 \leq t \leq 1,$$

for $a \in L^2(0, 1)$ and such that $\int_0^1 a(u)du = 0$. The process $V_n(t)$ was studied by Khmaladze and Parjanadze (1986) under the sequence of contiguous alternatives H_1 and H_1^η in the context of changepoint problems. For related earlier works on statistics based on sequential ranks we refer to Section 4.0, as well as to Chapter 8, where one of the processes considered is the two-time parameter empirical process of normalized sequential ranks.

In Section 4.1 we study the V_n processes in weighted supremum and L_p -distances under the null assumption that $X_1, X_2, \dots$ are i.i.d.r.v.'s and show that the optimal results of Chapter 1 in these weighted metrics are inherited by these partial sum processes of sequential ranks via the equivalent processes

$$\hat{V}_n(t) = n^{-1/2} \sum_{i=1}^{[nt]} a([iU_i]/i), \quad 0 \leq t \leq 1.$$

In Section 4.2 we show that the weak convergence of the process $V_n(t)/q(t)$ continues to hold under the contiguous alternatives H_1 for the same class of weight

functions as in the case of the null assumption that the X_i are identically distributed. When doing this, one would wish to use Le Cam's third lemma again, like in Chapter 2. However, while $\tilde{V}_n$ is equivalent to V_n in distribution, the joint distribution of $\tilde{V}_n$ and the log-likelihood process L_n (cf. (2.3.4)), i.e. that of $(\tilde{V}_n/q, L_n)$, does not give the right covariance function we would like to have, namely that of $(V_n/q, L_n)$. Combining results of Khmaladze and Parjanadze (1986), where they prove the weak convergence of (V_n, L_n) to the appropriate two-dimensional Gaussian process, with our results of Section 4.1, enables us to obtain the correct asymptotic distribution of the $V_n(t)/q(t)$ process under the desired contiguous alternatives. As a matter of fact, assuming H_0 , our $\|\cdot/q\|$ -metric approximation of the $V_n(t)$ process yields

$$n^{-1/2} \sum_{i=1}^{[nt]} a(\xi_i)/q(t) \xrightarrow{D} W(t)/q(t),$$

where $q \in \mathcal{Q}^*$ and (0.8) holds (cf. Corollary 4.1.1), while under the contiguous alternatives H_1 we obtain (cf. Theorem 4.2.1)

$$n^{-1/2} \sum_{i=1}^{[nt]} a(\xi_i)/q(t) \xrightarrow{D} \left(W(t) - \int_0^t \frac{Q(s)}{s} ds + Q(t) \right) / q(t)$$

for the same class of weight functions, where

$$Q(t) = \int_0^t \int_0^1 g(s, u) a(u) du ds.$$

We close Section 4.2 by describing the form of the limiting distribution in the case when change, as parameterized by the alternative H_1^q , has occurred.

In Chapter 5 we study the sequence of tied down partial sum processes $Z_n(t)$ as in (0.6). Assuming that $E|X_1|^r < \infty$ for some $r > 2$, Csörgő and Horváth (1988c) describe the asymptotic theory of $Z_n(t)$ in $\|\cdot/q\|$ -metrics under H_0 , namely

i.i.d. setup. Using the optimal $\|\cdot/q\|$ -metric results of Chapter 1 as building blocks for the tied down $\{Z_n(t), 0 \leq t \leq 1\}$ process, in Section 5.2 we improve their results under H_0 , assuming only two moments for X_1 . In Section 5.4 we extend them under the contiguous alternatives H_1 and H_1^q . The optimal weighted L_p -approximations for $Z_n(t)$ are established in Section 5.3. Clearly, under H_0 , the $Z_n(t)$ processes are to be approximated by a sequence of Brownian bridges $B_n(t)$, while under the alternatives H_1 we obtain

$$Z_n(t)/q(t) \xrightarrow{D} \left(B(t) + \int_0^t \int_0^1 g(s,u) F^{-1}(u) du ds - t \int_0^1 \int_0^1 g(s,u) F^{-1}(u) du ds \right) / q(t),$$

where $q \in \mathcal{Q}^*$ and (0.2) holds (cf. Theorem 5.4.1).

The tied down partial sum processes $\{Z_n(t), 0 \leq t \leq 1\}$ can be viewed as ready-made for detecting a change in the mean in general, or possibly a shift in the magnitude, of chronologically ordered observations. There are, of course, many different ways of manifesting change in distribution of various stochastic processes of interest.

With this in mind, Csörgő and Horváth (1988a) study tests based on processes of U -statistics. They consider the sequence (in n) of stochastic processes in k

$$(0.11) \quad Z_k = \sum_{1 \leq i \leq k} \sum_{k+1 \leq j \leq n} h(X_i, X_j), \quad 1 \leq k < n,$$

where $h(x, y)$ is a symmetric, or antisymmetric function, and study its limiting distribution in unweighted and weighted sup-norm metrics. They express Z_k as a linear combination of U -statistics and hence studying Z_k can be based on the Hoeffding (1948a, 1961), projection of a U -statistic on the basic observations (cf. Chapter 5 of Serfling 1980). The class of U -statistics was introduced by Hoeffding (1948a) as a generalization of the sample mean, that is, as a generalization of the

notion of forming an average, and is obtained by averaging a kernel h symmetrically over the observations

$$U_n = U(X_1, \dots, X_n) = \frac{1}{\binom{n}{m}} \sum_c h(X_{i_1}, \dots, X_{i_m}),$$

where $\sum_c$ denotes summation over the $\binom{n}{m}$ combinations of m distinct elements $\{i_1, \dots, i_m\}$ from $\{1, \dots, n\}$. Typical choices of h are $xy, (x - y)^2/2$ (sample variance), $|x - y|$ (Gini's mean difference), $\text{sign}(x + y)$ (Wilcoxon's one-sample statistics). Assuming

$$(0.12) \quad E|h(X_1, X_2)|^r < \infty \quad \text{for some } r > 2,$$

Csörgő and Horváth (1988a) establish weighted approximations for $\mathcal{Z}_k$ in supremum norm for the optimal class of weight functions $q \in \mathcal{Q}^*$ with (0.2), for both symmetric and antisymmetric kernel functions $h(\cdot, \cdot)$ in $\mathcal{Z}_k$, via proving first a type of approximation as in (0.6) for the $\mathcal{Z}_k$ process. This approach does not work for $r = 2$ in (0.12). Their symmetric kernel proof is based on Theorem 1 of Hall (1979) and that of the antisymmetric kernel case on Theorem 2.1 of Janson and Wichura (1983). Both of these theorems result in sup-norm approximations of $\mathcal{Z}_k$ by their appropriate linear combinations of U -statistics projections, on assuming (0.12) for $r = 2$ only. Noting that these approximations remain true in the stronger $\|\cdot/q\|$ -metric as well for $q \in \mathcal{Q}^*$ and such that $\lim_{t \downarrow 0} q(t)/t^{1/2} = \lim_{t \uparrow 1} q(t)/(1 - t)^{1/2} = \infty$, and since all $q \in \mathcal{Q}^*$ with (0.2) satisfy the latter condition, it is quite natural to hope that the method of Chapter 1 of approximating partial sums, having two moments only, should work also for $\mathcal{Z}_k$ with $r = 2$ in (0.12). We show in Sections 6.1 and 6.2 that this is indeed so, for both types of kernels, in weighted supremum norm and L_p -metrics for the optimal class of weight functions. Our best possible $\|\cdot/q\|$ -metric weak convergence results are shown to remain true under

contiguous measures in Section 6.3. For some preliminary results along these lines we refer to Szyszkowicz (1991a). We note that in the case of an antisymmetric kernel in (0.11), the asymptotic results for the $\mathcal{Z}_{[nt]}$ process coincide with those for the already considered tied down partial sum process $Z_n(t)$. In fact, taking $h(x, y) = x - y$ gives $\mathcal{Z}_k$ equal to $n(S_k - \frac{k}{n}S_n)$, or, equivalently, to

$$k(n-k) \left(\frac{1}{k} \sum_{i=1}^k X_i - \frac{1}{n-k} \sum_{i=k+1}^n X_i \right),$$

i.e. the comparison of two means (say, before and after a possible change) is a special case of the process $\mathcal{Z}_k$. Even though results for the particular case follow from the more general one, we preferred to prove directly all results for the $Z_n(t)$ process in Chapter 5, because they actually form the bases of our proofs when studying the more general $\mathcal{Z}_{[nt]}$ process with antisymmetric kernels.

Another example of an antisymmetric kernel which has gained special attention in the literature is $h(x, y) = \text{sign}(x - y)$ (cumulative rank tests) (cf. e.g. Pettitt, 1979). Pettitt's i.i.d. asymptotics remain true, of course, in $\|\cdot/q\|$ -metrics and under the changepoint alternative H_1^η we obtain, as $n \rightarrow \infty$,

$$\sup_{0 < t < 1} \left(12^{1/2} n^{-3/2} \sum_{i=1}^{[(n+1)t]} \sum_{j=[(n+1)t]+1}^n \text{sign}(X_i - X_j) \right) / q(t) \\ \xrightarrow{D} \sup_{0 < t < 1} \left(B(t) + (12)^{1/2} (-t(1-\lambda)\mathbf{1}\{t < \lambda\} - \lambda(1-t)\mathbf{1}\{t \geq \lambda\}) \int_0^1 u g(u) du \right) / q(t),$$

where $B(t)$ is a Brownian bridge and $q \in \mathcal{Q}^*$ is such that (0.2) holds.

In general, under H_0 , the $\mathcal{Z}_{[nt]}$, $0 \leq t < 1$, processes with antisymmetric kernels converge to a Brownian bridge B , while with symmetric kernels they converge to the Gaussian process

$$\Gamma(t) = (1-t)W(t) + t\{W(1) - W(t)\}, \quad 0 \leq t \leq 1,$$

where W is a standard Wiener process.

The idea of comparing the means before a possible change to those after, appear also in studying some archaeological sites.

When looking for a possible changepoint in the Land's End data, Kendall and Kendall (1980) considered test-statistics based on the stochastic processes

$$(0.13) \quad X_n(t) = \left(\frac{N(nt)}{nt} - \frac{N(n) - N(nt)}{n - nt} \right) \left(\frac{nt(n - nt)}{N(n)} \right)^{\frac{1}{2}}, \quad 0 < t < 1,$$

with $X_n(0) = X_n(1) = 0$, called by them *pontograms*, where $\{N(x), x \geq 0\}$ is a possibly nonhomogeneous, also called nonstationary, Poisson process with intensity function $\lambda(x)$, $x \geq 0$. We let

$$K_n(t) = n^{-\frac{1}{2}}(N(nt) - tN(n)), \quad 0 \leq t \leq 1$$

and note that

$$\left\{ \left(\frac{n}{N(n)} \right)^{\frac{1}{2}} \frac{K_n(t)}{(t(1-t))^{\frac{1}{2}}}, \quad 0 < t < 1, \quad n \geq 1 \right\} = \left\{ X_n(t), \quad 0 < t < 1, \quad n \geq 1 \right\}.$$

Assuming that $\lambda(x)$ has a constant, but unknown value $\lambda > 0$, Kendall and Kendall (1980) proved that, for all $0 < \epsilon_1 < 1 - \epsilon_2 < 1$, we have

$$\sup_{\epsilon_1 \leq t \leq 1 - \epsilon_2} \left(\frac{n}{N(n)} \right)^{\frac{1}{2}} \frac{K_n(t)}{(t(1-t))^{\frac{1}{2}}} \xrightarrow{D} \sup_{0 \leq s \leq \frac{1}{2} \log[(1 - \epsilon_2)/\epsilon_1][(1 - \epsilon_1)/\epsilon_2]} |V(s)|,$$

where $\{V(s), -\infty < s < \infty\}$ is an Ornstein-Uhlenbeck process with mean zero and covariance function $\exp(-|t - s|)$ and, on account of the problem at hand, called attention to the desirability of replacing ϵ_1 and ϵ_2 be zero, or by sequences $\epsilon_i(n) \rightarrow 0$ as $n \rightarrow \infty$, $i = 1, 2$. This problem was solved by Csörgő and Horváth (1987c) for various suprema and L_2 -functionals of pontograms (cf. Theorem 7.0.A and Proposition 7.2.2 with $p = 2$).

Along the lines of Szyszkowicz (1992d), in Chapter 7 we study weighted versions of the Kendall-Kendall (1980) pontograms of (0.13), which are of the form

$$(0.14) \quad \left(\frac{n}{N(n)} \right)^{1/2} K_n(t)/q(t) = \left(\frac{n}{N(n)} \right)^{1/2} \frac{K_n(t)}{(t(1-t))^{1/2}}/g(t), \quad 0 < t < 1,$$

where $\lim_{t \downarrow 0} g(t) = \lim_{t \uparrow 1} g(t) = \infty$, due to having the $\sup_{0 < t < 1}$ -functional of the Kendall-Kendall pontograms of (0.13) converge to ∞ in probability, as $n \rightarrow \infty$, even if $\lambda(x) \equiv \lambda > 0$ for all $x \geq 0$.

Assuming first the null assumption

$$H : \lambda(x) = \lambda > 0, \quad x \geq 0,$$

in Section 7.1 we give necessary and sufficient conditions for weight functions q for approximating the Poisson bridges $K_n(t)$ by a sequence of Brownian bridges $B_n(t)$ in $\|\cdot/q\|$ -metrics and characterize also the convergence in distribution of $\sup_{0 < t < 1} |K_n(t)|/q(t)$ for a larger class of functions (cf. Propositions 7.1.1 and 7.1.2). The optimal asymptotics of weighted L_p -approximations, $0 < p < \infty$, of $K_n(t)$ are studied in Section 7.2, where we extend also some L_2 -functional results of Csörgő and Horváth (1987c) for the Kendall-Kendall (1980) pontograms to L_p , $1 \leq p < \infty$, using the method of proof of similar results of Csörgő and Horváth (1988b, 1992) on weighted uniform and quantile processes.

As an alternative to the null assumption H , or simply on its own, it is natural to consider contiguous changes in the intensity function $\lambda(\cdot)$ of the underlying Poisson process $N(\cdot)$ of the $K_n(t)$ process. In Section 7.3 we study the asymptotic behaviour of the weighted Kendall-Kendall pontograms $(n/N(n))^{1/2} K_n(t)/q(t)$ of (0.14) under the sequence of contiguous alternatives

$$A : \lambda(x) = \lambda + \xi([x]/n)/n^{1/2}, \quad 0 \leq x \leq n, \quad n \geq 1,$$

where $\xi^2(t)$ is a Riemann integrable function.

In particular, A accommodates the changepoint alternative

A^η : there exists $\eta \in (0, 1)$ such that

$$\lambda(x) = \lambda \quad \text{for } 0 \leq x \leq n\eta$$

$$\lambda(x) = \lambda + \rho/n^{1/2} \quad \text{for } n\eta \leq x \leq n, \quad n \geq 1,$$

by taking $\xi(t) = \rho 1\{t \geq \eta\}$, $\rho \neq 0$. We note that $\rho > 0$ ($\rho < 0$) corresponds to a higher (lower) value for the intensity function $\lambda(x)$ of the Poisson process after the changepoint η of this parameterization, and one may want to pick an appropriate value for it, for a particular problem at hand.

We conclude in Section 7.3 that the weak convergence results in $\|\cdot/q\|$ -metrics proved for the Kendall-Kendall pontograms under the null assumption H can be transformed into statements under the contiguous alternatives A and A^η . In particular, under the alternatives A we obtain (cf. Corollary 7.3.1)

$$\left(\frac{n}{N(n)}\right)^{1/2} K_n(t)/q(t) \xrightarrow{D} \left(B(t) + \lambda^{-1/2} \left(\int_0^t \xi(s) ds - t \int_0^1 \xi(s) ds \right)\right)/q(t),$$

where $\{B(t), 0 \leq t \leq 1\}$ is a Brownian bridge and $q \in \mathcal{Q}^*$ is such that (0.8) holds. For further results along these lines for weighted compound Poisson bridges, we refer to Szyszkowicz (1992e).

Going back to the classical result of (0.1), that is having the weighted weak convergence of the uniform empirical process α_n in $\|\cdot/q\|$ -metrics for $q \in \mathcal{Q}^*$ if and only if (0.2) holds, in the light of our discussions so far and in that of the historical role this result has played in the development of the area of weighted approximations of stochastic processes, it is only natural to think about transforming also this optimal i.i.d. asymptotics into a contiguous one. This is indeed possible, for

example under the contiguous alternatives H_1 . Namely, for $q \in \mathcal{Q}^*$ and such that (0.8) holds, we obtain (cf. Theorem 8.9.1)

$$\alpha_n(s)/q(s) \xrightarrow{\mathcal{D}} \left(B(s) + \int_0^1 \int_0^s g(\tau, u) du d\tau \right) / q(s),$$

where $\{B(s), 0 \leq s \leq 1\}$ is a Brownian bridge. Preceding this consideration, in Chapter 8 we study three two-time parameter empirical processes on their own, as well as in terms of their role in handling change in distribution problems in weighted supremum-norm metrics.

The general idea of studying change in the distribution of a random sequence of observations via comparing the mean of the first k of them to that of the remaining $(n - k)$ of them appears to be even more appropriate when expressed in terms of comparing the empirical distribution function of the first k observations to that of the last $(n - k)$ observations. For example, Picard (1985), and Deshayes and Picard (1986) consider the statistics

$$\begin{aligned} (0.15) \quad & \sup_{1 \leq k < n} \sup_{x \in \mathbb{R}} n^{1/2} \left| \frac{1}{k} \sum_{i=1}^k \mathbf{1}\{X_i \leq x\} - \frac{1}{n-k} \sum_{i=k+1}^n \mathbf{1}\{X_i \leq x\} \right| \\ &= \sup_{1 \leq k < n} \sup_{x \in \mathbb{R}} \left| \sum_{i=1}^k \mathbf{1}\{X_i \leq x\} - \frac{k}{n} \sum_{i=1}^n \mathbf{1}\{X_i \leq x\} \right| / n^{1/2} \left(\frac{k}{n} \left(1 - \frac{k}{n} \right) \right). \end{aligned}$$

Since this sequence of random variables, and even the one where $\frac{k}{n} \left(1 - \frac{k}{n} \right)$ is replaced by $\left(\frac{k}{n} \left(1 - \frac{k}{n} \right) \right)^{1/2}$, converges to ∞ in probability as $n \rightarrow \infty$ already under H_0 , just like in case of similar partial sum processes, we are led to considering weight functions which are less severe on the tails than $(k/n(1 - k/n))^{1/2}$ is. This, in turn, leads to studying the “bridge” type (tied down to $t = 0$ and $t = 1$) two-time parameter empirical processes

$$(0.16) \quad (\alpha_n(s, t) - t\alpha_n(s, 1))/q(t), \quad 0 < s, t < 1,$$

for the sake of trying to see what kind of weight functions are possible for the otherwise degenerate statistics of (0.15). Here

$$\alpha_n(s, t) = n^{-1/2} \sum_{i=1}^{[nt]} (1\{F(X_i) \leq s\} - s), \quad 0 \leq s, t \leq 1,$$

and F is assumed to be a continuous distribution function. In Section 8.2 we study approximations of $\alpha_n(s, t)/q(t)$ in probability for various classes of weight functions q under H_0 by a Kiefer process (cf. Kiefer 1972, and Csörgő and Révész 1975, 1981) and then we use our results to prove the same kind of approximations for the weighted tied down processes of (0.16) in Section 8.3. In these studies the strong approximation result of Komlós, Major and Tusnády (1975) (cf. also Bonvalot and Castelle, 1991) of the two-time parameter empirical process $\alpha_n(s, t)$ by a Kiefer process $K(s, nt)$ plays a crucial role. Assuming H_0 we obtain (cf. Theorem 8.2.1)

$$(0.17) \quad \sup_{0 < t \leq 1} \sup_{0 \leq s \leq 1} \left| \alpha_n(s, t) - n^{-1/2} K(s, nt) \right| / q(t) = o_P(1)$$

for any q which is positive on $(0, 1]$ and such that

$$\lim_{t \downarrow 0} (t \log \log 1/t)^{1/2} / q(t) = 0.$$

As to the class of weight functions q , in addition to our own conditions on directly comparing them near zero and one to the function $(t \log \log 1/t)^{1/2}$, we call attention to other possibilities as well, by citing and inferring from the works of Deshayes and Picard (1983) (square-integrable weight functions), and Adler and Brown (1986). The latter actually reaches back to Chung (1949) and Kiefer (1961), resulting in the characterization of the class of weight functions for having (0.17) in terms of the integral

$$\mathcal{J}(q, c) = \int_0^1 t^{-1} q^2(t) \exp(-ct^{-1} q^2(t)) dt, \quad c > 0,$$

assuming that $q \in C(0, 1]$ and $q(t)/t^{1/2}$ is nonincreasing near zero.

The approximating process of that of (0.16) is in terms of the Kiefer process $K(s, nt)$, given by

$$\begin{aligned} & \left\{ n^{1/2}(K(s, nt) - tK(s, n)), 0 \leq s \leq 1, 0 \leq t \leq 1 \right\} \\ & \stackrel{D}{=} \{K(s, t) - tK(s, 1), 0 \leq s \leq 1, 0 \leq t \leq 1\} \text{ for each } n \geq 1 \\ & \stackrel{D}{=} \{\Gamma(s, t), 0 \leq s \leq 1, 0 \leq t \leq 1\}, \end{aligned}$$

that is to say $\Gamma(s, t)$ is a mean zero Gaussian process with covariance function

$$E\Gamma(s_1, t_1)\Gamma(s_2, t_2) = (s_1 \wedge s_2 - s_1 s_2)(t_1 \wedge t_2 - t_1 t_2),$$

which Parzen (1992) calls a pinned Brownian sheet (two-time parameter Brownian bridge). This limiting Gaussian process Γ also appeared in Hoeffding (1948b) and Blum, Kiefer and Rosenblatt (1961) in the context of testing for independence, as well as in Csörgö and Horváth (1987a), proposing nonparametric procedures for changepoint problems based on the empirical process of full sample ranks.

In Section 8.4 we consider the two-time parameter empirical process of normalized ranks $R_{1n}, \dots, R_{nn}$

$$\beta_n(s, t) = n^{-1/2} \sum_{i=1}^{[nt]} (1\{R_{in} \leq s\} - \frac{[ns]}{n}), \quad 0 \leq s, t \leq 1,$$

where

$$R_{kn} = n^{-1} \sum_{i=1}^n 1\{X_i \leq X_k\}, \quad k = 1, \dots, n.$$

Noting first that we have

$$\beta_n(s, t) = \alpha_n(U_n(s), t) - \frac{[nt]}{n} \alpha_n(U_n(s), 1),$$

where $U_n(\cdot)$ is the empirical quantile function of (8.4.1), which shows that $\beta_n(s, t)$ has, as it is, a two-time parameter bridge type structure, though random in s , we

then prove approximations for the $\beta_n(s, t)/q(t)$ process which are like those of the $(\alpha_n(s, t) - t\alpha_n(s, 1))/q(t)$ process.

The two-time parameter empirical process $\gamma_n(s, t)$ of sequential ranks $\xi_i = i^{-1} \sum_{j=1}^i 1\{X_j < X_i\}$, $i = 1, \dots, n$, defined by

$$\gamma_n(s, t) = n^{-1/2} \sum_{i=1}^{[nt]} (1\{\xi_i \leq s\} - \frac{[is]}{s}), \quad 0 \leq s, t \leq 1,$$

are ready-made for analyzing chronological observations. Csörgő and Horváth (1987b) use these empirical processes to construct sequential procedures for change detection. In Section 8.5 we study approximations of the weighted versions $\gamma_n(s, t)/q(t)$ under H_0 with continuous F , and those of their tied down forms $(\gamma_n(s, t) - \gamma_n(s, 1))/q(t)$ in Section 8.6, for the sake of constructing tools, statistics for sequential studies of chronologically ordered finite samples of size n , as $n \rightarrow \infty$. In the light of the already mentioned representation $\xi_k \stackrel{D}{=} [kU_k]/k$, $k = 1, 2, \dots$, of sequential ranks in terms of independent uniform- $(0, 1)$ random variables $U_1, U_2, \dots$, it is reasonable that under H_0 we get the same Gaussian weighted approximations like in case of the processes in (0.16). All in all, our considerations and results in Sections 8.2-8.6 amount to a unified treatment under H_0 of the two-time parameter empirical processes $\alpha_n(s, t)$, $\beta_n(s, t)$ and $\gamma_n(s, t)$ of the data, their ranks and sequential ranks, respectively, in $\sup_{0 < t < 1} \sup_{0 \leq s \leq 1} |\cdot, \cdot/q(t)|$ -metrics. The same can be said about our studying them under the contiguous alternatives H_1 and H_1^n in Sections 8.7 and 8.8 respectively, where we succeed in transforming our i.i.d. based asymptotics into contiguous asymptotics with the same class of weight functions. For example, assuming that q is positive on $(0, 1)$ and such that

$$\lim_{t \downarrow 0} \left(t(1-t) \log \log 1/(t(1-t)) \right)^{1/2} / q(t) = 0$$

and

$$\lim_{t \uparrow 1} \left(t(1-t) \log \log 1/(t(1-t)) \right)^{1/2} / q(t) = 0,$$

under H_1 we obtain (cf. Theorem 8.7.2)

$$(\alpha_n(s, t) - t\alpha_n(s, 1))/q(t) \xrightarrow{\mathcal{D}} (\Gamma(s, t) + d(s, t))/q(t),$$

where

$$d(s, t) = \int_0^t \int_0^s g(\tau, u) du d\tau - t \int_0^1 \int_0^s g(\tau, u) du d\tau.$$

The $\|\cdot/q\|$ -metric convergence of the $\beta_n(s, t)$ process is similar. It inherits the limiting distributions from $\alpha_n(s, t) - t\alpha_n(s, 1)$. However this is not so in the case of the empirical processes of sequential ranks under H_1 (cf. Theorems 8.7.1 and 8.7.2). In the case of contiguous alternatives, our proofs with $q \equiv 1$ constitute an alternative proof of Theorem 1 of Pardzhanadze and Khamaladze (1986).

Each chapter of this dissertation has an introduction. Hence it is inevitable that certain elements of this global introduction will reappear time and again, here and there, in the text. Our reason for writing this general introduction was to provide a somewhat more comprehensive view, in the retrospect, of what is accomplished in this thesis, hopefully well enough, as a result of our studies and research of random phenomena in the last three years.

CHAPTER 1

Weighted approximations of partial sums of i.i.d.r.v.'s

1.0. Introduction

Given a probability space $(\Omega, \mathcal{A}, P)$, let $X_1, X_2, \dots$ be independent, identically distributed random variables with partial sums $S_n = X_1 + \dots + X_n$. Donsker's theorem says that, assuming the existence of two moments, say $EX_1 = 0$ and $EX_1^2 = \sigma^2$, $0 < \sigma^2 < \infty$, the sequence of properly normalized partial sum processes with values in $D[0, 1]$ converges in distribution to a standard Brownian motion on $[0, 1]$, i.e.,

$$\frac{1}{\sigma\sqrt{n}} S_{[nt]} \xrightarrow{\mathcal{D}} W(t) \quad \text{in } D[0, 1],$$

where $[x]$ stands for the integer part of the real number x . The question arises under what conditions does weak convergence continue to hold for weighted partial sum processes. Namely, one would want to study the problem of

$$\frac{1}{\sigma\sqrt{n}} S_{[nt]}/q(t) \xrightarrow{\mathcal{D}} W(t)/q(t),$$

where $q(t)$ is a nonnegative function on $(0, 1]$ approaching zero as $t \rightarrow 0$.

The motivation for studying weighted partial sum processes comes from earlier studies of Rényi (1953), Chibisov (1964), Pyke and Shorack (1968), O'Reilly (1974) and others, concerning the asymptotic behaviour of weighted empirical and quantile processes. Based on a refinement of the Csörgő and Révész (1978) inequality for the strong approximation of the uniform quantile processes by a sequence of Brownian bridges, which in turn is based on the Komlós, Major and Tusnády [KMT] (1976) inequality for the strong approximation of the partial sum process by a Wiener process, Csörgő, Csörgő, Horváth and Mason [CsCsHM] (1986) gave

complete characterizations for describing the asymptotics of the weighted uniform quantile process in terms of optimal classes of weight functions for the supremum norm. They also established similar results for the uniform empirical process. It is clear from their proof, and they also underline this strongly, that the optimal class of weight functions for the weak convergence of these weighted processes is determined not by these empirical processes themselves, but by the weighted tail behaviour of the limiting Brownian bridge process. Hence the motivation for a detailed study of the local behaviour of the Brownian bridge and Brownian motion processes with weight functions.

In the light of the mentioned results it was only natural for Csörgő and Horváth (1988c) to extend the KMT (1976) approximation of partial sums by a Wiener process to their weighted approximation in terms of the optimal class of weight functions as in CsCsHM (1986), via assuming the existence of more than two moments, just like in the corresponding Komlós, Major and Tusnády theorems. Such a method of proof, however, cannot be carried out under the initial assumption of two moments only for having Donsker's theorem. Earlier O'Reilly (1974) proved weak convergence of weighted partial sum processes assuming the existence of three moments. He asserted also, without proof, that the third moment condition could be dropped and the theorem would continue to hold.

In Section 2 we establish weighted approximations of partial sum processes, assuming the existence of two moments only (cf. also Szyszkowicz 1991d, 1992a). The obtained result is optimal in terms of moment conditions and the class of weight functions. Since the limiting process is a standard Brownian motion, Section 1 is devoted to studying its local behavior. Using results of CsCsHM (1986)

and Csörgő, Shao and Szyszkowicz (1991), we sort out necessary and sufficient conditions for

$$\limsup_{t \downarrow 0} |W(t)|/q(t) < \infty \quad \text{a.s.}$$

and

$$\limsup_{t \downarrow 0} |W(t)|/q(t) = \beta \quad \text{a.s.} \quad 0 \leq \beta < \infty,$$

versus the conditions of the classical Kolmogorov test for the upper- and lower-class functions of a Wiener process.

In a similar vein, in Section 3 we present the optimal condition for the finiteness of the random integral of a weighted Wiener process $\int_0^1 |W(t)|^p/q(t)dt$, where $0 < p < \infty$. In the case of $p = 2$, the result is due to Shepp (1964). It was reproved for the general case by Csörgő, Horváth and Shao (1991), where they also find a necessary and sufficient condition for the weak convergence of the uniform empirical and quantile processes to Brownian bridge in weighted L_p -distances. As in the sup-metric case, their method of proof would work also for partial sum processes, but only under the assumption of more than two moments being finite. Assuming the existence of two moments only, in Section 4 we prove optimal results for the approximation of partial sum processes in weighted L_p -distances, $0 < p < \infty$ (cf. also Szyszkowicz 1992b,c).

1.1. Local functions of a Wiener process

A function q defined on $(0, 1]$ will be called *positive* if $\inf_{\delta \leq s \leq 1} q(s) > 0$ for all $0 < \delta < 1$.

Let $\mathcal{Q}$ be the class of those positive functions on $(0, 1]$ which are non-decreasing in a neighbourhood of zero.

Given a standard Wiener process $\{W(t), t \geq 0\}$ and a function $q \in \mathcal{Q}$, by Blumenthal's 0 – 1 law (cf. Itô and McKean, 1965), or by the direct 0 – 1 law for Brownian motion as in Doob (1984) (cf. Theorem A.3.1 in our Appendix), we have

$$(1.1.1) \quad P\left\{\limsup_{t \downarrow 0} |W(t)|/q(t) < \infty\right\} = 0 \text{ or } 1.$$

A function $q \in \mathcal{Q}$ will be called a *local function of a standard Wiener process* $\{W(t), t \geq 0\}$ if

$$(1.1.2) \quad \limsup_{t \downarrow 0} |W(t)|/q(t) < \infty \quad \text{a.s.}$$

An application of any of the just mentioned 0 – 1 laws shows that (1.1.2) holds true if and only if there exists a constant $0 \leq \beta < \infty$ such that

$$(1.1.3) \quad \limsup_{t \downarrow 0} |W(t)|/q(t) = \beta \quad \text{a.s.}$$

i.e., a positive function q on $(0, 1]$ which is non-decreasing in a neighbourhood of zero, is a *local function of a standard Brownian motion starting at zero* if and only if (1.1.3) obtains (cf Proposition A.3.1).

A *local function* will be called a *Chibisov-O'Reilly local function* if $\beta = 0$ in (1.1.3).

Given a function $r \in \mathcal{Q}$, by the 0 – 1 law we have

$$P\{|W(t)| < r(t), t \downarrow 0\} = 0 \text{ or } 1.$$

A function $r \in \mathcal{Q}$ is said to belong to the *local upper-class* if this probability is 1 and to the *local lower-class* otherwise.

We note in passing that excluding from $\mathcal{Q}$ all positive functions which are not non-decreasing near zero, i.e. the decreasing ones, constitutes no loss of generality in the above definitions and statements when $t \downarrow 0$, since we start W at zero with probability one.

As in Itô and McKean (1965, p.33), Kolmogorov's test states that if r is a positive function in $C(0,1]$, non-decreasing near zero and if $r(t)/t^{1/2}$ is non-increasing for small $t > 0$, then r belongs to the local upper or to the local lower-class according as

$$\int_0^1 t^{-3/2} r(t) \exp(-r^2(t)/(2t)) dt$$

converges or diverges.

The statement (1.1.3) with $\beta > 0$ is equivalent to saying that for any $\varepsilon > 0$ we have

$$(1.1.4) \quad P\{|W(t)| < (\beta + \varepsilon)q(t), t \downarrow 0\} = 1$$

and

$$(1.1.5) \quad P\{|W(t)| < (\beta - \varepsilon)q(t), t \downarrow 0\} = 0,$$

i.e., $q \in \mathcal{Q}$ is a local function for W if and only if $(\beta + \varepsilon)q$ belongs to its local upper-class and $(\beta - \varepsilon)q$ belongs to its local lower-class.

We introduce the following integrals:

$$(1.1.6) \quad E(q, c) := \int_0^1 t^{-3/2} q(t) \exp(-ct^{-1}q^2(t)) dt$$

and

$$(1.1.7) \quad I(q, c) := \int_0^1 t^{-1} \exp(-ct^{-1}q^2(t)) dt$$

with some constant $0 < c < \infty$.

The integral $E(q, c)$ appeared in the works of Kolmogorov, Petrovski, Erdős and Feller. For details we refer to Itô and McKean (1965, Section 1.8).

The integral $I(q, c)$ appeared in the works of Chibisov (1964) and O'Reilly (1974).

For further comments on these two integrals, as well as for the proof of the next three theorems we refer to CsCsHM (1986). We have (cf. Proposition 3.1, and Theorems 3.3 and 3.4, respectively, of CsCsHM, 1986):

Theorem 1.1.A. (i) *Whenever the integral $I(q, c) < \infty$ for $q \in \mathcal{Q}$, then $E(q, c + \varepsilon) < \infty$ for every $\varepsilon > 0$ and $q(t)/t^{1/2} \rightarrow \infty$ as $t \downarrow 0$.*

(ii) *Whenever $E(q, c) < \infty$ and $q(t)/t^{1/2} \rightarrow \infty$ as $t \downarrow 0$ for $q \in \mathcal{Q}$, then $I(q, c) < \infty$.*

Theorem 1.1.B. *A function $q \in \mathcal{Q}$ is a local function of a standard Wiener process starting at zero if and only if the integral $I(q, c) < \infty$ for some $c > 0$ or, equivalently, if and only if the integral $E(q, c) < \infty$ for some $c > 0$ and $\lim_{t \downarrow 0} q(t)/t^{1/2} = \infty$.*

Theorem 1.1.C. *A function $q \in \mathcal{Q}$ is a Chibisov-O'Reilly local function if and only if the integral $I(q, c) < \infty$ for all $c > 0$ or, equivalently, if and only if the integral $E(q, c) < \infty$ for all $c > 0$ and $\lim_{t \downarrow 0} q(t)/t^{1/2} = \infty$.*

We note that in O'Reilly (1974), as well as in CsCsHM (1986), a Chibisov-O'Reilly local function q on $(0, 1)$ is defined by requiring $q(t)$ and $q(1 - t)$ to be Chibisov-O'Reilly local functions on $(0, 1/2]$. The definition of a local function is extended to $(0, 1)$ in a similar way in CsCsHM (1986). This convention enables

one to talk about *local* functions of a Brownian bridge as well, and it will be used later on.

Remarks. Connecting the two integrals $E(q, c)$ and $I(q, c)$, Theorem A allows for the characterization of local functions of a Wiener process in terms of any of these integrals, equivalently.

Theorem 1.1.B constitutes a test for $q \in \mathcal{Q}$ being (or not) a local function of a standard Wiener process, i.e., for the two statements in (1.1.1), by saying that with $q \in \mathcal{Q}$

$$(1.1.8) \quad P\{\limsup_{t \downarrow 0} |W(t)|/q(t) < \infty\} = 1,$$

i.e. $q \in \mathcal{Q}$ is a *local function* of W , if and only if there is a $c > 0$ such that $I(q, c) < \infty$ (or, $E(q, c) < \infty$ and $\lim_{t \downarrow 0} q(t)/t^{1/2} = \infty$), while we have

$$(1.1.9) \quad P\{\limsup_{t \downarrow 0} |W(t)|/q(t) < \infty\} = 0,$$

i.e. $q \in \mathcal{Q}$ is not a local function of W , if and only if $I(q, c) = \infty$ (or $E(q, c) = \infty$) for all $c > 0$.

Theorem 1.1.C is an extension of O'Reilly's (1974) Proposition 2.1 for possibly discontinuous q functions in $\mathcal{Q}$ and constitutes a test for $q \in \mathcal{Q}$ being (or not) a Chibisov-O'Reilly function by saying that with $q \in \mathcal{Q}$

$$P\left\{\limsup_{t \downarrow 0} |W(t)|/q(t) = 0\right\} = 1,$$

i.e. $q \in \mathcal{Q}$ is a Chibisov-O'Reilly function of W , if and only if for all $c > 0$ we have $I(q, c) < \infty$ (or $E(q, c) < \infty$ and $\lim_{t \downarrow 0} q(t)/t^{1/2} = \infty$), while we have

$$P\left\{\limsup_{t \downarrow 0} |W(t)|/q(t) = 0\right\} = 0$$

i.e. $q \in \mathcal{Q}$ is not a Chibisov-O'Reilly function of W , if and only if there is a $c > 0$ such that $I(q, c) = \infty$ (or $E(q, c) = \infty$).

Naturally, if with $q \in \mathcal{Q}$ we have

$$(1.1.10) \quad P\{|W(t)| < q(t), t \downarrow 0\} = 1,$$

then we have also (1.1.8), i.e. all local upper-class functions $q \in \mathcal{Q}$ of W are local functions of W and hence, for any local upper-class function $q \in \mathcal{Q}$ of W there is a $c > 0$ such that $I(q, c) < \infty$ (or, $E(q, c) < \infty$ and $\lim_{t \downarrow 0} q(t)/t^{1/2} = \infty$). Conversely, however, we can only conclude (cf. Theorem 1.1.B) that $q \in \mathcal{Q}$ is a local function of W , i.e. that we have (1.1.8) but not necessarily also (1.1.10) at the same time. Thus such a $q \in \mathcal{Q}$ may not be a local upper-class function of W . If it is, however, then it does not have to be in $C(0, 1]$, nor does it have to be such that $q(t)/t^{1/2}$ is non-decreasing for small $t > 0$, as required by the classical Erdős-Feller-Kolmogorov-Petrovski [EFKP] test as quoted above from Itô and McKean (1965).

While proving Theorem 1.1.B (cf. Proof of Theorem 3.3 of CsCsHM, 1986), CsCsHM (1986) actually established more than what the second part of the latter theorem claims. Namely they proved also

Theorem 1.1.D. *Let $q \in \mathcal{Q}$, and $W(\cdot)$ be a standard Wiener process starting at zero. Then $I(q, c) < \infty$ implies*

$$(1.1.11) \quad \limsup_{t \downarrow 0} |W(t)|/q(t) \leq (2c)^{1/2} \quad \text{a.s.,}$$

and if (1.1.3) holds true for some $0 \leq \beta < \infty$, then $I(q, c) < \infty$ for any $c > 8/\beta^2$.

Hence CsCsHM (1986) conjectured that (1.1.3) holding true for some $\beta \geq 0$ should imply $I(q, c) < \infty$ for any $c > \beta^2/2$, for then the latter combined with

(1.1.11) and Theorem 1.1.A would amount to saying that when testing for (1.1.4) and (1.1.5), the assumptions inherited from the classical EFKP local upper-lower functions integral test that q be continuous, and be such that $q(t)/t^{1/2}$ is non-increasing for small $t > 0$, could be dropped.

This conjecture turned out to be true. Namely we have (cf. Theorem 1.1. of Csörgő, Shao and Szyszkowicz, 1991)

Theorem 1.1.E. *Let $q \in \mathcal{Q}$, and $W(\cdot)$ be a standard Wiener process starting at zero. Then (1.1.3) holds true with some $\beta \geq 0$, i.e., we have*

$$\limsup_{t \downarrow 0} |W(t)|/q(t) = \beta \quad a.s.$$

if and only if

$$I(q, c) \begin{cases} < \infty & \text{for any } c > \beta^2/2 \\ = \infty & \text{for any } c < \beta^2/2, \end{cases}$$

or, equivalently, if and only if

$$E(q, c) \begin{cases} < \infty & \text{for any } c > \beta^2/2, \\ = \infty & \text{for any } c < \beta^2/2, \end{cases}$$

while $\lim_{t \downarrow 0} q(t)/t^{1/2} = \infty$.

Another, somewhat more convenient version of Theorem 1.1.E reads as follows.

Theorem 1.1.E*. *Let $q \in \mathcal{Q}$, and $W(\cdot)$ be a standard Wiener process starting at zero. Then with some $c \geq 0$ we have*

$$(1.1.12) \quad \limsup_{t \downarrow 0} |W(t)|/q(t) = (2c)^{1/2} \quad a.s.$$

if and only if for any $\varepsilon > 0$ we have $I(q, c + \varepsilon) < \infty$ and $I(q, c - \varepsilon) = \infty$ or, equivalently, if and only if $E(q, c + \varepsilon) < \infty$ and $E(q, c - \varepsilon) = \infty$ with $\lim_{t \downarrow 0} q(t)/t^{1/2} = \infty$.

Consequently, requiring only that $q \in \mathcal{Q}$, with $\varepsilon > 0$ we conclude that $((2c)^{1/2} + \varepsilon)q$ belongs to the local upper-class (resp. $((2c)^{1/2} - \varepsilon)q$ belongs to the local lower-class) for W if and only if $I(q, c + \varepsilon) < \infty$ (resp. $I(q, c - \varepsilon) = \infty$) or, equivalently, if and only if $E(q, c + \varepsilon) < \infty$ (resp. $E(q, c - \varepsilon) = \infty$) with $\lim_{t \downarrow 0} q(t)/t^{1/2} = \infty$.

Khinchin's celebrated law of the iterated logarithm states that

$$r(t) = (1 + \epsilon) \left(2t \log \log \frac{1}{t} \right)^{\frac{1}{2}}$$

belongs to the upper class if $\epsilon > 0$ and to the lower class if $\epsilon < 0$, i.e.,

$$P \left\{ \limsup_{t \downarrow 0} |W(t)| / \left(2t \log \log \frac{1}{t} \right)^{\frac{1}{2}} = 1 \right\} = 1.$$

Theorem 1.1.E is an extension of Khinchin's law and constitutes a test for

$$P \left\{ \limsup_{t \downarrow 0} |W(t)|/q(t) = \beta \right\} = 1$$

for some $0 \leq \beta < \infty$. It achieves testing for $(\beta + \epsilon)q(t) \in \mathcal{Q}$, respectively for $(\beta - \epsilon)q(t) \in \mathcal{Q}$, with $\epsilon > 0$ being a local upper-class, respectively local lower-class function of a standard Wiener process, i.e., that $q \in \mathcal{Q}$ is a local function of W , without requiring that q be continuous and $q(t)/t^{1/2}$ be nonincreasing for small $t > 0$. Whether Kolmogorov's classical test for

$$P \left\{ |W(t)| < r(t), t \downarrow 0 \right\} = 0 \quad \text{or} \quad 1$$

is feasible without requiring the just mentioned continuity and monotonicity assumptions, i.e. assuming only that $q \in \mathcal{Q}$, remains an open question. In the light of Theorem 1.1.E, Csörgő, Shao and Szyszkowicz (1991) conjectured that the answer to this question should be in the affirmative. We note also that requiring for $q \in \mathcal{Q}$ to be so that $q(t)/t^{1/2}$ is nonincreasing near zero is a considerable restriction on the class of possible weight functions $q \in \mathcal{Q}$ for which we have $I(q, c) < \infty$ for all $c > 0$. For examples and further discussion along these lines we refer to Csörgő, Csörgő and Horváth (1986).

1.2. Sup-metric approximation of partial sums

Let $X_1, X_2, \dots$ be independent, identically distributed random variables and put $S(nt) = X_1 + \dots + X_{[nt]}$, $0 \leq t \leq 1$, where $[x]$ denotes the integer part of x . Assuming the existence of two moments only, we prove approximations of $n^{-1/2}S(nt)$ in $\|\cdot/q\|$ -metrics (weighted supremum norm) for the optimal class of weight functions.

As in Section 1.1, let $\mathcal{Q}$ be the class of functions q defined on $(0, 1]$ which are nondecreasing in a neighbourhood of zero and

$$(1.2.1) \quad \inf_{\delta \leq t \leq 1} q(t) > 0 \quad \text{for all } 0 < \delta < 1,$$

and once again we define the following integrals:

$$E(q, c) = \int_0^1 t^{-3/2} q(t) \exp(-ct^{-1}q^2(t)) dt,$$

and

$$I(q, c) = \int_0^1 t^{-1} \exp(-ct^{-1}q^2(t)) dt$$

for some constant $0 < c < \infty$.

Theorem 1.2.1. *Let $X_1, X_2, \dots$ be independent, identically distributed random variables such that*

$$(1.2.2) \quad EX_1 = 0, \quad EX_1^2 = 1,$$

and for each $n \geq 1$ let $S(nt) = X_1 + \dots + X_{[nt]}$ with $S(0) = 0$. Then a standard Wiener process $\{W(t), 0 \leq t < \infty\}$ can be constructed in such a way that the following statements hold true.

(a) Let $q \in \mathcal{Q}$. Then

$$(1.2.3) \quad \sup_{0 < t \leq 1} |n^{-1/2}(S(nt) - W(nt))|/q(t) = o_P(1)$$

if and only if $I(q, c) < \infty$ for all $c > 0$.

(b) Let $q \in \mathcal{Q}$. Then

$$(1.2.4) \quad \sup_{0 < t \leq 1} |n^{-1/2}(S(nt) - W(nt))|/q(t) = O_P(1)$$

if and only if $I(q, c) < \infty$ for some $c > 0$.

Remark 1.2.1. According to Theorem 1.1.C, the statement (1.2.3) of Theorem 1.2.1 is also equivalent to

$$E(q, c) < \infty \text{ for all } c > 0 \text{ and } \lim_{t \downarrow 0} q(t)/t^{1/2} = \infty.$$

Similarly, according to Theorem 1.1.B, the statement (1.2.4) of Theorem 1.2.1 is also equivalent to

$$E(q, c) < \infty \text{ for some } c > 0 \text{ and } \lim_{t \downarrow 0} q(t)/t^{1/2} = \infty.$$

Corollary 1.2.1. Let $X_1, X_2, \dots$ be independent, identically distributed random variables such that $EX_1 = 0$, $EX_1^2 = 1$, and $\{W(t) : 0 \leq t \leq 1\}$ be a standard Wiener process.

(a) Let $q \in \mathcal{Q}$. Then

$$S(nt)/(n^{1/2}q(t)) \xrightarrow{\mathcal{D}} W(t)/q(t) \text{ in } D[0, 1]$$

if and only if $I(q, c) < \infty$ for all $c > 0$.

(b) Let $q : (0, 1] \rightarrow (0, \infty)$ be such that (1.2.1) holds. Then

$$S(nt)/(n^{1/2}q(t)) \xrightarrow{\mathcal{D}} W(t)/q(t) \text{ in } D[\delta, 1]$$

for any $\delta \in (0, 1)$.

Part (a) of Corollary 1.2.1 was proved by O'Reilly (1974) for continuous weight functions under the assumption of $E|X_1|^3 < \infty$. He asserted also, without proof, that the third moment condition could be dropped to two, and his theorem would remain true. For an extension of the Komlós, Major and Tusnády (1975, 1976) approximation of partial sums of i.i.d.r.v.'s having $r > 2$ moments, to weighted supremum norm approximations which improve also the just mentioned result of O'Reilly (1974) in terms of the optimal class of weight functions as obtained in CsCsHM (1986), we refer to Csörgő and Horváth (1988c) and the references given there. Using the same class of optimal functions we are proving approximations of $n^{-1/2}S(nt)$ in $\|\cdot/q\|$ -metrics, assuming the existence of two moments only. In our proof we will use the following result of Major (1979).

Theorem 1.2.A. *Let a distribution $F(x)$ be given with $\int x dF(x) = 0$, $\int x^2 dF(x) = 1$. Define*

$$\sigma_k^2 = \int_{-\sqrt{2^n}}^{\sqrt{2^n}} x^2 dF(x) - \left(\int_{-\sqrt{2^n}}^{\sqrt{2^n}} x dF(x) \right)^2 \quad \text{if } 2^n \leq k < 2^{n+1}, \quad n = 1, 2, \dots$$

A sequence of i.i.d.r.v.'s $X_1, X_2, \dots$ with distribution function $F(x)$, and a sequence of independent normal random variables $Y_1, Y_2, \dots$ with $EY_k = 0$, $EY_k^2 = \sigma_k^2$ can be constructed in such a way that the partial sums $S(n) = X_1 + \dots + X_n$, $T(n) = Y_1 + \dots + Y_n$, $n = 1, 2, \dots$ satisfy the relation

$$|S(n) - T(n)| \stackrel{a.s.}{=} o(n^{1/2}).$$

We note here that in this theorem the partial sum sequence $\{S(n)\}$ is approximated by a Gaussian process $\{T(n)\}$ and not by a Wiener process $\{W(n)\}$. In fact it cannot be approximated by a Wiener process at that rate as proved by

Major (1976b) (cf. Theorem 2.2.3 in Csörgő and Révész, 1981). He showed that for any sequence $\{a_n\}$ of real numbers with $a_n \nearrow \infty$ there exists a distribution function F with mean 0 and variance 1 such that for any i.i.d. sequence $\{X_i\}$ having the distribution F and for any Wiener process $W(t)$ we have

$$\overline{\lim}_{n \rightarrow \infty} a_n |S(n) - W(n)| / (n \log \log n)^{1/2} \stackrel{a.s.}{=} \infty.$$

In other words this result is saying that if we assume the existence of two moments only then the rate of approximation of Strassen's theorem (cf. Strassen, 1964) is the best possible if $S(n)$ is to be approximated by a Wiener process. In the same time, Theorem 1.2.A is that strong invariance principle which implies Donsker's theorem, as well as the Strassen (1964) strong invariance principle for the law of the iterated logarithm. Namely, as in Csörgő and Révész (1981), p.112, we take a Wiener process $W(t)$ such that

$$(1.2.5) \quad W(n) = \sum_{i=1}^n Y_i / \sigma_i, \quad n = 1, 2, \dots,$$

and let $T(nt) = Y_1 + \dots + Y_{[nt]}$, $0 \leq t \leq 1$. Since $\sigma_n \uparrow 1$, by Kolmogorov's inequality

$$(1.2.6) \quad \begin{aligned} & \sup_{0 \leq t \leq 1} |n^{-1/2}(T(nt) - W([nt]))| \\ &= \sup_{0 \leq t \leq 1} \left| n^{-1/2} \left(\sum_{i=1}^{[nt]} Y_i - \sum_{i=1}^{[nt]} Y_i / \sigma_i \right) \right| \xrightarrow{P} 0. \end{aligned}$$

Hence, using Theorem 1.2.A and continuity of a Wiener process, we have also

$$(1.2.7) \quad \sup_{0 \leq t \leq 1} |n^{-1/2}(S(nt) - W(nt))| \xrightarrow{P} 0.$$

In fact, it will now be seen that Theorem 1.2.A is strong enough to imply also the "weighted" Donsker's theorem. Namely we prove the weighted version of (1.2.7)

for the optimal class of weight functions as stated in Theorem 1.2.1. For further use in the sequel, we prove first a lemma.

Lemma 1.2.1. *Let $\{W(t), 0 \leq t < \infty\}$ be a standard Wiener process. Then we have, as $n \rightarrow \infty$*

$$\sup_{1/n \leq t \leq 1} n^{-1/2} |W(nt) - W([nt])| / t^{1/2} \stackrel{a.s.}{=} O(1).$$

Proof. First we note that

$$\begin{aligned} & \sup_{1/n \leq t \leq 1} n^{-1/2} |W(nt) - W([nt])| / t^{1/2} \\ & \leq \sup_{1/n \leq t \leq 1} \sup_{0 \leq s \leq 1} |W([nt] + s) - W([nt])| / (nt)^{1/2}. \end{aligned}$$

By Theorem 1.2.1 of Csörgő and Révész (1981) we have

$$\sup_{0 \leq x \leq [nt]-1} \sup_{0 \leq s \leq 1} |W(x+s) - W(x)| / (2 \log nt)^{1/2} = 1 \quad a.s.$$

as $nt \rightarrow \infty$, which implies

$$\sup_{0 \leq s \leq 1} |W([nt] + s) - W([nt])| = O((\log nt)^{1/2}) \quad a.s.,$$

and consequently

$$\sup_{1 \leq nt \leq n} \sup_{0 \leq s \leq 1} |W([nt] + s) - W([nt])| / (nt)^{1/2} = O(1) \quad a.s..$$

It is of interest to note again that Theorem 1.2.A implies also Strassen's strong invariance principle for the law of iterated logarithm, namely that there exists a Wiener process W such that (cf. Strassen, 1964)

$$\sup_{0 \leq t \leq 1} |S(nt) - W(nt)| / (n \log \log n)^{1/2} \stackrel{a.s.}{=} o(1).$$

In fact, taking $X_1, X_2, \dots, Y_1, Y_2, \dots$ as in Theorem 1.2.A with $S(k) = X_1 + \dots + X_k$, $T(k) = Y_1 + \dots + Y_k$, $1 \leq k \leq n$, and a Wiener process $W(t)$ such that (1.2.5) holds, we have

$$\begin{aligned} & \{T(nt) - W([nt]), 0 \leq nt < \infty\} \\ &= \left\{ \sum_{i=1}^{[nt]} \left(1 - \frac{1}{\sigma_i}\right) Y_i, 0 \leq nt < \infty \right\} \\ &\stackrel{D}{=} \left\{ W\left(\sum_{i=1}^{[nt]} (\sigma_i - 1)^2\right), 0 \leq nt < \infty \right\}. \end{aligned}$$

Since, as $nt \rightarrow \infty$,

$$\left| W\left(\sum_{i=1}^{[nt]} (\sigma_i - 1)^2\right) \right| \stackrel{a.s.}{=} O\left(\left(\sum_{i=1}^{[nt]} (\sigma_i - 1)^2 \log \log \sum_{i=1}^{[nt]} (\sigma_i - 1)^2\right)^{1/2}\right),$$

we obtain

$$\begin{aligned} & \sup_{0 \leq t \leq 1} |T(nt) - W([nt])| / (n \log \log n)^{1/2} \\ &= \sup_{0 \leq t \leq 1} \left| \sum_{i=1}^{[nt]} \left(1 - \frac{1}{\sigma_i}\right) Y_i \right| / (n \log \log n)^{1/2} \\ &\stackrel{a.s.}{=} \sup_{0 \leq t \leq 1} \left| W\left(\sum_{i=1}^{[nt]} (\sigma_i - 1)^2\right) \right| / (n \log \log n)^{1/2} \\ &\stackrel{a.s.}{=} O\left(\left(\left(\sum_{i=1}^n (\sigma_i - 1)^2 \log \log \sum_{i=1}^n (\sigma_i - 1)^2\right) / (n \log \log n)\right)^{1/2}\right) \\ &= o(1)O(1), \end{aligned}$$

since $\sigma_i \uparrow 1$. Consequently, using now Theorem 1.2.A and proof of Lemma 1.2.1, with a Wiener process $W(t)$ such that (1.2.5) holds, we have

$$\begin{aligned} & \sup_{0 \leq t \leq 1} |S(nt) - W(nt)| / (n \log \log n)^{1/2} \\ &\leq \sup_{0 \leq t \leq 1} |S(nt) - T(nt)| / (n \log \log n)^{1/2} \\ &+ \sup_{0 \leq t \leq 1} |T(nt) - W([nt])| / (n \log \log n)^{1/2} \\ &+ \sup_{0 \leq t \leq 1} |W([nt]) - W(nt)| / (n \log \log n)^{1/2} \\ &\stackrel{a.s.}{=} o\left((\log \log n)^{-1/2}\right) + o(1) + O\left((\log n)^{1/2} (n \log \log n)^{-1/2}\right) \\ &= o(1). \end{aligned}$$

Another way of normalizing partial sums $S(k) = X_1 + \dots + X_k$ is via Darling-Erdős (1956) theorem. Let

$$a_n = (2 \log \log n)^{1/2}$$

$$b_n = 2 \log \log n + \frac{1}{2} \log \log \log n - \frac{1}{2} \log \pi.$$

Theorem 1.2.B. *Let $X_1, X_2, \dots$ be i.i.d.r.v.'s such that*

$$EX_1 = 0, \quad EX_1^2 = 1 \quad \text{and} \quad E|X_1|^3 < \infty.$$

Then we have:

$$(a) \quad a_n \sup_{0 < t < 1} n^{-1/2} S(nt)/t^{1/2} - b_n \xrightarrow{D} E_v,$$

$$(b) \quad a_n \sup_{0 < t < 1} n^{-1/2} |S(nt)|/t^{1/2} - b_n \xrightarrow{D} E_v^2,$$

where E_v and E_v^2 are the random variables with distribution functions $\exp(-\exp(-t))$ and $\exp(-2\exp(-t))$, respectively.

The above theorem does not follow from the just stated invariance principles for partial sum processes assuming two moments only. Actually, more than two moments are needed. As to what exactly the Darling-Erdős (1956) theorem is all about in terms of two moments and a bit more, we refer to Einmahl (1989).

Proof of Theorem 1.2.1. Let $X_1, X_2, \dots$ and $Y_1, Y_2, \dots$ be as in Theorem 1.2.A and let $\{W(t), 0 \leq t < \infty\}$ be a Wiener process as in (1.2.5). Since for each $n \geq 1$

$$(1.2.8) \quad \{n^{-1/2} W(nt); 0 \leq t \leq 1\} \stackrel{D}{=} \{W(t); 0 \leq t \leq 1\},$$

by Theorems 1.1.C and 1.1.B respectively, with $q \in \mathcal{Q}$ we obtain

$$(1.2.9) \quad \sup_{0 < t < 1/n} n^{-1/2} |W(nt)|/q(t) = \begin{cases} o_P(1) & \text{if and only if } I(q, c) < \infty \text{ for all } c > 0 \\ O_P(1) & \text{if and only if } I(q, c) < \infty \text{ for some } c > 0. \end{cases}$$

By Theorem 1.2.A we have

$$|S(nt) - T(nt)| \stackrel{a.s.}{=} o((nt)^{1/2}), \quad \text{as } nt \rightarrow \infty,$$

and consequently

$$\sup_{1 \leq nt < \infty} |S(nt) - T(nt)|/(nt)^{1/2} \stackrel{a.s.}{=} O(1), \quad \text{as } n \rightarrow \infty.$$

Let $\delta \in (0, 1)$ be fixed and n be such that $1/n < \delta$. Then, a.s. as $n \rightarrow \infty$,

$$\begin{aligned} (1.2.10) \quad & \sup_{1/n \leq t < \delta} |n^{-1/2}(S(nt) - T(nt))|/q(t) \\ & \leq O(1) \sup_{0 < t < \delta} t^{1/2}/q(t). \end{aligned}$$

Using Theorem 1.2.A once again, we get

$$(1.2.11) \quad \sup_{\delta \leq t < 1} |n^{-1/2}(S(nt) - T(nt))|/q(t) \stackrel{a.s.}{=} o(1)$$

for any $\delta \in (0, 1)$, due to the positivity of q (cf. 1.2.1)). Taking $\delta > 0$ arbitrarily small, by (1.2.10) and (1.2.11) we conclude

$$(1.2.12) \quad \sup_{1/n \leq t \leq 1} |n^{-1/2}(S(nt) - T(nt))|/q(t) \stackrel{a.s.}{=} o(1)$$

for any $q \in \mathcal{Q}$ such that

$$(1.2.13) \quad \lim_{t \downarrow 0} t^{1/2}/q(t) = 0.$$

Next we note that for each $n \geq 1$

$$\left\{ n^{-1/2} \sum_{i=1}^{[nt]} \left(1 - \frac{1}{\sigma_i}\right) Y_i, \quad 0 \leq t \leq 1 \right\} \stackrel{D}{=} \left\{ W\left(\frac{1}{n} \sum_{i=1}^{[nt]} (\sigma_i - 1)^2\right), \quad 0 \leq t \leq 1 \right\}.$$

Let $\delta > 0$ be small enough, such that q is already nondecreasing and let n be such that $1/n < \delta$. Since $\sum_{i=1}^{[nt]} (\sigma_i - 1)^2 \leq nt$, we have

$$\begin{aligned}
& \sup_{1/n \leq t < \delta} |n^{-1/2}(T(nt) - \sum_{i=1}^{[nt]} Y_i/\sigma_i)|/q(t) \\
& \stackrel{D}{=} \sup_{1/n \leq t < \delta} |W\left(\frac{1}{n} \sum_{i=1}^{[nt]} (\sigma_i - 1)^2\right)|/q(t) \\
& \leq \sup_{1/n \leq t < \delta} |W\left(\frac{1}{n} \sum_{i=1}^{[nt]} (\sigma_i - 1)^2\right)|/q\left(\frac{1}{n} \sum_{i=1}^{[nt]} (\sigma_i - 1)^2\right) \\
(1.2.14) \quad & \leq \sup_{0 < t < \delta} |W(t)|/q(t).
\end{aligned}$$

By (1.2.6) we get

$$(1.2.15) \quad \sup_{\delta \leq t \leq 1} |n^{-1/2}\left(T(nt) - \sum_{i=1}^{[nt]} Y_i/\sigma_i\right)|/q(t) = o_P(1)$$

for any $\delta \in (0, 1)$. Since $\delta > 0$ can be taken arbitrarily small, using again Theorems 1.1.C and 1.1.B, by (1.2.14) and (1.2.15) we conclude

$$\begin{aligned}
(1.2.16) \quad & \sup_{1/n \leq t \leq 1} |n^{-1/2}\left(T(nt) - \sum_{i=1}^{[nt]} Y_i/\sigma_i\right)|/q(t) \\
& = \begin{cases} o_P(1) & \text{if } I(q, c) < \infty \text{ for all } c > 0 \\ O_P(1) & \text{if } I(q, c) < \infty \text{ for some } c > 0. \end{cases}
\end{aligned}$$

Since

$$\sup_{0 \leq t \leq 1} |W([nt]) - W(nt)| \stackrel{a.s.}{=} O((\log n)^{1/2}),$$

we have with $q \in \mathcal{Q}$

$$\sup_{\delta \leq t \leq 1} |n^{-1/2}(W([nt]) - W(nt))/q(t)| \stackrel{a.s.}{=} o(1)$$

for any $\delta > 0$, due to (1.2.1). Using Lemma 1.2.1 we obtain

$$\begin{aligned} & \sup_{1/n \leq t \leq \delta} |n^{-1/2}(W([nt]) - W(nt))|/q(t) \\ & \leq \sup_{1/n \leq t \leq \delta} |n^{-1/2}(W([nt]) - W(nt))|/t^{1/2} \sup_{0 < t \leq \delta} t^{1/2}/q(t) \\ & \stackrel{a.s.}{=} o(1) \end{aligned}$$

for any $q \in \mathcal{Q}$ such that (1.2.13) holds. Taking $\delta > 0$ arbitrarily small, we get

$$(1.2.17) \quad \sup_{1/n \leq t < 1} |n^{-1/2}(W([nt]) - W(nt))|/q(t) \stackrel{a.s.}{=} o(1)$$

for any $q \in \mathcal{Q}$ such that (1.2.13) holds. Since Theorem 1.1.B implies (1.2.13) for any $q \in \mathcal{Q}$ such that $I(q, c) < \infty$ for some $c > 0$, by combining (1.2.9), (1.2.12), (1.2.16) and (1.2.17) we obtain

$$\begin{aligned} & \sup_{0 < t \leq 1} |n^{-1/2}(S(nt) - W(nt))|/q(t) \\ & \leq \sup_{0 < t < 1/n} |n^{-1/2}W(nt)|/q(t) \\ & + \sup_{1/n \leq t \leq 1} |n^{-1/2}(S(nt) - T(nt))|/q(t) \\ & + \sup_{1/n \leq t \leq 1} \left| n^{-1/2} \left(T(nt) - \sum_{i=1}^{[nt]} Y_i / \sigma_i \right) \right| / q(t) \\ & + \sup_{1/n \leq t \leq 1} \left| n^{-1/2} \left(\sum_{i=1}^{[nt]} Y_i / \sigma_i - W(nt) \right) \right| / q(t) \\ & = \begin{cases} o_P(1) & \text{if } I(q, c) < \infty \text{ for all } c > 0 \\ O_P(1) & \text{if } I(q, c) < \infty \text{ for some } c > 0. \end{cases} \end{aligned}$$

The converse parts of (a) and (b) follow from (1.2.9), since

$$\sup_{0 \leq t \leq 1} |n^{-1/2}(S(nt) - W(nt))|/q(t) \geq \sup_{0 < t < 1/n} |n^{-1/2}W(nt)|/q(t).$$

Proof of Corollary 1.2.1. Let $S_n(t) = S(nt)/n^{1/2}$.

(a) Obviously, with $q \in \mathcal{Q}$ and $I(q, c) < \infty$ for all $c > 0$, by Theorem 1.2.1 (a) we have

$$(1.2.18) \quad S_n(t)/q(t) \xrightarrow{\mathcal{D}} W(t)/q(t) \quad \text{in } D[0, 1],$$

where W is a standard Wiener process.

Assuming now (1.2.18) with $q \in \mathcal{Q}$, by the Skorohod-Dudley-Wichura theorem (cf. Theorem A.2.1) there exists S_n^* , $n \geq 1$, and W^* such that

$$S_n^* \xrightarrow{\mathcal{D}} S_n, \quad n \geq 1, \quad W^* \xrightarrow{\mathcal{D}} W$$

and

$$\sup_{0 < t \leq 1} |S_n^*(t) - W^*(t)|/q(t) = o_P(1).$$

Hence, using the technique of Lemma 3.1.1 in Csörgő (1983) (cf. Section A.1), there exists a sequence of Wiener processes $\{W_n(t), 0 \leq t \leq 1\}$ such that

$$\sup_{0 < t \leq 1} |S_n(t) - W_n(t)|/q(t) = o_P(1),$$

which, by Theorem 1.2.1 (a) implies $I(q, c) < \infty$ for all $c > 0$.

(b) Using Theorem 1.2.A (or Theorem 1.2.1 (a) with $q \equiv 1$), we have

$$\sup_{0 < t \leq 1} |S_n(t) - n^{-1/2}W(nt)| = o_P(1),$$

and hence with any $q : (0, 1] \rightarrow (0, \infty)$ such that $\inf_{\delta \leq t \leq 1} q(t) > 0$ we have

$$\sup_{\delta \leq t \leq 1} |S_n(t) - n^{-1/2}W(nt)|/q(t) = o_P(1)$$

for any $\delta \in (0, 1)$.

1.3. Random integrals

We say that q is positive on (a, b) if $\inf_{a+\delta \leq t \leq b-\delta} q(t) > 0$ for all $0 < \delta < (b-a)/2$. Similarly, q is positive on $(a, b]$, if $\inf_{a+\delta \leq t \leq b} q(t) > 0$ for all $0 < \delta < b-a$.

First we present a very simple condition which implies the almost sure finiteness of $\int_0^1 |W(t)|^p / q(t) dt$, where $\{W(t), 0 \leq t < \infty\}$ is a Wiener process (cf. Lemma 3.1 in Csörgő and Horváth, 1992).

Lemma 1.3.A. *We assume that $0 < p < \infty$ and q is positive on $(0, 1]$. Then the statements*

$$(1.3.1) \quad \int_0^1 t^{p/2} / q(t) dt < \infty$$

and

$$(1.3.2) \quad \int_0^1 |W(t)|^p / q(t) dt < \infty \quad a.s.$$

are equivalent.

In case of $p = 2$, Lemma 1.3.A is due to Shepp (1966), whose proof uses computations of Radon-Nikodym derivatives of Gaussian processes. This method of proof does not work for the general case of $0 < p < \infty$. A simpler method of proof for Lemma 1.3.A was developed in Csörgő, Horváth and Shao (1991), where they find a necessary and sufficient condition for the weak convergence of the uniform empirical and quantile processes to Brownian bridge in weighted L_p -distances. Under the same condition, weighted L_p -functionals of the uniform empirical and quantile processes converge in distribution to the corresponding functionals of a Brownian bridge. Csörgő, Horváth and Shao (1991) prove also some general dichotomy theorems for integrals of stochastic processes and we present here two of their results which we will use later on.

Theorem 1.3.A. Let $0 < p < \infty$, $-\infty \leq a < b \leq \infty$, q be a positive function on (a, b) and $\{\eta(t), a < t < b\}$ be a Gaussian process with $E\eta(t) = m(t)$ and $\sigma^2(t) = \text{var } \eta(t)$. We assume

$$(1.3.3) \quad \sup_{a < t < b} |m(t)/\sigma(t)| < \infty.$$

Then

$$\int_a^b |\eta(t)|^p / q(t) dt < \infty \quad \text{a.s.}$$

if and only if

$$\int_a^b \sigma^p(t) / q(t) dt < \infty.$$

Corollary 1.3.A. Let $0 < p < \infty$, $-\infty \leq a < b \leq \infty$ and $\{\eta(t), a < t < b\}$ be a Gaussian process with $E\eta(t) = 0$ and $\sigma^2(t) = \text{var } \eta(t)$. Then

$$P\left\{\int_a^b |\eta(t)|^p dt < \infty\right\} = \begin{cases} 1, & \text{if } \int_a^b \sigma^p(t) dt < \infty \\ 0, & \text{if } \int_a^b \sigma^p(t) dt = \infty. \end{cases}$$

1.4. L_p -approximations of partial sums

As before, let $\mathcal{Q}$ be the class of those positive functions on $(0, 1]$ which are nondecreasing in a neighbourhood of zero.

Theorem 1.4.1 Let $X_1, X_2, \dots$ be independent, identically distributed random variables such that

$$EX_1 = 0 \quad \text{and} \quad EX_1^2 = 1.$$

We assume that $0 < p < \infty$ and $q \in \mathcal{Q}$.

(a) A standard Wiener process $\{W(t), 0 \leq t < \infty\}$ can be constructed in such a way that

$$(1.4.1) \quad \int_0^1 |n^{-1/2}(S(nt) - W(nt))|^p / q(t) dt = o_P(1)$$

if and only if

$$(1.4.2) \quad \int_0^1 t^{p/2}/q(t)dt < \infty.$$

(b) Let $\{W(t), 0 \leq t < \infty\}$ be a standard Wiener process. Then

$$(1.4.3) \quad \int_0^1 |n^{-1/2}S(nt)|^p/q(t)dt \xrightarrow{P} \int_0^1 |W(t)|^p/q(t)dt$$

if and only if (1.4.2) holds.

We call attention to the fact that (a) and (b) are equivalent.

Proof. (a) Let $X_1, X_2, \dots$ and $Y_1, Y_2, \dots$ be as in Theorem 1.2.A and $\{W(t), 0 \leq t < \infty\}$ be a Wiener process as in (1.2.5), i.e. $W(n) = \sum_{i=1}^n Y_i/\sigma_i, n = 1, 2, \dots$

We put $S(nt) = X_1 + \dots + X_{[nt]}, T(nt) = Y_1 + \dots + Y_{[nt]}$ and $S(0) = T(0) = 0$.

First we assume that $q \in \mathcal{Q}$ and (1.4.2) holds. We note that for each $n \geq 1$ we have $\{n^{-1/2}W(nt), 0 \leq t \leq 1\} \stackrel{D}{=} \{W(t), 0 \leq t \leq 1\}$. Hence, using Lemma 1.3.A we obtain

$$(1.4.4) \quad \int_0^{1/n} |n^{-1/2}W(nt)|^p/q(t)dt = o_P(1).$$

As noted already (cf. (1.2.7)) we have

$$\sup_{0 \leq t \leq 1} |n^{-1/2}(S(nt) - W(nt))| = o_P(1),$$

which implies

$$(1.4.5) \quad \sup_{\delta \leq t \leq 1} |n^{-1/2}(S(nt) - W(nt))|/q(t) = o_P(1),$$

for any $\delta \in (0, 1)$ due to the positivity of $q \in \mathcal{Q}$. Consequently

$$(1.4.6) \quad \int_{\delta}^1 |n^{-1/2}(S(nt) - W(nt))|^p/q(t)dt = o_P(1)$$

for any $\delta \in (0, 1)$. By Theorem 1.2.A we have

$$|S(nt) - T(nt)| \stackrel{a.s.}{=} o((nt)^{1/2}), \quad \text{as } nt \rightarrow \infty.$$

Hence, as $n \rightarrow \infty$,

$$\sup_{1 \leq nt < \infty} |n^{-1/2}(S(nt) - T(nt))|/t^{1/2} \stackrel{a.s.}{=} O(1),$$

and we have for any $\delta \in (0, 1)$

$$\begin{aligned} & \int_{1/n}^{\delta} |n^{-1/2}(S(nt) - T(nt))|^p / q(t) dt \\ & \leq \sup_{1/n \leq t \leq 1} \left(|n^{-1/2}(S(nt) - T(nt))|/t^{1/2} \right)^p \int_0^{\delta} t^{p/2} / q(t) dt \\ (1.4.7) \quad & \stackrel{a.s.}{=} O(1) \int_0^{\delta} t^{p/2} / q(t) dt. \end{aligned}$$

We note that

$$\begin{aligned} & \left\{ \sum_{i=1}^{[nt]} \left(1 - \frac{1}{\sigma_i}\right) Y_i; \quad 0 \leq t \leq 1, \quad n = 1, 2, \dots \right\} \\ & \stackrel{D}{=} \left\{ W\left(\sum_{i=1}^{[nt]} (\sigma_i - 1)^2\right); \quad 0 \leq t \leq 1, \quad n = 1, 2, \dots \right\} \end{aligned}$$

and hence, for each $n \geq 1$

$$\left\{ n^{-1/2} \sum_{i=1}^{[nt]} \left(1 - \frac{1}{\sigma_i}\right) Y_i; \quad 0 \leq t \leq 1 \right\} \stackrel{D}{=} \left\{ W\left(\frac{1}{n} \sum_{i=1}^{[nt]} (\sigma_i - 1)^2\right); \quad 0 \leq t \leq 1 \right\}.$$

Let $\delta > 0$ be small enough, such that q is already nondecreasing, and let n be such that $\frac{1}{n} < \delta$. Since $\sum_{i=1}^{[nt]} (\sigma_i - 1)^2 \leq nt$, we have

$$\begin{aligned} & \int_{1/n}^{\delta} \left| n^{-1/2} \left(T(nt) - \sum_{i=1}^{[nt]} Y_i / \sigma_i \right) \right|^p / q(t) dt \\ & \stackrel{D}{=} \int_{1/n}^{\delta} \left| W\left(\frac{1}{n} \sum_{i=1}^{[nt]} (\sigma_i - 1)^2\right) \right|^p / q(t) dt \\ & \leq \int_{1/n}^{\delta} \left| W\left(\frac{1}{n} \sum_{i=1}^{[nt]} (\sigma_i - 1)^2\right) \right|^p / q\left(\frac{1}{n} \sum_{i=1}^{[nt]} (\sigma_i - 1)^2\right) dt \\ (1.4.8) \quad & \leq \int_0^{\delta} |W(t)|^p / q(t) dt. \end{aligned}$$

By Lemma 1.2.1 we obtain

$$(1.4.9) \quad \begin{aligned} & \int_{1/n}^{\delta} |n^{-1/2}(W([nt]) - W(nt))|^p / q(t) dt \\ & \leq O(1) \int_0^{\delta} t^{p/2} / q(t) dt. \end{aligned}$$

We note that

$$\begin{aligned} & \int_{1/n}^{\delta} |n^{-1/2}(S(nt) - W(nt))|^p / q(t) dt \\ & \leq 2^p \int_{1/n}^{\delta} |n^{-1/2}(S(nt) - T(nt))|^p / q(t) dt + 2^p \int_{1/n}^{\delta} |n^{-1/2}(T(nt) - W(nt))|^p / q(t) dt \end{aligned}$$

and also

$$\begin{aligned} & \int_{1/n}^{\delta} |n^{-1/2}(T(nt) - W(nt))|^p / q(t) dt \\ & \leq 2^p \int_{1/n}^{\delta} \left| n^{-1/2}(T(nt) - \sum_{i=1}^{[nt]} Y_i / \sigma_i) \right|^p / q(t) dt \\ & \quad + 2^p \int_{1/n}^{\delta} \left| n^{-1/2} \left(\sum_{i=1}^{[nt]} Y_i / \sigma_i - W(nt) \right) \right|^p / q(t) dt. \end{aligned}$$

Since $\delta > 0$ can be taken arbitrarily small, by combining (1.4.6)-(1.4.9) and using our assumption (1.4.2) as well as Lemma 1.3.A, we get

$$\int_{1/n}^1 |n^{-1/2}(S(nt) - W(nt))|^p / q(t) dt = o_P(1),$$

which together with (1.4.4) yields (1.4.1).

Next we assume (1.4.1) to hold. Since

$$\int_0^1 |n^{-1/2}(S(nt) - W(nt))|^p / q(t) dt \geq \int_0^{1/n} |n^{-1/2}W(nt)|^p / q(t) dt \stackrel{D}{=} \int_0^{1/n} |W(t)|^p / q(t) dt,$$

for n large enough there exist positive constants δ and K such that

$$P\left\{ \int_0^{\delta} |W(t)|^p / q(t) dt > K \right\} \leq 1/2.$$

hence the 0 - 1 law for a Wiener process implies

$$P\left\{\int_0^{\delta} |W(t)|^p/q(t)dt < \infty\right\} = 0 \text{ or } 1,$$

which in turn gives

$$\int_0^{\delta} |W(t)|^p/q(t)dt < \infty \quad \text{a.s. .}$$

Since

$$\int_0^1 |W(t)|^p dt < \infty \quad \text{a.s.}$$

and $q(t)$ is positive, i.e. $\inf_{\delta \leq t \leq 1} q(t) > 0$ for any $\delta \in (0, 1)$, we have also

$$\int_{\delta}^1 |W(t)|^p/q(t)dt < \infty \quad \text{a.s.}$$

Consequently

$$\int_0^1 |W(t)|^p/q(t)dt < \infty \quad \text{a.s.}$$

Now Lemma 1.3.A implies (1.4.2).

(b) We assume that (1.4.2) holds. Using (1.4.5) we get

$$(1.4.10) \quad \int_{\delta}^1 |n^{-1/2} S(nt)|^p/q(t)dt \xrightarrow{P} \int_{\delta}^1 |W(t)|^p/q(t)dt$$

for all $\delta \in (0, 1)$. By Lemma 1.3.A we have

$$(1.4.11) \quad \lim_{\delta \rightarrow 0} \int_0^{\delta} |W(t)|^p/q(t)dt = 0 \quad \text{a.s. .}$$

We note that

$$\begin{aligned} \int_0^{\delta} |n^{-1/2} S(nt)|^p/q(t)dt &\leq 2^p \int_0^{\delta} |n^{-1/2} W(nt)|^p/q(t)dt \\ &\quad + 2^p \int_0^{\delta} |n^{-1/2} (S(nt) - W(nt))|^p/q(t)dt. \end{aligned}$$

Using (1.4.1) and (1.4.11) we obtain

$$(1.4.12) \quad \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} P\left\{\int_0^{\delta} |n^{-1/2} S(nt)|^p/q(t)dt > \epsilon\right\} = 0$$

for all $\epsilon > 0$. Combining (1.4.10), (1.4.11) and (1.4.12) gives the result.

Assuming now that (1.4.3) holds we have

$$\int_0^1 |W(t)|^p/q(t)dt < \infty \quad \text{a.s.,}$$

which yields (1.4.2) by Lemma 1.3.A.

CHAPTER 2

Weighted partial sum processes under contiguous measures

2.0. Introduction

The notion of contiguity, which was introduced by Lucien Le Cam (1960) as a criterion of nearness of sequences of probability measures, has an enduring place among the fundamental concepts of Mathematical Statistics. From the point of view of mathematics, contiguity can be thought of as an asymptotic version of the idea of absolute continuity of measures. Concerning its role in statistics, Roussas (1972, Introduction) writes: "In addition to its purely mathematical interest, contiguity is a powerful and very useful tool in statistics, when one is concerned with asymptotic theory, leading to elegant derivations of the asymptotic properties of tests and estimates under less restrictive assumptions than usual. Of course, this presupposes that one subscribes to the usefulness of large sample theory results. Frequently, however, this is the only alternative open to the statistician in an unfriendly real world."

The concept of contiguity has been highly popularized in the area of rank tests through the work of Jaroslav Hájek and his associates. For details along these lines we refer to Chapter VI of Hájek and Šidák (1967). For a more global review of contiguity we refer to the monograph of Roussas (1972). More recent developments are covered in the book of Greenwood and Shiriyayev (1985) as well as in Le Cam (1986) and Le Cam and Yang (1990).

Essentially all the important results are already given in Le Cam (1960), which were then used, extended, modified and clarified by many authors. In addition to the works mentioned already, and their references, we found the papers Oosterhoff

and Van Zwet (1975), Hall and Loynes (1977) and the book of Shorack and Wellner (1986) of great help.

In Section 2.1 we introduce the concept of contiguity, discuss its relationship to some other modes of nearness of sequences of probability measures. The various characterizations of contiguity provide alternative methods to establishing the presence (or absence) of contiguity in a given case. We cite some examples for illustrative purposes, as well as a number of important results, based on assumption of contiguity. The real power of contiguity theory in applications is that convergence properties of statistic of interest can be determined under contiguous alternatives by proving their joint convergence together with the log-likelihood ratio process under the null hypothesis of i.i.d. sampling. Hence our focus on the celebrated Le Cam's third lemma.

In Section 2.2 we consider contiguity of product probability measures. We introduce Hellinger distance and present the necessary and sufficient conditions of Oosterhoff and Van Zwet (1975) for contiguity of product probability measures in terms of their marginals.

In Section 2.3 we present a parameterization of contiguity which was used by Khmaladze and Parjanadze (1986). Under such contiguity conditions, in Section 2.4 we establish weak convergence of partial sum processes. In Section 2.5 we extend this result to using weight functions and show that the weak convergence of weighted partial sum processes continues to hold under these contiguous assumptions for the same class of weight functions as in the case of i.i.d.r.v.'s (cf. also Szyszkowicz, 1991d, 1992a). The obtained results are optimal in terms of moment conditions and the class of weight functions.

2.1. The notion of contiguity

Let $\{(\mathcal{X}_n, \mathcal{A}_n)\}$, $n \geq 1$, be a sequence of measurable spaces and let P_n, Q_n be probability measures on $(\mathcal{X}_n, \mathcal{A}_n)$.

The sequence of probability measures $\{Q_n\}$ is said to be *contiguous* to the sequence $\{P_n\}$, denoted $\{Q_n\} \triangleleft \{P_n\}$, if for every sequence of sets $\{A_n\}$, $A_n \in \mathcal{A}_n$,

$$\lim_{n \rightarrow \infty} P_n(A_n) = 0 \text{ implies } \lim_{n \rightarrow \infty} Q_n(A_n) = 0.$$

This definition is not symmetric in P and Q . Initially, the concept of contiguity, as introduced by Le Cam (1960) and then adopted by Roussas (1972), was used to describe a symmetric relation - that may be termed as *mutual contiguity*. Namely we will say that $\{Q_n\}$ and $\{P_n\}$ are mutually contiguous, denoted $\{Q_n\} \triangleleft\triangleright \{P_n\}$, if both $\{Q_n\} \triangleleft \{P_n\}$ and $\{P_n\} \triangleleft \{Q_n\}$, i.e., if for every sequence of sets $\{A_n\}$, $A_n \in \mathcal{A}_n$,

$$\lim_{n \rightarrow \infty} P_n(A_n) = 0 \text{ if and only if } \lim_{n \rightarrow \infty} Q_n(A_n) = 0.$$

The one-sided definition is used by Hájek and Šidák (1967), Greenwood and Shirayev (1985) and also in the new book by Le Cam and Yang (1990). The notation is due to Witting and Nölle (1970) (cf. Oosterhoff and Van Zwet, 1975).

Remark 2.1.A. We note that contiguity implies that *any sequence of random variables $\{T_n\}$ which converges to zero in P_n -probability, converges to zero also in Q_n -probability*. This provides an alternative definition of contiguity and in fact it is used as such by Le Cam (1960), Le Cam and Yang (1990) and Roussas (1972).

Characterizations of contiguity are usually given in terms of some properties of the likelihood ratio sequence.

Having $\{P_n\}$ and $\{Q_n\}$ on $(\mathcal{X}_n, \mathcal{A}_n)$, let μ_n be a σ -finite measure dominating P_n and Q_n , i.e., P_n and Q_n are absolutely continuous with respect to μ_n for each n ($P_n \ll \mu_n$ and $Q_n \ll \mu_n$). Let X_n be the identity map from $\mathcal{X}_n$ onto $\mathcal{X}_n$ for each n . Let the random variables

$$p_n = \frac{dP_n}{d\mu_n} = \frac{dP_n}{d\mu_n}(X_n), \quad q_n = \frac{dQ_n}{d\mu_n} = \frac{dQ_n}{d\mu_n}(X_n)$$

be finite versions of the corresponding Randon-Nikodym derivatives, and let

$$\lambda_n = \begin{cases} q_n/p_n & \text{if } p_n > 0 \\ \infty & \text{if } p_n = 0 < q_n \\ 0 & \text{if } p_n = 0 = q_n. \end{cases}$$

Remark 2.1.B. If $Q_n \ll P_n$ then λ_n is simply a version of the Radon-Nikodym derivative dQ_n/dP_n .

Let F_n be the distribution function of $\lambda_n = \lambda_n(X_n)$ under P_n , i.e.

$$F_n(x) = P_n(\lambda_n \leq x).$$

The following result is called **Le Cam's first lemma** in Hájek and Šidák (1967).

Lemma 2.1.A. Assume that F_n converges to a distribution function F at its points of continuity and

$$\int_0^\infty x dF(x) = 1.$$

Then the sequence $\{Q_n\}$ is contiguous to $\{P_n\}$.

The asymptotic distribution F of the sequence of likelihoods $\{\lambda_n\}$ will "regularly" be log-normal. We say that a random variable Y has log-normal (μ, σ^2) distribution if $\log Y$ is a normal (μ, σ^2) random variable, that is $\log Y \stackrel{D}{=} N(\mu, \sigma^2)$. The condition under which we obtain

$$EY = \int x dF = 1$$

for a log-normal random variable Y is given by

$$\mu = -\frac{1}{2}\sigma^2.$$

Namely we have

$$EY = E \exp(\log Y) = \exp\left(\mu + \frac{1}{2}\sigma^2\right).$$

the value of the moment generating function of a $N(\mu, \sigma^2)$ random variable at 1, which is equal to 1 at $\mu = -\frac{1}{2}\sigma^2$. Hence, in this regular case, we have the following version of Le Cam's first lemma.

Corollary 2.1.A. *Let $\Lambda_n = \log \lambda_n$. If*

$$\Lambda_n \xrightarrow{D} N\left(-\frac{1}{2}\sigma^2, \sigma^2\right).$$

then $\{Q_n\}$ is contiguous to $\{P_n\}$.

In order to acquire some intuition for the connection between contiguity and the limiting expected value of λ_n being equal to 1, we note first that for each $n \geq 1$

$$(2.1.1) \quad E_n \lambda_n = \int \lambda_n dP_n \leq 1.$$

This follows immediately from the Lebesgue decomposition, which gives for all $A_n \in \mathcal{A}_n$

$$(2.1.2) \quad Q_n(A_n) = \int_{A_n} \lambda_n dP_n + Q_n(A_n \cap \{\lambda_n = \infty\}),$$

and, consequently,

$$\int_{A_n} \lambda_n dP_n \leq Q_n(A_n).$$

(cf. Greenwood and Shiryayev, 1985, p.8).

Also, when taking $A_n = \{\lambda_n \geq a\}$, we have

$$1 \geq \int_{\{\lambda_n \geq a\}} \lambda_n dP_n \geq aP_n(\lambda_n \geq a),$$

which gives

$$(2.1.3) \quad P_n(\lambda_n \geq a) \leq a^{-1}, \quad a > 0.$$

We recall also the following two facts related to uniform integrability (for the proof see, e.g., Billingsley, 1968).

Lemma 2.1.B. (a) *The sequence (X_n, P_n) is uniformly integrable, i.e.,*

$$\lim_{N \rightarrow \infty} \sup_n \int_{\{|X_n| \geq N\}} |X_n| dP_n = 0,$$

if and only if $\sup_n E_n |X_n| = \sup_n \int |X_n| dP_n < \infty$ and, as $n \rightarrow \infty$,

$$\int_{\{A_n\}} |X_n| dP_n \rightarrow 0 \quad \text{if} \quad P_n(A_n) \rightarrow 0, \quad A_n \in \mathcal{A}_n.$$

(b) *Let $X_n \xrightarrow{D} X$ under P_n . Then (X_n, P_n) is uniformly integrable if and only if*

$$E|X| = \int |X| dP < \infty$$

and

$$E_n |X_n| = \int |X_n| dP_n \rightarrow E|X| = \int |X| dP, \quad n \rightarrow \infty,$$

where P is the distribution of X .

Let us assume now that λ_n converges weakly under P_n and that the sequence $\{Q_n\}$ is contiguous to $\{P_n\}$. By (2.1.1) we have already

$$\sup_n E_n \lambda_n < \infty.$$

and from the Lebesgue decomposition (cf. (2.1.2))

$$E_n \lambda_n 1\{A_n\} \leq Q_n(A_n),$$

which $\rightarrow 0$ if $P_n(A_n) \rightarrow 0$, because of the contiguity $\{Q_n\} \triangleleft \{P_n\}$. Hence, Lemma 2.1.B (a) implies that (λ_n, P_n) is uniformly integrable. Since we assume the weak convergence of λ_n under P_n , to show that the limiting expected value is equal to 1, by Lemma 2.1.B (b) it remains to show that

$$\lim_{n \rightarrow \infty} E_n \lambda_n = 1.$$

Using again the Lebesgue decomposition, we have

$$\begin{aligned} E_n \lambda_n &= \int \lambda_n dP_n = \int_{\{\lambda_n \geq n\}} \lambda_n dP_n + \int_{\{\lambda_n < n\}} \lambda_n dP_n \\ &= \int_{\{\lambda_n \geq n\}} \lambda_n dP_n + Q_n(\lambda_n < n) \\ &= 1 + \int_{\{\lambda_n \geq n\}} \lambda_n dP_n - Q_n(\lambda_n \geq n). \end{aligned}$$

Since (λ_n, P_n) is uniformly integrable,

$$\lim_{n \rightarrow \infty} \int_{\{\lambda_n \geq n\}} \lambda_n dP_n = 0.$$

Further, by (2.1.3), $P_n(\lambda_n \geq n) \leq n^{-1}$, which together with the contiguity $\{Q_n\} \triangleleft \{P_n\}$ implies

$$\lim_{n \rightarrow \infty} Q_n(\lambda_n \geq n) = 0.$$

Consequently, on account of the above decomposition of $E_n \lambda_n$, we conclude

$$\lim_{n \rightarrow \infty} E_n \lambda_n = \lim_{n \rightarrow \infty} \int \lambda_n dP_n = 1.$$

What we really obtained here is the following fact (cf. Theorem 1 of Hall and Loynes, 1977; Greenwood and Shiryayev, 1985; p. 38), an "if and only if version" of Lemma 2.1.A.

Lemma 2.1.C. *Assume that $\lambda_n \xrightarrow{D} \lambda$ under P_n and let P be the distribution of λ . In order that $\{Q_n\}$ be contiguous to $\{P_n\}$ it is necessary and sufficient that $E\lambda = \int \lambda dP = 1$.*

Postulating the weak convergence of λ_n under P_n in this lemma is not necessary for concluding $\{Q_n\}$ to be contiguous to $\{P_n\}$. Rather, for any subsequence of λ_n converging weakly with respect to the corresponding subsequence of $\{P_n\}$, the limiting expected value being equal to 1 is essential. Other conditions equivalent to contiguity in terms of λ_n can be summarized as follows (cf. Theorem 2.1 of Le Cam, 1960; Greenwood and Shirayev, 1985, p.31).

Theorem 2.1.A. *The following conditions are equivalent:*

- 1) $\{Q_n\}$ is contiguous to $\{P_n\}$
- 2) (λ_n, P_n) is uniformly integrable and

$$Q_n(\Lambda_n = \infty) \rightarrow 0 \quad \text{as } n \rightarrow \infty$$

$$3) \lambda_n \text{ is tight under } \{Q_n\}, \text{ i.e. } \lim_{N \rightarrow \infty} \limsup_{n \rightarrow \infty} Q_n(\lambda_n \geq N) = 0$$

$$4) \log \lambda_n \text{ is tight under } \{Q_n\}$$

5) *For every subsequence $\{\lambda_{n_k}\}$ converging weakly with respect to $\{P_{n_k}\}$, its weak limit λ is such that $E\lambda = 1$.*

Remark 2.1.B. (cf. Greenwood and Shirayev, 1985, p.42). We note that λ_n is always tight under $\{P_n\}$. This can be easily seen from the bound (2.1.3) according to which

$$P_n(\lambda_n \geq N) \leq N^{-1} \text{ for all } n \geq 1 \text{ and } N > 0.$$

In fact one of the properties of contiguity is that it preserves tightness. More exactly, let $\{Q_n\}$ be contiguous to $\{P_n\}$ and suppose that X_n is tight under $\{P_n\}$, i.e.

$$\lim_{N \rightarrow \infty} \limsup_n P_n(|X_n| \geq N) = 0.$$

Then X_n is also tight under $\{Q_n\}$, i.e.

$$\lim_{N \rightarrow \infty} \limsup_n Q_n(|X_n| \geq N) = 0.$$

The concept of contiguity is a concept expressing "closeness" or "nearness" between the sequences of probability measures $\{P_n\}$ and $\{Q_n\}$ in the sense of Definition 2.1.A. In the case of mutual contiguity, $\{Q_n\} \triangleleft \{P_n\}$ and $\{P_n\} \triangleleft \{Q_n\}$, one can say that it expresses nearness of these probability measures in the sense that they produce distributions in which random elements either do escape to infinity simultaneously, or they do not do so simultaneously.

When relating this notion to other, more familiar, concepts of the same nature, it can be shown that "nearness" of two sequences of probability measures expressed by mutual contiguity is weaker than that expressed by convergence in L_1 -norm. Let η and ν be two arbitrary finite positive measures on a σ -field $\mathcal{A}$. Let μ denote a measure, with respect to which both η and ν are absolutely continuous, for example take $\mu = \eta + \nu$. The L_1 norm of $\eta - \nu$, denoted by $\|\eta - \nu\|$, is defined by the integral

$$\|\eta - \nu\| = \int |f - g| d\mu,$$

where $f = d\eta/d\mu$ and $g = d\nu/d\mu$ are the Radon-Nikodym derivatives of η and ν with respect to μ (cf., e.g. Le Cam, 1960). If η and ν are probability measures, then $\|\eta - \nu\|$ is equal to $2\sup\{|\eta(A) - \nu(A)|; A \in \mathcal{A}\}$ on account of Scheffé's (1947) theorem (cf. Theorem 1 in Devroye and Györfi, 1985, Chap. 1). Of course, if

$\|P_n - Q_n\| \rightarrow 0$, then the sequences $\{P_n\}$ and $\{Q_n\}$ are as close together as they can be, apart from having identical elements.

Lemma 2.1.D. (cf. Roussas, 1972, p.8). *If $\|P_n - Q_n\| \rightarrow 0$ then $\{P_n\}$ and $\{Q_n\}$ are mutually contiguous.*

The converse is not true.

Another distance that is more appropriate to describe contiguity is the Hellinger distance. If P and Q are positive measures on a σ -field $\mathcal{A}$, their Hellinger distance $H(P, Q)$ is defined by (cf. Le Cam and Yang, 1990, p.25)

$$H^2(P, Q) = \frac{1}{2} \int ((dP)^{1/2} - (dQ)^{1/2})^2.$$

For probability measures this can be also written as

$$H^2(P, Q) = 1 - \int (fg)^{1/2} d\mu,$$

where $f = dP/d\mu$, $g = dQ/d\mu$ for any measure μ that dominates both P and Q . Since the Hellinger distance is very useful when passing to direct product of probability measures, we will discuss it more closely in the next section.

To illustrate the concept of contiguity itself, we now give a few examples (cf. Roussas, 1972, Chap. 1).

Example 2.1.A. Let $(\mathcal{X}, \mathcal{A}_n) = (\mathbf{R}, \mathcal{B})$, $P_n = U(-1/n, 1)$, the uniform measure over $(-1/n, 1)$ and $Q_n = U(0, 1 + 1/n)$, the uniform measure over $(0, 1 + 1/n)$. Then, if

$$f_n = dP_n/d\ell \quad \text{and} \quad g_n = dQ_n/d\ell,$$

where ℓ is the Lebesgue measure on $\mathbf{R}$, we have

$$\|P_n - Q_n\| = \int |f_n - g_n| d\ell = 2/(n + 1) \rightarrow 0.$$

Thus $\{P_n\}$ and $\{Q_n\}$ are mutually contiguous, by Lemma 2.1.D.

The above example shows also that mutual contiguity need not imply mutual absolute continuity (neither for all n nor only for all sufficiently large n). On the other hand, mutual absolute continuity need not imply contiguity either, as demonstrated by the following example.

Example 2.1.B. Let

$$(\mathcal{X}, \mathcal{A}_n) = (\mathbf{R}, \mathcal{B}), \quad P_n = N(\mu_n, 1) \quad \text{and} \quad Q_n = N(\mu'_n, 1)$$

where $\mu_n \rightarrow -\infty$ and $\mu'_n \rightarrow \infty$. (The notation $N(\mu, \sigma^2)$ is used for the distribution of a normal (μ, σ^2) random variable or to denote such a random variable itself). Then clearly $\{P_n\}$ and $\{Q_n\}$ are mutually absolutely continuous for each n (i.e. $P_n \ll Q_n$ and $Q_n \ll P_n$ for all $n \geq 1$). Consider the set A_n defined by $A_n = (\mu_n - 1, \mu_n + 1)$. Then $P_n(A_n) = c$, a positive constant ($c \approx 0.68$), so that $P_n(A_n)$ does not tend to zero. But, clearly $Q_n(A_n) \rightarrow 0$. Thus, $\{P_n\}$ and $\{Q_n\}$ cannot be contiguous.

From this example one might (wrongly) conclude that the absence of contiguity of $\{P_n\}$ and $\{Q_n\}$ is due to the lack of tightness. (It is clear that the above sequences are not tight). That tightness is not a necessary condition for contiguity is shown in Example 2.1.C, where the sequences involved are not tight (for a special choice of the parameters) and yet contiguous. In the same example and for another choice of the parameters, it is also seen that tightness is not a sufficient condition for contiguity.

Example 2.1.C. Let $(\mathcal{X}, \mathcal{A}_n) = (\mathbf{R}, \mathcal{B})$. We take

$$P_n = N(\mu_n, \sigma_n^2) \quad \text{and} \quad Q_n = N(\mu'_n, \sigma_n^2).$$

Then we have

$$\log \lambda_n = \Lambda_n = \frac{\mu'_n - \mu_n}{\sigma_n^2} X + \frac{\mu_n^2 - \mu_n'^2}{2\sigma_n^2},$$

so that

$$\Lambda_n \stackrel{D}{=} N\left(-\frac{(\mu'_n - \mu_n)^2}{2\sigma_n^2}, \frac{(\mu'_n - \mu_n)^2}{\sigma_n^2}\right) \text{ under } P_n,$$

$$\Lambda_n \stackrel{D}{=} N\left(\frac{(\mu'_n - \mu_n)^2}{2\sigma_n^2}, \frac{(\mu'_n - \mu_n)^2}{\sigma_n^2}\right) \text{ under } Q_n.$$

Suppose that $\left\{\frac{\mu'_n - \mu_n}{\sigma_n}\right\}$ is bounded, and let the subsequence $\left\{\frac{\mu'_m - \mu_m}{\sigma_m}\right\}$ converge to σ , say. Then, clearly,

$$\Lambda_n \xrightarrow{D} N\left(-\frac{1}{2}\sigma^2, \sigma^2\right) \text{ under } P_n \text{ and } \Lambda_n \xrightarrow{D} N\left(\frac{1}{2}\sigma^2, \sigma^2\right) \text{ under } Q_n.$$

Hence $\{P_n\}$ and $\{Q_n\}$ are mutually contiguous. As a special case, let $\sigma_n = 1$, $\mu_n \rightarrow \mu$, $\mu'_n \rightarrow \mu'$, where both μ, μ' are finite and $\mu \neq \mu'$. Then $\{P_n\}$ and $\{Q_n\}$ are mutually contiguous and, clearly, $\|P_n - Q_n\| \rightarrow 0$. Also, let $\sigma_n = 1$ as above, and let $\mu'_n = \mu_n + c$, where c is a constant, $c \neq 0$ and $\mu_n \rightarrow -\infty$. Then again $\{P_n\}, \{Q_n\}$ are contiguous, but clearly, they are not tight.

Finally, let $\mu'_n = \mu_n + c$, with $c \neq 0$ (a constant), let $\mu_n \rightarrow \mu$, finite, and let $\sigma_n \rightarrow 0$. Then $\{P_n\}, \{Q_n\}$ are not contiguous and yet they are tight.

Example 2.1.D. Let $(\mathcal{X}, \mathcal{A}_n) = (\mathbb{R}, \mathcal{B})$. Let $\{P_n\}$ and $\{Q_n\}$ be Poisson measures given by

$$P_n(\{x\}) = \exp(-\theta_n) \frac{\theta_n^x}{x!},$$

$$Q_n(\{x\}) = \exp(-\theta'_n) \frac{\theta_n'^x}{x!}, \quad x = 0, 1, \dots$$

Then, with $a_n = \log(\theta'_n/\theta_n)$ and $b_n = \theta_n - \theta'_n$, we have

$$\log \lambda_n = \Lambda_n = a_n X + b_n,$$

so that

$$\begin{aligned} P_n(\Lambda_n = y_n) &= \exp(-\theta_n) \frac{\theta_n^x}{x!}; \quad y_n = a_n x + b_n, \quad x = 0, 1, \dots, \\ Q_n(\Lambda_n = y_n) &= \exp(-\theta'_n) \frac{\theta'_n^x}{x!}; \quad y_n = a_n x + b_n, \quad x = 0, 1, \dots. \end{aligned}$$

Suppose that $\{\theta_n\}$ and $\{\theta'_n\}$ are bounded from above and away from 0. Then, if $\{m\}$ is a subsequence of $\{n\}$ such that $\theta_m \rightarrow \theta$, $\theta'_m \rightarrow \theta'$, it follows that $\Lambda_m \xrightarrow{P} \Lambda$ under P_m and $\Lambda_m \xrightarrow{P} \Lambda'$ under Q_m , where the distribution P of Λ and the distribution Q of Λ' are over the points

$$y = ax + b, \quad x = 0, 1, \dots, \text{ with } a = \log(\theta'/\theta), \quad b = \theta - \theta'.$$

Therefore Λ_n is tight under $\{P_n\}$ and also under $\{Q_n\}$. Hence $\{P_n\}$ and $\{Q_n\}$ are mutually contiguous on account of Theorem 2.1.A ((1) $\Leftrightarrow$ (4)). Contiguity can be also concluded by noting that the limiting expected value of λ_n is equal to 1, namely

$$\begin{aligned} &\int \exp y \, dP(y) \\ &= \sum_{x=0}^{\infty} \exp(ax + b) \exp(-\theta) \frac{\theta^x}{x!} \\ &= \exp(-\theta') \sum_{x=0}^{\infty} \frac{1}{x!} (\theta \exp a)^x \\ &= \exp(-\theta') \exp(\theta \exp a) = 1. \end{aligned}$$

One of the most useful consequences of the contiguity properties is summarized in the following proposition (cf. Le Cam 1960; Le Cam and Yang, 1990, p. 22).

Proposition 2.1.A. *Let $\{T_n\}$ be a sequence of random variables. Assume that $\{Q_n\}$ is contiguous to $\{P_n\}$ and the (joint) distributions of (T_n, Λ_n) under*

$\{P_n\}$ tend to a limit F on $(-\infty, +\infty) \times [-\infty, \infty]$, where $\Lambda_n = \log \lambda_n$. Then the distributions of (T_n, Λ_n) under $\{Q_n\}$ tend to a limit G and

$$G(dt, d\lambda) = e^\lambda F(dt, d\lambda)$$

Note. The result is also valid for variables T_n that are vector valued, or for random elements that take values in Polish spaces. For more general results see Le Cam (1986), page 90.

Proposition 2.1.A applies in particular to Λ_n itself by taking $T_n = 1$. Suppose that the distribution of Λ_n under P_n converge to the Gaussian $F = N(\mu, \sigma^2)$. Then under Q_n ($\{Q_n\}$ is contiguous to $\{P_n\}$) the distribution of Λ_n under Q_n will converge to G such that $G(d\lambda) = e^\lambda F(d\lambda)$. Namely, we have

$$\begin{aligned} \int \Lambda dG &= \int \exp \Lambda dF = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} \exp \left[x - \frac{1}{2\sigma^2}(x - \mu)^2 \right] dx \\ &= \exp \left(\mu + \frac{1}{2}\sigma^2 \right) \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} \exp \left[-\frac{1}{2\sigma^2}(x - (\mu + \sigma^2))^2 \right] dx. \end{aligned}$$

Since G must be a probability measure, it follows that $\mu + \frac{1}{2}\sigma^2 = 0$ and that $G = N(\mu + \sigma^2, \sigma^2)$, i.e.

$$\Lambda_n \xrightarrow{D} N\left(-\frac{1}{2}\sigma^2, \sigma^2\right) \text{ under } P_n$$

and

$$\Lambda_n \xrightarrow{D} N\left(\frac{1}{2}\sigma^2, \sigma^2\right) \text{ under } Q_n.$$

Consider another particular case. Assume that (T_n, Λ_n) converge under P_n to a bivariate normal distribution. If $\{Q_n\}$ is contiguous to $\{P_n\}$ then the distributions of (T_n, Λ_n) under Q_n also tend to a limit, which is a bivariate normal distribution that has the same covariance structure as the limit of (T_n, Λ_n) under

P_n . The vector of means is changed by adding the appropriate covariance terms of the limiting distribution.

This result is very useful in statistical applications and such a form is known as **Le Cam's third lemma** (cf. Hájek and Šidak, 1967).

We say that a pair (T_n, Λ_n) is asymptotically jointly normal $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \sigma_{12})$ if it converges in distribution to a normal vector (Z_1, Z_2) such that $EZ_i = \mu_i$, $\text{var}Z_i = \sigma_i^2$, $i = 1, 2$, and $\text{cov}(Z_1, Z_2) = \sigma_{12}$

Lemma 2.1.E. *Assume that the pair (T_n, Λ_n) under P_n is asymptotically jointly normal $(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \sigma_{12})$ with $\mu_2 = -\frac{1}{2}\sigma_2^2$. Then T_n under Q_n is asymptotically normal $(\mu_1 + \sigma_{12}, \sigma_1^2)$, i.e. if*

$$\begin{pmatrix} T_n \\ \Lambda_n \end{pmatrix} \xrightarrow{\mathcal{D}} N\left(\mu_1, -\frac{1}{2}\sigma_2^2, \sigma_1^2, \sigma_2^2, \sigma_{12}\right) \text{ under } P_n$$

then

$$T_n \xrightarrow{\mathcal{D}} N\left(\mu_1 + \sigma_{12}, \sigma_1^2\right) \text{ under } Q_n.$$

We note that even if contiguity is not assumed here explicitly, it is implied by the requirement that $\Lambda_n \xrightarrow{\mathcal{D}} N\left(-\frac{1}{2}\sigma^2, \sigma^2\right)$, due to Le Cam's first lemma (cf. Corollary 2.1.A).

The multivariate version of Le Cam's third lemma reads as follows (cf. Shorack and Wellner, 1986, p.156).

Lemma 2.1.F. *If*

$$\begin{pmatrix} \Lambda_n \\ \vec{T}_n \end{pmatrix} \xrightarrow{\mathcal{D}} N\left(\begin{bmatrix} -\frac{1}{2}\sigma^2 \\ \vec{\mu} \end{bmatrix}, \begin{bmatrix} \sigma^2 & \vec{\sigma} \\ \vec{\sigma} & \Sigma \end{bmatrix}\right) \text{ under } P_n$$

then

$$\vec{T}_n \xrightarrow{\mathcal{D}} N([\vec{\mu} + \vec{\sigma}], [\Sigma]) \text{ under } Q_n.$$

Note. It is clear that when T_n is a stochastic process then the weak convergence of T_n under Q_n can be derived if we have the joint weak convergence of Λ_n and T_n under P_n with a specified limit. Namely, the multivariate version of Le Cam's third lemma will give us limits of all finite dimensional distributions under Q_n , and tightness will be implied by the weak convergence of T_n under P_n and contiguity of $\{Q_n\}$ to $\{P_n\}$ (cf. Remark 2.1.B).

There are other situations (see Le Cam, 1960; 1986) where convergence of the distributions of (T_n, Λ_n) under P_n to F yields a limit, say G , of the distributions of (T_n, Λ_n) under Q_n that is easily identifiable. For instance if F is infinitely divisible and the corresponding marginal of T in F is Gaussian, then the first marginal of G is also normal with the same variance as in F and with a mean changed by adding the covariance term from the normal component of F .

2.2. Contiguity of product measures

In this Section we characterize contiguity of product probability measures in terms of their marginals. The main purpose is to state necessary and sufficient conditions for the asymptotic normality of the log-likelihood ratio which we will need later on. Here we follow closely Oosterhoff and Van Zwet (1975) and refer for all details and proofs to them.

For $n = 1, 2, \dots$ let $(\mathcal{X}_{n1}, \mathcal{A}_{n1}), \dots, (\mathcal{X}_{nn}, \mathcal{A}_{nn})$ be arbitrary measurable spaces. Let P_{ni} and Q_{ni} be probability measures defined on $(\mathcal{X}_{ni}, \mathcal{A}_{ni})$, $i = 1, \dots, n$; $n = 1, 2, \dots$ and let

$$P_n^{(n)} = \prod_{i=1}^n P_{ni} \text{ and } Q_n^{(n)} = \prod_{i=1}^n Q_{ni}, \quad n = 1, 2, \dots$$

denote their product probability measures. For each i and n let X_{ni} be the identity map from $\mathcal{X}_{ni}$ onto $\mathcal{X}_{ni}$. Then P_{ni} and Q_{ni} represent two possible distributions

of the random element X_n , as well as the probability measures of the underlying probability space. Obviously $X_1, \dots, X_n$ are independent under both $P_n^{(n)}$ and $Q_n^{(n)}$ ($n = 1, 2, \dots$).

The sequence $\{Q_n^{(n)}\}$ is said to be *contiguous* to the sequence $\{P_n^{(n)}\}$ if

$$\lim_{n \rightarrow \infty} P_n^{(n)}(A_n) = 0 \text{ implies } \lim_{n \rightarrow \infty} Q_n^{(n)}(A_n) = 0$$

for any sequence of measurable sets A_n .

We introduce the Hellinger distance $H(P, Q)$ between two probability measures P and Q on the same σ -field, defined by

$$H(P, Q) = \left\{ \int (p^{1/2} - q^{1/2})^2 d\mu \right\}^{1/2} = \left\{ 2 - 2 \int p^{1/2} q^{1/2} d\mu \right\}^{1/2},$$

where $p = dP/d\mu$, $q = dQ/d\mu$ and μ is any σ -finite measure dominating $P + Q$.

The Hellinger distances of the product measures and their marginals are connected by the relationship

$$H^2(P_n^{(n)}, Q_n^{(n)}) = 2 - 2 \prod_{i=1}^n \left\{ 1 - \frac{1}{2} H^2(P_{ni}, Q_{ni}) \right\}.$$

The following two relations, which connect the Hellinger distance and contiguity, can be shown.

1) $\sum_{i=1}^n H^2(P_{ni}, Q_{ni}) = o(1)$ as $n \rightarrow \infty$, implies that $\{P_n^{(n)}\}$ and $\{Q_n^{(n)}\}$ are mutually contiguous.

2) The relation that $\{Q_n\}$ is contiguous to $\{P_n\}$ implies

$$\sum_{i=1}^n H^2(P_{ni}, Q_{ni}) = O(1) \text{ for } n \rightarrow \infty.$$

The necessary and sufficient conditions for contiguity of the product measures were established by Oosterhoff and Van Zwet (1975) via adding another condition to $\sum_{i=1}^n H^2(P_{ni}, Q_{ni}) = O(1)$, $n \rightarrow \infty$, as will now be seen in Theorem 2.2.A.

Let μ_{ni} be a σ -finite measure on $(\mathcal{X}_{ni}, \mathcal{A}_{ni})$ dominating $P_{ni} + Q_{ni}$ and write $p_{ni} = dP_{ni}/d\mu_{ni}$ and $q_{ni} = dQ_{ni}/d\mu_{ni}$ ($i = 1, \dots, n$; $n = 1, 2, \dots$).

Theorem 2.2.A. $\{Q_n^{(n)}\}$ is contiguous to $\{P_n^{(n)}\}$ if and only if

$$(2.2.1) \quad \limsup_{n \rightarrow \infty} \sum_{i=1}^n H^2(P_{ni}, Q_{ni}) < \infty$$

and

$$(2.2.2) \quad \lim_{n \rightarrow \infty} \sum_{i=1}^n Q_{ni}(q_{ni}(X_{ni})/p_{ni}(X_{ni}) \geq c_n) = 0$$

whenever $c_n \rightarrow \infty$.

Corollary 2.2.A. $\{P_n^{(n)}\}$ and $\{Q_n^{(n)}\}$ are mutually contiguous if and only if (2.2.1) and (2.2.2) are satisfied and

$$\lim_{n \rightarrow \infty} \sum_{i=1}^n P_{ni}(p_{ni}(X_{ni})/q_{ni}(X_{ni}) \geq c_n) = 0$$

whenever $c_n \rightarrow \infty$.

Since asymptotic normality of the log-likelihood ratio plays a crucial role in applications, we define

$$\Lambda_n = \sum_{i=1}^n \log\{q_{ni}(X_{ni})/p_{ni}(X_{ni})\}, \quad n = 1, 2, \dots,$$

and state necessary and sufficient conditions as proved by Oosterhoff and Van Zwet (1975) for

$$\Lambda_n \xrightarrow{D} N\left(-\frac{1}{2}\sigma^2, \sigma^2\right) \quad \text{under } P_n^{(n)},$$

which by Le Cam's first lemma implies that $\{Q_n^{(n)}\}$ is contiguous to $\{P_n^{(n)}\}$.

Theorem 2.2.B. For any $\sigma \geq 0$

$$(2.2.3) \quad \Lambda_n \xrightarrow{P} N\left(-\frac{1}{2}\sigma^2, \sigma^2\right) \text{ under } P_n^{(n)}$$

and

$$(2.2.4) \quad \lim_{n \rightarrow \infty} \max_{1 \leq i \leq n} P_{n_i}(|\log\{q_{n_i}(X_{n_i})/p_{n_i}(X_{n_i})\}| \geq \epsilon) = 0$$

for every $\epsilon > 0$ if and only if

$$(2.2.5) \quad \lim_{n \rightarrow \infty} \sum_{i=1}^n H^2(P_{n_i}, Q_{n_i}) = \frac{1}{4}\sigma^2.$$

$$(2.2.6) \quad \lim_{n \rightarrow \infty} \sum_{i=1}^n Q_{n_i}(q_{n_i}(X_{n_i})/p_{n_i}(X_{n_i}) \geq 1 + \epsilon) = 0,$$

$$(2.2.7) \quad \lim_{n \rightarrow \infty} \sum_{i=1}^n P_{n_i}(p_{n_i}(X_{n_i})/q_{n_i}(X_{n_i}) \geq 1 + \epsilon) = 0,$$

or equivalently, if and only if (2.2.5) holds and for every $\epsilon > 0$

$$(2.2.8) \quad \lim_{n \rightarrow \infty} \sum_{i=1}^n \int_{|q_{n_i} - p_{n_i}| \geq \epsilon p_{n_i}} (q_{n_i}^{1/2} - p_{n_i}^{1/2})^2 d\mu_{n_i} = 0.$$

The intrinsic nature of this theorem is that it is a Lindeberg-Feller type central limit theorem for Λ_n whose summands are to be asymptotically uniformly negligible (cf. (2.2.4)).

It is sometimes more convenient to study the asymptotic normality of Λ_n via that of the statistic

$$W_n = 2 \sum_{i=1}^n \{(q_{n_i}(X_{n_i})/p_{n_i}(X_{n_i}))^{1/2} - 1\}$$

whose summands always have finite variances. The following lemma, due to Le Cam (cf. Hájek and Šidák, 1967, p.205), shows that asymptotic normality of

Λ_n under $P_n^{(n)}$ may be established by proving asymptotic normality of W_n and assuming asymptotic negligibility of summands à la (2.2.4) again.

Le Cam's second lemma. *Assume that*

$$\lim_{n \rightarrow \infty} \max_{1 \leq i \leq n} P_{n,i} \left(\left| \frac{q_{n,i}(X_i)}{p_{n,i}(X_i)} - 1 \right| > \epsilon \right) = 0$$

and that

$$W_n \xrightarrow{D} N\left(-\frac{1}{4}\sigma^2, \sigma^2\right) \quad \text{under } P_n^{(n)}.$$

Then

$$\Lambda_n \xrightarrow{D} N\left(-\frac{1}{2}\sigma^2, \sigma^2\right) \quad \text{under } P_n^{(n)}.$$

2.3. A parametrization of contiguity

Let $X_1, X_2, \dots$ be independent random variables. We wish to test the null hypothesis

$$H_0 : X_i, \ 1 \leq i \leq n, \text{ have the same distribution } F,$$

versus the alternative hypothesis

$$H_1 : X_i, \ 1 \leq i \leq n, \text{ have the respective distribution functions } F_{i,n},$$

where we assume that all $F_{i,n}$ are absolutely continuous with respect to the distribution function F and

$$\begin{aligned} \left[\frac{dF_{i,n}}{dF}(F^{-1}(u)) \right]^{1/2} &= 1 + \frac{1}{2n^{1/2}} h_n(t, F^{-1}(u)) \\ (2.3.1) \qquad \qquad \qquad &\equiv 1 + \frac{1}{2n^{1/2}} g_n(t, u), \qquad \frac{i-1}{n} < t \leq \frac{i}{n}, \end{aligned}$$

where $F^{-1}(u) = \inf\{x : F(x) \geq u\}$, $0 < u \leq 1$, $F^{-1}(0) = F^{-1}(0+)$. We assume also that there exists a function $g \in L^2[0, 1]^2$ such that

$$(2.3.2) \qquad \int_0^1 g(t, u) \, du = 0 \quad \text{for almost all } t \in [0, 1]$$

and

$$(2.3.3) \quad \int_0^1 \int_0^1 \left[g_n(t, u) - g\left(\frac{[nt]}{n}, u\right) \right]^2 du dt \rightarrow 0, \quad n \rightarrow \infty.$$

The sequence of direct products $F_{1n} \times \dots \times F_{nn}$, $n = 1, 2, \dots$ is contiguous to the sequence $F \times \dots \times F$. One way of seeing this is via Corollary 2.1.A since, as it will be shown (cf. proof of Lemma 2.3.1), the log-likelihood ratio is converging weakly to $N\left(-\frac{1}{2}\sigma^2, \sigma^2\right)$. This description of contiguous alternatives was used by Khmaladze and Parjanadze (1986) in the context of studying changepoint problems using linear statistics of sequential ranks.

We assume without loss of generality that under H_0 the X_i are of the form $F^{-1}(\bar{U}_i)$, where the $\bar{U}_i$ are independent identically distributed uniform $-(0, 1)$ random variables, and so $X_i \stackrel{D}{=} F^{-1}(\bar{U}_i)$, $i = 1, 2, \dots$.

Define the centered log-likelihood ratio process $L_n(\cdot)$ by

$$(2.3.4) \quad L_n(t) = \sum_{i=1}^{[nt]} \left\{ \log \frac{dF_{1n}}{dF}(F^{-1}(\bar{U}_i)) - E \log \frac{dF_{1n}}{dF}(F^{-1}(\bar{U}_i)) \right\}$$

Let L be a Gaussian process with mean zero and covariance function $Q_L(t_1 \wedge t_2)$, where

$$(2.3.5) \quad Q_L(t) = \int_0^t \int_0^1 g^2(s, u) du ds.$$

Lemma 2.3.1. *Let $X_1, X_2, \dots$ be independent random variables. We assume that H_0 holds and $\int x dF = 0$, $\int x^2 dF = 1$. Then*

$$L_n(t) \xrightarrow{D} L(t) \quad \text{in } D[0, 1].$$

Proof. For proving the weak convergence of L_n we use the results of Oosterhoff and Van Zwet (1975) as summarized in Section 2.2. In terms of the

parametrization of alternatives of H_1 used here (see (2.3.1)), the conditions of their Theorem 2 (cf. Theorem 2.2.B) for having

$$\Lambda_n \xrightarrow{D} N\left(-\frac{1}{2}\sigma^2, \sigma^2\right),$$

where

$$\Lambda_n = \sum_{i=1}^n \log \frac{dF_{in}}{dF}(F^{-1}(\bar{U}_i)),$$

reduce to showing (cf. 2.2.5) and (2.2.8))

$$(2.3.6) \quad \lim_{n \rightarrow \infty} \sum_{i=1}^n \int_0^1 \frac{g_n^2(t, u)}{4n} du = \frac{1}{4}\sigma^2,$$

and

$$(2.3.7) \quad \lim_{n \rightarrow \infty} \sum_{i=1}^n \int_{\left|\frac{t}{\sqrt{n}} + \frac{u^2}{4n}\right| \geq \epsilon} \frac{g_n^2(t, u)}{4n} du = 0, \text{ for any } \epsilon > 0.$$

The latter, in turn, are easily seen to be true with $\sigma^2 = \int_0^1 \int_0^1 g^2(t, u) dudt$ (cf. (2.3.5)) by (2.3.2) and (2.3.3). We notice that, due to having (2.3.1) and (2.3.3), $L_n(t)$ defined by (2.3.4) is a process of partial sums of independent random variables with mean zero and finite variances for which Lindeberg's condition holds. Now we can use Problem 1 of page 77 in Billingsley (1968) to conclude that $L_n \xrightarrow{D} L$.

2.4. Partial sums under contiguous measures

In this section we consider the asymptotic behaviour of partial sum processes under the sequence of contiguous alternatives of H_1 as parametrized in (2.3.1)–(2.3.3). Let $\{W(t), 0 \leq t \leq 1\}$ be a standard Wiener process.

Theorem 2.4.1. *Let $X_1, X_2, \dots$ be independent random variables and $S(nt) = X_1 + \dots + X_{[nt]}$. We assume that under H_0 $\int x dF = 0$ and $\int x^2 dF = 1$. Then, as $n \rightarrow \infty$, under H_1 we have*

$$n^{-1/2} S(nt) \xrightarrow{D} W(t) + C(t) \quad \text{in } D[0, 1],$$

where

$$C(t) = \int_0^t \int_0^1 g(s, u) F^{-1}(u) du ds.$$

In order to prove Theorem 2.4.1 we define the centered log-likelihood ratio process $L_n(\cdot)$ as in (2.3.4), namely

$$L_n(t) = \sum_{i=1}^{[nt]} \left\{ \log \frac{dF_{in}}{dF} (F^{-1}(U_i)) - E \log \frac{dF_{in}}{dF} (F^{-1}(U_i)) \right\},$$

where $\bar{U}_i$, $i = 1, 2, \dots$ are i.i.d. uniform-(0, 1) random variables.

Let (W, L) be the two-dimensional Gaussian process, where W is a standard Wiener process, L is a Gaussian process with mean zero and covariance function $Q_L(t_1 \wedge t_2)$, where

$$(2.4.1) \quad Q_L(t) = \int_0^t \int_0^1 g^2(s, u) du ds,$$

and the covariance function of W and L is

$$(2.4.2) \quad Q_{W,L}(t_1, t_2) = EW(t_1)L(t_2) = C(t_1 \wedge t_2)$$

with

$$(2.4.3) \quad C(t) = \int_0^t \int_0^1 g(s, u) F^{-1}(u) du ds.$$

Lemma 2.4.1. *Let $X_1, X_2, \dots$ be i.i.d.r.v.'s with distribution F such that $\int x dF = 0$, $\int x^2 dF = 1$ and $S(nt) = X_1 + \dots + X_{[nt]}$. Then, as $n \rightarrow \infty$*

$$\left(n^{-1/2} S(nt), L_n(t) \right) \xrightarrow{\mathcal{D}} (W(t), L(t)) \quad \text{in } D^2[0, 1].$$

Proof. Put $S_n(t) = n^{-1/2} S(nt)$. The weak convergence of both marginal processes S_n and L_n to their desired limits follows from Donsker's theorem and Lemma 2.3.1, respectively. Because of the structure of both processes S_n and

L_n , it is easy to compute their covariance function, which is seen to converge to $Q_{W,L}(t_1, t_2)$, defined by (2.4.2). It is also clear that the sum $S_n + L_n$ converges weakly to $W + L$, which then yields the weak convergence of the vector (S_n, L_n) to (W, L) (cf. Ethier and Kurtz, 1986, p. 153, Problem 22(b)), or we use the Cramér-Wold theorem (cf. Billingsley, 1968).

Proof of Theorem 2.4.1. By Lemma 2.4.1 and Le Cam's third lemma, all finite dimensional distributions of S_n converge to those of $W + d$, where $d(t) = \text{Cov}(W(t), L(1)) = C(t)$. Since contiguity preserves tightness, the result holds.

2.5. Weighted partial sums under contiguous measures

In Section 1.2 we have proved the weak convergence of weighted partial sum process of independent random variables with finite second moment under the null hypothesis H_0 of $X_1, X_2, \dots$ being identically distributed. Here we show that under the sequence of contiguous alternatives of H_1 as parametrized in (2.3.1)–(2.3.3), the convergence in distribution of weighted partial sum processes holds for the same optimal class of weight functions as in the case of H_0 .

As before, let $\mathcal{Q}$ be the class of functions q defined on $(0, 1]$ which are nondecreasing near zero and

$$(2.5.1) \quad \inf_{\delta \leq t \leq 1} q(t) > 0 \quad \text{for all } 0 < \delta < 1.$$

Recall also the integrals

$$E(q, c) = \int_0^1 t^{-3/2} q(t) \exp(-ct^{-1} q^2(t)) dt$$

and

$$I(q, c) = \int_0^1 t^{-1} \exp(-ct^{-1} q^2(t)) dt, \quad 0 < c < \infty.$$

Theorem 2.5.1. Let $X_1, X_2, \dots$ be independent random variables and $S(nt) = X_1 + \dots + X_{[nt]}$. We assume that under H_0 $\int x dF = 0$, $\int x^2 dF = 1$. Then, as $n \rightarrow \infty$, under H_1 the following hold.

(a) Let $q \in \mathcal{Q}$. If $I(q, c) < \infty$ for all $c > 0$, then

$$n^{-1/2} S(nt)/q(t) \xrightarrow{\mathcal{D}} (W(t) + C(t))/q(t) \quad \text{in } D[0, 1],$$

where

$$C(t) = \int_0^t \int_0^1 g(s, u) F^{-1}(u) \, du \, ds.$$

(b) Let $q : (0, 1) \rightarrow (0, \infty)$ be such that (2.5.1) holds. Then

$$n^{-1/2} S(nt)/q(t) \xrightarrow{\mathcal{D}} (W(t) + C(t))/q(t) \quad \text{in } D[\delta, 1]$$

for any $\delta \in (0, 1)$.

Remark 2.5.1. According to Theorem 1.1.C, in part (a) of Theorem 2.5.1 we can equivalently assume $E(q, c) < \infty$ for all $c > 0$ and $\lim_{t \downarrow 0} q(t)/t^{1/2} = \infty$ instead of $I(q, c) < \infty$ for all $c > 0$.

Proof. (a) Let $q \in \mathcal{Q}$ and $I(q, c) < \infty$ for all $c > 0$. Since $g \in L^2[0, 1]^2$ and $\int x^2 dF = 1$, $C(t)$ is finite. In fact we have

$$\begin{aligned} |C(t)| &= \left| \int_0^t \int_0^1 g(s, u) F^{-1}(u) \, du \, ds \right| \\ &\leq \int_0^t \left(\left(\int_0^1 g^2(s, u) \, du \right)^{1/2} \left(\int_0^1 (F^{-1}(u))^2 \, du \right)^{1/2} \right) ds \\ &= \int_0^t \left(\int_0^1 g^2(s, u) \, du \right)^{1/2} ds \\ &\leq \left(\int_0^t \int_0^1 g^2(s, u) \, du \, ds \right)^{1/2} \left(\int_0^t ds \right)^{1/2} \\ &= \left(\int_0^t \int_0^1 g^2(s, u) \, du \, ds \right)^{1/2} t^{1/2}. \end{aligned}$$

which implies

$$\sup_{0 < t \leq 1} |C(t)|/q(t) \leq \sup_{0 < t \leq 1} \left(\int_0^t \int_0^1 g^2(s, u) du ds \right)^{1/2} \sup_{0 < t \leq 1} t^{1/2}/q(t) < \infty,$$

due to the fact that $q(t)$ is positive, i.e. $\inf_{\delta \leq t \leq 1} q(t) > 0$ for all $\delta \in (0, 1)$ and because $I(q, c) < \infty$ for some $c > 0$ implies $\lim_{t \downarrow 0} q(t)/t^{1/2} = 0$ (cf. Theorem 1.1.B).

Put $S_n(t) = n^{-1/2} S(nt)$. By Theorem 1.2.1

$$S_n(t)/q(t) \xrightarrow{D} W(t)/q(t) \quad \text{in } D[0, 1].$$

Let $L_n(t)$ be defined as in (2.3.4).

Using Lemma 2.4.1 and the fact that

$$E \frac{S_n(t_1)}{q(t_1)} L_n(t_2) = \frac{1}{q(t_1)} E S_n(t_1) L_n(t_2) \rightarrow \frac{1}{q(t_1)} C(t_1 \wedge t_2),$$

we conclude that

$$(S_n(t)/q(t), L_n(t)) \xrightarrow{D} (W(t)/q(t), L(t)) \quad \text{in } D^2[0, 1].$$

Now Le Cam's third lemma implies the result (cf. Note following Lemma 2.1.F).

(b) By Theorem 2.1.1 we have

$$S_n(t) \xrightarrow{D} W(t) + C(t) \quad \text{in } D[0, 1].$$

By the Skorohod-Dudley-Wichura theorem (cf. Theorem A.2.1) there exists versions S_n^* , $n \geq 1$, and W^* such that

$$(2.5.2) \quad S_n^* \stackrel{D}{=} S_n, \quad n \geq 1, \quad W^* \stackrel{D}{=} W + C$$

and

$$\sup_{0 \leq t \leq 1} |S_n^*(t) - W^*(t)| = o_P(1).$$

Hence for any $q : (0, 1] \rightarrow (0, \infty)$ such that $\inf_{\delta \leq t \leq 1} q(t) > 0$ for all $\delta \in (0, 1)$ we have

$$\sup_{\delta \leq t \leq 1} |S_n^*(t) - W^*(t)|/q(t) = o_P(1),$$

which implies

$$S_n^*/q \xrightarrow{D} W^*/q \quad \text{in } D[\delta, 1].$$

Now, on account of (2.5.2), we conclude

$$S_n(t)/q(t) \xrightarrow{D} (W(t) + C(t))/q(t) \quad \text{in } D[\delta, 1]$$

for all $\delta \in (0, 1)$.

CHAPTER 3

Changepoint Problem

3.0. Introduction

The problem of abrupt parameter changes arises in a variety of experimental and mathematical sciences. For instance, in epidemiology one may be interested in testing whether the incidence of a disease has remained constant over time, and if not, in estimating the time(s) of change(s) in order to suggest possible causes. Another example is rhythm analysis in electrocardiograms, where the use of change detection methods constitutes a part of pattern recognition analysis. Detection of possible changepoints is also of interest in ethological time series, in quality control, in seismic signal processing and in the study of archaeological sites.

Such situations can usually be modeled by saying that we have a random process that generates independent observations indexed by time and we wish to detect whether a change could have occurred in the distribution that governs this random process as time goes by.

We study such phenomena in terms of various stochastic (probabilistic) processes. The construction of these processes is such that statistical tests can be based on them for detecting a possible change in their distribution, a deviation from whatever was being assumed to be a normal (acceptable) flow of events when studying a given random phenomenon.

When studying problems of detecting abrupt changes, in Probability and Statistics one usually describes them as changepoint problems. These have originally arisen in the context of quality control, where one typically observes the output of a production line and would wish to signal deviation from an acceptable

output level while observing the data. When one observes such a random process sequentially and stops observing at a random time of detecting change, then one speaks of a sequential procedure. Otherwise, one usually observes a large, chronologically ordered, finite sequence for the sake of determining possible change during the data collection. Thus, in this case, one studies a fixed, usually large sample, in various sequential ways, frequently keeping the chronological order of data as they came, or ordering them via some other schemes for further sequential studies. Most such "fixed sample size sequential procedures" are however described in terms of asymptotic results ("infinite" sample size, i.e., as $n \rightarrow \infty$). Hence we prefer to call these non-sequential procedures. Based on the chronologically ordered set of data, the main idea is to construct stochastic processes which should be sensitive to changes in their probabilistic behaviour.

The nature of our approach to studying these stochastic processes is such that we approximate them by their ideal asymptotic versions, usually Gaussian processes, and we wish to measure their distance from these under the most stringent possible conditions. Speaking in somewhat more mathematical terms, for measuring the nearness of stochastic processes we wish to use metrics which should be able to accommodate violent changes early enough, or not too late, in time, without ruining our ability to measure their distance (deviation) from each other. Thus we aim at studying the nearness of approximations of various empirical type processes by appropriate Gaussian processes in weighted supremum (uniform) and L_p (integral) norms (distances) under optimal weighing conditions (largest possible class of weight functions). Consequently, the construction of these processes will be such that statistical tests based on various weighted functionals (weighted statistics) of them should be sensitive to deviations from some postulated, acceptable stochastic (distributional) behaviour.

Changepoint problems have been extensively studied in the literature from the parametric as well as the nonparametric point of view. Parametric results are summarized in Siegmund (1985). Nonparametric results are reviewed in Wolfe and Schechtman (1984), and Csörgő and Horváth (1988c). See also the bibliographies of Shaban (1980) and Zacks (1983). Parzen (1992) proposes a general approach to change analysis based on a unified framework for statistical methods that has been evolving since Parzen (1979). A collection of articles edited by Basseville and Benveniste (1986) based on presentations in a conference on the subject of "Detection of Abrupt Changes in Signals and Dynamical Systems" held in Paris on March 21-22, 1984, provides an excellent survey of mathematical tools and applications.

Sequential as well as non-sequential procedures for changepoint problems are usually based on parametric or nonparametric models, allowing at most one change (AMOC) or, possibly, more than one change.

Here we are concerned with the nonparametric AMOC setting only. Being strongly influenced by the works of Csörgő and Horváth in the area of changepoint detection, we find it useful and convenient here to present this topic following the lines of their review paper Csörgő and Horváth (1988c), and to indicate our own contributions along these lines.

Csörgő and Horváth (1988c) used the following general parameterization of a changepoint problem.

Let $X_1, \dots, X_n$ be independent random variables. They consider the null hypothesis

$$H_0 : X_i, 1 \leq i \leq n, \text{ have the same distribution.}$$

2

PM-1 3 1/2" x 4" PHOTOGRAPHIC MICROCOPY TARGET
NBS 1010a ANSI/ISO #2 EQUIVALENT

The alternative hypothesis is that there is a changepoint in the sequence $X_1, \dots, X_n$, namely that we have

$$\begin{aligned} H_A : & \text{There is an } \eta \in (0, 1) \text{ such that } P\{X_1 \leq t\} = P\{X_2 \leq t\} \\ &= \dots = P\{X_{[n\eta]} \leq t\}, P\{X_{[n\eta]+1} \leq t\} = \dots = P\{X_n \leq t\}, \\ &-\infty < t < \infty, \text{ and } P\{X_{[n\eta]} \leq t_0\} \neq P\{X_{[n\eta]+1} \leq t_0\} \text{ for some } t_0. \end{aligned}$$

Let

$$H(t) = P\{X_{[n\eta]} \leq t\}, \quad G(t) = P\{X_{[n\eta]+1} \leq t\}$$

and

$$F(t) = \eta H(t) + (1 - \eta)G(t).$$

Thus we are testing for $\eta = 1$ (no changepoint) versus $\eta \in (0, 1)$, and under the null hypothesis we have $F = H$.

We wish to test the null hypothesis

$$H_0 : X_i, \ 1 \leq i \leq n, \text{ have the same distribution } F,$$

versus the alternative hypothesis

$$H_1 : X_i, \ 1 \leq i \leq n, \text{ have the respective distribution functions } F_{in},$$

where we assume that all F_{in} are absolutely continuous with respect to the distribution function F and

$$(3.0.1) \quad \left[\frac{dF_{in}}{dF} (F^{-1}(u)) \right]^{1/2} = 1 + \frac{1}{2n^{1/2}} g_n(t, u), \quad \frac{i-1}{n} < t \leq \frac{i}{n},$$

where $F^{-1}(u) = \inf\{x : F(x) \geq u\}$, $0 \leq u \leq 1$, $F^{-1}(0) = F^{-1}(0^+)$ (cf (2.3.1)).

We assume also that there exists a function $g \in L^2[0, 1]^2$ such that

$$(3.0.2) \quad \int_0^1 g(t, u) \, du = 0 \quad \text{for almost all } t \in [0, 1]$$

and

$$(3.0.3) \quad \int_0^1 \int_0^1 \left[g_n(t, u) - g\left(\frac{[nt]}{n}, u\right) \right]^2 du dt \rightarrow 0, \quad n \rightarrow \infty.$$

The sequence of direct products $F_{1n} \times \dots \times F_{nn}$, $n = 1, 2, \dots$ of such F_{in} is contiguous to the sequence $F \times \dots \times F$ (cf. Section 2.3). In particular, for the changepoint problem we assume that there exists $\eta \in (0, 1)$ such that

$$(3.0.4) \quad g(t, u) = 1\{t \geq \eta\} g(u)$$

for some square integrable function g .

In this formulation then, H_A changes into

H_1^η : There is an $\eta \in (0, 1)$ such that $P\{X_1 \leq t\} = P\{X_2 \leq t\} = \dots = P\{X_{[n\eta]} \leq t\}$ and $X_{[n\eta]+1}, \dots, X_n$ have distributions $F_{([n\eta]+1)n}, \dots, F_{nn}$ respectively, parameterized as in (3.0.1)–(3.0.3) with the function $g(\cdot, \cdot)$ as in (3.0.4).

This description of alternatives was used by Khmaladze and Parjanadze (1986) in the context of studying changepoint problems using linear statistics of sequential ranks.

3.1. Change in the mean

The basic AMOC problem (at most one change) of $X_1, \dots, X_n$ independent, continuous random variables such that X_i , $i = 1, \dots, k$ has distribution function $F(x)$, and X_i , $i = k + 1, \dots, n$ has distribution function $F(x - \Delta)$ with $-\infty < \Delta < \infty$ was first considered by Page (1954, 1955). He considered testing the null hypothesis of no change, that is, $H_0^{(1)} : \Delta = 0$, against either one- or two-sided alternatives, under the assumption that the initial mean, say θ_0 , of the process

(i.e. the mean of X_1) is known a priori. Letting $S_0^* = 0$, $S_k^* = \sum_{j=1}^k V_j$, $k = 1, \dots, n$, and

$$V_j = \begin{cases} a & \text{if } X_j > \theta_0, \\ -b & \text{if } X_j \leq \theta_0. \end{cases}$$

where $a > 0$, $b > 0$ are constants, possibly dependent on F , chosen so that $EV_j = 0$, $j = 1, \dots, n$, Page's decision rule rejects $H_0^{(1)} : \Delta = 0$ in favour of the alternative of one change and $\Delta > 0$ if

$$T_n^* = \max_{0 \leq k \leq n} \left\{ S_k^* - \min_{0 \leq j \leq k} S_j^* \right\}$$

is too large. Special emphasis was given to the choice of $a = b = 1$, in which case the V_j are simply the signs associated with the $(X_j - \theta_0)$. The resulting changepoint test is nonparametric, distribution free over the class of continuous random variables. Indeed under $H_0^{(1)}$, on choosing a and b so that $EV_j = 0$, $j = 1, \dots, n$, where it is not necessary any more that θ_0 should be EX_1 , we have $E(S_n^*)^2 = nab$, and Donsker's invariance principle (cf. for example Billingsley, 1968) yields

$$\left(S_{[nt]}^* - \inf_{0 \leq j \leq [nt]} S_j^* \right) / (nab)^{1/2} \xrightarrow{\mathcal{D}} W(t) - \inf_{0 \leq s \leq t} W(s) \quad \text{in } D[0, 1],$$

where W is a standard Wiener process. It is well known (cf. Chung and Williams, 1983) that

$$\left\{ W(t) - \inf_{0 \leq s \leq t} W(s), 0 \leq t \leq 1 \right\} \stackrel{\mathcal{D}}{=} \{|W(t)|, 0 \leq t \leq 1\}.$$

Hence, (cf. Feller, 1966)

$$\begin{aligned} \lim_{n \rightarrow \infty} P \left\{ T_n^* / (nab)^{1/2} \leq x \right\} &= P \left\{ \sup_{0 \leq t \leq 1} |W(t)| \leq x \right\} \\ (3.1.1) \quad &= 1 - 4 \sum_{k=1}^{\infty} (-1)^{k+1} \Phi(-(2k-1)x), \quad x > 0, \end{aligned}$$

where Φ is the standard normal distribution function.

To make the above changepoint test more powerful for detecting change that may have occurred at the beginning, it may be desirable to introduce weight functions. As an immediate application of Theorem 1.2.1 to changepoint problems we have, without additional assumptions on X_i , the following result

$$\left(S_{[nt]}^* - \inf_{0 \leq j \leq [nt]} S_j^* \right) / (nab)^{1/2} q(t) \xrightarrow{D} W(t)/q(t) - \inf_{0 \leq s \leq t} W(s)/q(s) \quad \text{in } D[0, 1],$$

where $q \in \mathcal{Q}$ and $I(q, c) < \infty$ for all $c > 0$.

Again, for the case where the initial mean $EX_1 = \theta_0$ is known, it appears to be natural to reject $H_0^{(1)} : \Delta = 0$ in favour of $H_1^{(1)} : \Delta \neq 0$ for large values of

$$(3.1.2) \quad T_n = \max_{0 \leq k \leq n} |S_k - k\theta_0|,$$

where $S_0 = 0$ and $S_k = \sum_{i=1}^k X_i$, $k = 1, \dots, n$. Using again Donsker's invariance principle we have

$$(S_{[nt]} - [nt]\theta_0) / (n^{1/2}\sigma) \xrightarrow{D} W(t) \quad \text{in } D[0, 1],$$

where W is a standard Wiener process, and $\sigma^2 = E(X_1 - \theta_0)^2$. Hence with any consistent sequence of estimators $\{\sigma_n\}$ for σ we get

$$\lim_{n \rightarrow \infty} P\{T_n / (n^{1/2}\sigma_n) \leq x\} = P\left\{ \sup_{0 \leq t \leq 1} |W(t)| \leq x \right\},$$

just like in (3.1.1). Thus Page's T_n^* and T_n of (3.1.2) are asymptotically equivalent in distribution. Also, if in the case of T_n we wished to introduce weight functions to make the statistical test of (3.1.2) more sensitive near 0, then using again Theorem 1.2.1, we get

$$(S_{[nt]} - [nt]\theta_0) / (n^{1/2}\sigma q(t)) \xrightarrow{D} W(t)/q(t) \quad \text{in } D[0, 1],$$

where $q \in \mathcal{Q}$ and $I(q, c) < \infty$ for all $c > 0$.

In their review Csörgö and Horváth (1988c) illustrate that a large number of nonparametric as well as parametric modelling of AMOC problems result in the same test statistic, namely in

$$(3.1.3) \quad \max_{1 \leq k \leq n} \left\{ |S_k - kS_n/n| / (k(1 - k/n))^{1/2} \right\}.$$

For example, it is used when testing for shift in distribution like $H_0^{(1)}$ versus $H_1^{(1)}$, as well as when testing for $H_0^{(2)} : EX_1 = \dots = EX_n = \theta_0$ versus $H_1^{(2)}$: there is a $\eta \in (0, 1)$ such that $EX_1 = \dots = EX_{[n\eta]} = \theta_0 \neq \theta_1 = EX_{[n\eta]+1} = \dots = EX_n$. Assuming for a moment that $\sigma^2 = 1$ and that in $H_1^{(2)}$ $[n\eta] = k$ is known, our changepoint problem in this special case becomes a two population test to decide whether the mean θ_0 of the first k observations equals to the mean θ_1 of the last $n - k$ observations. The standard test statistic is the normalized difference between the mean of the first k observations, S_k/k and the overall mean, S_n/n . This is just (3.1.3) with the max, which accounts for the fact that λ is actually unknown. The test statistic (3.1.3) is frequently concluded also in the case of parametric modelling of changepoint problems. Assuming normality, for example, in case of $H_0^{(i)}$ versus $H_1^{(i)}$ ($i = 1, 2$) it is easily shown that the square root of the log-likelihood ratio statistic is proportional to the test statistic of (3.1.3) (cf., for example, Sen and Srivastava, 1975; Yao and David, 1984; Siegmund, 1986). Assuming exponentiality in $H_0^{(i)}$ versus $H_1^{(i)}$ ($i = 1, 2$), Haccou, Meelis and van de Geer (1985) conclude that a linear transformation of the square root of the log likelihood ratio statistic is asymptotically proportional to the test statistic of (3.1.3). Matthews, Farewell and Pyke (1985) arrived at a similar conclusion when studying a changepoint parameter τ for the hazard rate at which it changes from the constant hazard $\lambda(t) = \lambda$, ($0 \leq t < \tau$) to $\lambda(t) = (1 - \xi)\lambda$, ($t \geq \tau$), where

$0 \leq \xi < 1$ and $\tau \geq 0$. For an illuminating presentation and extension of this so-called sample Brownian bridge problem we refer to Parzen (1992).

Clearly then, when considering changepoint problems, the study of the limiting behaviour of the test statistic (3.1.3) is of interest. Unfortunately, given H_0 , as $n \rightarrow \infty$ we have

$$(3.1.4) \quad \max_{1 \leq k < n} \left(S_k - \frac{k}{n} S_n \right) / n^{1/2} \left(\frac{k}{n} \left(1 - \frac{k}{n} \right) \right)^{1/2} \xrightarrow{P} \infty.$$

Hence, in order to have the above test statistics converge in distribution to a nondegenerate random variable, we have to renormalize it, or introduce weight functions which are less severe on the tails than $((k/n)(1 - k/n))^{1/2}$ is.

In the case of θ_0 unknown, and then estimated by the sample mean, T_n has the form

$$(3.1.5) \quad \max_{1 \leq k < n} |S_k - k S_n / n|.$$

Pettit (1980) also considered the statistic of (3.1.5) for the changepoint problem with zero-one observations. The asymptotic behaviour of the statistic of (3.1.5) is well known, namely under the null hypothesis Donsker's invariance principle again yields

$$(S_{[nt]} - ([nt]/n) S_n) / (n^{1/2} \sigma) \xrightarrow{D} B(t) \quad \text{in } D[0, 1],$$

where $\{B(t), 0 \leq t \leq 1\}$ is a Brownian bridge. Hence with any sequence of consistent estimators $\{\sigma_n\}$ for σ we have

$$\begin{aligned} \lim_{n \rightarrow \infty} P \left\{ \max_{0 \leq k \leq n} |S_k - k S_n / n| / (n^{1/2} \sigma_n) \leq x \right\} \\ = P \left\{ \sup_{0 \leq t \leq 1} |B(t)| \leq x \right\} = 1 - \sum_{k \neq 0} (-1)^{k+1} \exp(-2k^2 x^2), \quad x > 0. \end{aligned}$$

The latter distribution function is well-known as the limiting distribution of the two-sided Kolmogorov-Smirnov statistic. Consequently, just like the classical Kolmogorov-Smirnov statistics, (3.1.5) should be more powerful for detecting changes that occur near $n/2$ than noticing the ones occurring near the endpoints 0 and n . Thus weighted versions of (3.1.5) should be modifications of (3.1.3), which would emphasize changes which may have occurred near the endpoints, while retaining sensitivity to possible changes in the middle as well.

Assuming that $E|X_1|^\nu < \infty$ for some $\nu > 2$, the asymptotic theory of (3.1.3) under the null hypothesis of i.i.d. observations was described in $\|\cdot/q\|$ -metrics by Csörgő and Horváth (1988c) (cf. their Theorem 2.1 and Corollary 2.1). In Chapter 5 we improve on their result under H_0 , assuming only two moments for X_1 , and extend it as well under the contiguous alternatives H_1 (cf. also Szyszkowicz 1992b,c) for the optimal class of weight functions. We establish also optimal L_p -approximations for tied-down partial sum processes.

3.2. U -statistic type processes

In the context of the changepoint problem, Csörgő and Horváth (1988a) study tests based on processes of U -statistics which are generalizations of Wilcoxon-Mann-Whitney type statistics. They consider

$$(3.2.1) \quad Z_k = \sum_{1 \leq i \leq k} \sum_{k+1 \leq j \leq n} h(X_i, X_j), \quad 1 \leq k < n,$$

where $h(x, y)$ is a symmetric or an antisymmetric function and study its limiting behaviour in unweighted and weighted sup-norm metrics. Using $\|\cdot/q\|$ -metrics, Csörgő and Horváth (1988a) obtain their results under the null hypothesis H_0 , assuming $E|h(X_1, X_2)|^\nu < \infty$ for some $\nu > 2$. In Chapter 6 we improve on these results assuming two moments only. We also study L_p -functionals of weighted Z_k

processes. Finally we consider the behaviour of such process under the class of contiguous and changepoint alternatives. Taking $h(x, y) = x - y$ gives $\mathcal{Z}_k$ equal to

$$n \left(S_k - \frac{k}{n} S_n \right), \quad 1 \leq k < n,$$

where $S_k = \sum_{i=1}^k X_i$, or, equivalently, to

$$(3.2.2) \quad k(n-k) \left(\frac{1}{k} \sum_{i=1}^k X_i - \frac{1}{n-k} \sum_{i=k+1}^n X_i \right).$$

A statistic based on (3.2.2) was used by Hinkley (1970) to estimate a changepoint (see also Pettitt, 1980a). Another example of antisymmetric kernel which has gained special attention in the literature is $h(x, y) = \text{sign}(x - y)$. Pettitt (1979) studied the process

$$(3.2.3) \quad U_k = \sum_{i=1}^k \sum_{j=k+1}^n \text{sign}(X_i - X_j)$$

for the sake of testing $H_0^{(1)}$ versus $H_1^{(1)}$ of Section 3.1. In particular, under H_0 he concluded (cf. also Csörgő and Horváth, 1988a) that $(12)^{1/2} n^{-3/2} U_{[(n+1)t]}$ converges weakly to a Brownian bridge in $D[0, 1]$. Sen and Srivastava (1975), as well as Schechtman and Wolfe (1981), suggest, as an analog to some parametric likelihood ratio procedure to reject $H_0^{(1)}$ in favour of $H_1^{(1)}$ for large values of

$$D_n = (12)^{1/2} \max_{1 \leq k \leq n} |U_k| / (k(n-k+1)n)^{1/2}.$$

However, $D_n \xrightarrow{P} \infty$ even under H_0 . Results of Csörgő and Horváth (1988a) indicate what kind of weight functions are possible for the process U_k under H_0 . From the results of Section 6.3, where we study behaviour of $\mathcal{Z}_{[nt]}$ processes under

contiguous alternatives, we conclude in particular that under the changepoint alternatives H_1^η

$$\begin{aligned} & \sup_{0 < t < 1} \left(12^{1/2} n^{-3/2} \sum_{i=1}^{[(n+1)t]} \sum_{j=[(n+1)t]+1}^n \text{sign}(X_i - X_j) \right) / \left(t(1-t) \right)^{1/2} \log \log 1 / (t(1-t)) \\ & \xrightarrow{D} \sup_{0 < t < 1} \left(B(t) + 12^{1/2} (-t(1-\lambda) \mathbf{1}\{t < \lambda\} \right. \\ & \quad \left. - \lambda(1-t) \mathbf{1}\{t \geq \lambda\}) \int_0^1 u g(u) du \right) / \left(t(1-t) \right)^{1/2} \log \log 1 / (t(1-t)). \end{aligned}$$

We note also that (3.2.3) is equivalent to

$$2 \sum_{i=1}^k R_i^n - k(n+1),$$

where R_i^n are the ranks of $X_1, \dots, X_n$. Hence we can conclude similarly weighted results under H_0 and also under the sequence of contiguous alternatives H_1 for cumulatives of ranks as well. For further results on the applications of rank statistics to the changepoint problem we refer to Lombard (1981, 1983, 1986, 1987).

3.3. Sequential ranks

Statistical tests based on sequential ranks are frequently used because of their convenient properties. Bhattacharya and Frierson (1981) propose a control chart based on partial sums of sequential ranks for detecting small changes in the distribution of a random sample. Khmaladze and Parjanadze (1986) study changepoint problems under contiguous alternatives using linear statistics of sequential ranks. They consider the process

$$V_n(t) = n^{-1/2} \sum_{i=1}^{[nt]} a(\xi_k), \quad 0 \leq t \leq 1,$$

for $a \in L^2(0,1)$ and such that $\int_0^1 a(u) du = 0$, where $\xi_1, \xi_2, \dots$ are normalized sequential ranks. For further results along these lines we refer to Lombard (1983).

In Chapter 4 we study the asymptotic behaviour of weighted V_n process for the optimal class of weight functions.

Here we wish to point out the generality of the Khmaladze and Parjanadze (1986) description of changepoint alternatives as given in (3.0.1)–(3.0.4). The limiting distribution of different processes under various “close” alternatives has been derived by many authors. For example, Bhattacharya and Frierson (1981) derive the limiting distribution of the partial sum process of sequential ranks under the assumption that the distribution changes from F to G_N after the first $[N\theta]$ of $X_1, \dots, X_N$ variables, where $\theta \in (0, 1]$, F and G_N are continuous distribution functions and the change from F to G_N is small in the sense that

$$\begin{aligned} \xi_N &= \int F(x) dG_N(x) = 1/2 + \delta N^{-1/2} + o(N^{-1/2}), \quad \delta > 0 \\ (3.3.1) \quad \mathcal{L}_N &= 2 \int F(x) G_N(x) dG_N(x) = 2/3 + O(N^{-1/2}), \\ \beta_N &= \int F^2(x) dG_N(x) = 1/3 + O(N^{-1/2}). \end{aligned}$$

Concerning the process

$$S_N(t) = (12/N)^{1/2} \{S_{[Nt]} + (Nt - [Nt])Z_{[Nt]+1}\}, \quad 0 \leq t \leq 1,$$

where $S_k = \sum_{i=1}^k Z_i$, $k = 1, \dots, N$ and Z_i 's are the normalized sequential ranks centralized at the mean. Bhattacharya and Frierson (1981) derive its weak limit, which is necessary to describe the control chart proposed by them, and study its properties when a change in distribution occurs. The main result of their paper states that, with respect to uniform convergence in $C[0, 1]$

$$(3.3.2) \quad S_N(t) \xrightarrow{D} W(t) + (12)^{1/2} \delta \theta \log(t/\theta) 1\{t \geq \theta\}, \quad 0 \leq t \leq 1,$$

where $W(t)$ is a standard Brownian motion and $\delta > 0$ is determined from (3.3.1).

Considering the original observations, let M_F and $\sigma_F^2 < \infty$ denote the mean and the variance of F and define M_G and $\sigma_G^2 < \infty$ likewise. Suppose the change from F to G_N is such that (3.3.1) holds and $\sqrt{N}(M_{G_N} - M_F)/\sigma_F \rightarrow \Delta > 0$ and $\sigma_{G_N} \rightarrow \sigma_F$ as $N \rightarrow \infty$. If M_F and σ_F^2 of the initial distribution are known, the control chart based on the cumulative sums of the original observations is described in terms of the limiting distribution of the process

$$S_N^*(t) = N^{-1/2} \left\{ \sum_{i=1}^{[Nt]} Z_i^* + (Nt - [Nt])Z_{[Nt]+1}^* \right\}, \quad 0 \leq t \leq 1,$$

where $Z_i^* = (X_i - M_F)/\sigma_F$. We have

$$(3.3.3) \quad S_N^*(t) \xrightarrow{D} W(t) + \Delta(t - \theta)1\{t \geq \theta\}, \quad 0 \leq t \leq 1,$$

where $W(t)$ is a standard Brownian motion and Δ is as given above.

We note that $S_N^*(t)$ is a continuous version of $S_n(t)$ process considered in Chapter 1 and $S_N(t)$ corresponds to the $V_n(t)$ process considered by Khmaladze and Parjanadze (1986), as well as by us in Chapter 4. A look at Corollary 4.2.1 with $q \equiv 1$, and Theorem 2.4.1 with g as in (3.0.4), shows that, using the parametrization of the changepoint problem via contiguous alternatives as given in (3.0.1)–(3.0.4) together with Le Cam's theory of contiguity, we arrive at the same kind of shifted distributions of Brownian motion like those of (3.3.2) and (3.3.3), respectively. For further results along these lines we refer to Lombard (1983).

3.4. Kendall-Kendall pontograms

This section is devoted to a somewhat different problem than that of testing H_0 versus H_A (or H_1) discussed so far. However, the nature of the resulting tests will be familiar and related to the ones mentioned in Section 3.1.

Let $\{N(x), x \geq 0\}$ be a nonhomogeneous, also called nonstationary, Poisson process with intensity function $\lambda(x), x \geq 0$. We consider the problem of testing the null hypothesis of $\lambda(x)$ having a constant value against the alternative that $\lambda(x)$ is a function of x .

When looking for a possible changepoint in the Land's End data, Kendall and Kendall (1980) considered test-statistics based on the stochastic process

$$\left(\frac{n}{N(n)}\right)^{1/2} \frac{K_n(t)}{(t(1-t))^{1/2}}, \quad 0 \leq t \leq 1$$

called by them a pontogram, where

$$K_n(t) = n^{-1/2}(N(nt) - tN(n)), \quad 0 \leq t \leq 1.$$

Assuming that $\lambda(x)$ has a constant, but unknown value $\lambda > 0$, Kendall and Kendall (1980) proved that, for all $0 < \epsilon_1 < 1 - \epsilon_2 < 1$, we have

$$\sup_{\epsilon_1 \leq t \leq 1 - \epsilon_2} \left(\frac{n}{N(n)}\right)^{\frac{1}{2}} \frac{|K_n(t)|}{(t(1-t))^{\frac{1}{2}}} \xrightarrow{D} \sup_{0 \leq s \leq 1/2 \log[(1-\epsilon_2)/\epsilon_1][(1-\epsilon_1)/\epsilon_2]} |V(s)|,$$

where $\{V(s), -\infty < s < \infty\}$ is an Ornstein-Uhlenbeck process with covariance function $\exp(-|t-s|)$. The problem of replacing ϵ_1 and ϵ_2 by zero, or by sequences $\epsilon_i(n) \rightarrow 0$ as $n \rightarrow \infty$, $i = 1, 2$, was solved by Csörgő and Horváth (1987c) via the following result (cf. Theorem 1.1. of Csörgő and Horváth, 1987c).

Theorem 3.4.A. *Let $\{N(x), x \geq 0\}$ be a Poisson process with intensity parameter $\lambda > 0$. Then there exists a sequence of Brownian bridges $\{B_n(t); 0 \leq t \leq 1\}$ such that*

$$n^\tau \sup_{1/(n+1) \leq t \leq n/(n+1)} \left| n^{-1/2}(N(nt) - tN(n)) - \lambda^{1/2} B_n(t) \right| / (t(1-t))^{1/2-\tau} = O_P(1)$$

where $0 \leq \tau < 1/2$.

In particular they obtain (cf. Theorem 7.0.A) that, as $n \rightarrow \infty$

$$(3.4.1) \quad (2 \log \log n)^{1/2} \left(\frac{n}{N(n)} \right)^{\frac{1}{2}} \sup_{0 < t < 1} |K_n(t)| / (t(1-t))^{1/2} \\ - 2 \log \log n - \frac{1}{2} \log \log \log n + \frac{1}{2} \log \pi \xrightarrow{D} Y,$$

where Y is a random variable with distribution function $\exp(-2\exp(-x))$, $-\infty < x < \infty$.

Here we consider the asymptotic behaviour of the process

$$\left\{ \left(\frac{n}{N(n)} \right)^{1/2} K_n(t)/q(t), \quad 0 < t < 1 \right\}$$

for the largest possible class of weight functions q (cf. also Szyszkowicz 1992d,e).

Studying the asymptotic behaviour of the stochastic process $K_n(t)$, in Section 7.1 we assume that $\lambda(x) \equiv \lambda > 0$, and describe necessary and sufficient conditions for weight functions q to have the weak convergence of $K_n(t)$ process, and then that of its sup-functionals for a larger class of functions q . For example, we have, as $n \rightarrow \infty$

$$\sup_{0 < t < 1} \left(\frac{n}{N(n)} \right)^{\frac{1}{2}} \frac{|K_n(t)|}{(t(1-t))^{1/2}} / (\log \log 1/(t(1-t)))^{1/2} \\ \rightarrow \sup_{0 < t < 1} |B(t)| / (t(1-t) \log \log 1/(t(1-t)))^{1/2},$$

where $B(t)$ is a Brownian bridge (cf. Proposition 7.1.2(b)).

The asymptotics of weighted L_p norms, $0 < p < \infty$, of $K_n(t)$ are studied in Section 7.2. The following result is an L_p -analog of (3.4.1) (cf. Proposition 7.2.2 and for $p = 2$, Csörgő and Horváth 1987c). Assuming that $\lambda(x) \equiv \lambda > 0$ and $p \geq 1$, we have as $n \rightarrow \infty$

$$(4D \log n)^{-1/2} \left(\int_0^1 \left| \left(\frac{n}{N(n)} \right)^{\frac{1}{2}} K_n(t) \right|^p / (t(1-t))^{1+\frac{p}{2}} dt - 2M \log n \right) \\ \xrightarrow{D} N(0, 1),$$

where

$$M = M(p) = (2\pi)^{-1/2} \int_{-\infty}^{\infty} |x|^p \exp(-x^2/2) dx$$

and

$$D = D(p) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |xy|^p \times \left\{ \frac{1}{2\pi(1 - e^{-2|u|})^{1/2}} \exp\left(-\frac{x^2 + y^2 - 2e^{-|u|}xy}{2(1 - e^{-2|u|})^{1/2}}\right) - \phi(x)\phi(y) \right\} dx dy du.$$

Having the description of the behaviour of $K_n(t)/q(t)$ under the null hypothesis

$$H: \lambda(x) = \lambda > 0, \quad x \geq 0,$$

in Section 7.3 we study its asymptotic behaviour under the sequence of contiguous alternatives

$$A: \lambda(x) = \lambda + \xi\left(\frac{[x]}{n}\right)/n^{1/2}, \quad 0 \leq x \leq n, \quad n \geq 1,$$

where $\xi^2(t)$ is a Riemann integrable function.

In particular, A accommodates the so-called changepoint alternative

A^η : there exists $\eta \in (0, 1)$ such that

$$\lambda(x) = \lambda \quad \text{for } 0 \leq x < n\eta,$$

$$\lambda(x) = \lambda + \rho/n^{1/2} \quad \text{for } n\eta \leq x \leq n,$$

by taking $\xi(t) = \rho \mathbf{1}\{t \geq \eta\}$, $\rho \neq 0$. Under such assumptions we obtain, for example,

$$\begin{aligned} & \sup_{0 < t < 1} \left(\frac{n}{N(n)} \right)^{\frac{1}{2}} \frac{|K_n(t)|}{(t(1-t))^{1/2}} / \log \log 1/(t(1-t)) \\ & \xrightarrow{D} \sup_{0 < t < 1} \left(B(t) + \lambda^{-1/2} [-t\rho(1-\eta)\mathbf{1}\{t < \eta\} \right. \\ & \quad \left. - (1-t)\rho\eta\mathbf{1}\{t \geq \eta\}] \right) / (t(1-t))^{1/2} \log \log 1/(t(1-t)), \end{aligned}$$

where $B(t)$ is a Brownian bridge. The value of ρ is not known a priori. One may want to pick an appropriate value for it, for a particular problem at hand. We note only that $\rho > 0$ ($\rho < 0$) corresponds to a higher (lower) value of the intensity function $\lambda(x)$ of the Poisson process after the changepoint η of this parameterization.

3.5. Empirical processes

Many statistical procedures used for detecting a possible change in distribution of a random process are based on empirical processes. In Chapter 8 we describe the weighted asymptotic behaviour of three basic empirical processes in a unified way, under the null assumption of no change and also under a sequence of contiguous alternatives which accommodate the possible occurrence of a changepoint in the series of measurements.

Let $R_{1n}, \dots, R_{nn}$ denote the normalized ranks

$$R_{kn} = n^{-1} \sum_{i=1}^n 1\{X_i \leq X_k\}, \quad k = 1, \dots, n,$$

and let $\xi_1, \dots, \xi_n$ be the normalized sequential ranks

$$\xi_k = k^{-1} \sum_{i=1}^k 1\{X_i < X_k\}, \quad k = 1, \dots, n,$$

of the first n of the random variables $X_1, X_2, \dots$.

The three processes we consider in Chapter 8 are the two-time parameter empirical processes

$$\begin{aligned} \alpha_n(s, t) &= n^{-1/2} \sum_{i=1}^{[nt]} (1\{F(X_i) \leq s\} - s), \\ \beta_n(s, t) &= n^{-1/2} \sum_{i=1}^{[nt]} (1\{R_{in} \leq s\} - \lambda_n(s)), \\ \gamma_n(s, t) &= n^{-1/2} \sum_{i=1}^{[nt]} (1\{\xi_i \leq s\} - \lambda_i(s)), \end{aligned}$$

where $0 \leq s, t \leq 1$, and F is assumed to be continuous. Here $\lambda_n(s) = [ns]/n$ and $\lambda_t(x) = [is]/i$ denote the distribution functions of the normalized rank R_{in} and the normalized sequential rank ξ_i , respectively, under the null assumption H_0 . The asymptotic behaviour of the empirical process

$$\xi_n(x, t) = n^{-1/2} \sum_{i=1}^{[nt]} \left(\mathbf{1}\{X_i \leq x\} - F(x) \right), \quad -\infty < x < \infty, \quad 0 \leq t \leq 1.$$

which is of practical interest, may be concluded from that of $\alpha_n(s, t)$ due to the fact that for any random variable X having a continuous distribution function F , the random variable $F(X)$ is uniformly distributed on $[0, 1]$, and hence

$$\{\xi_n(F^{-1}(s), t), \quad 0 \leq s, t \leq 1, n \geq 1\} \stackrel{D}{=} \{\alpha_n(s, t), \quad 0 \leq s, t \leq 1, n \geq 1\}$$

Such empirical processes are studied by Pardzhanadze and Khmaladze (1986) under the class of contiguous alternatives H_1 . The process $\alpha_n(s, t)$ is a two-time parameter version of the usual empirical process. Csörgő and Horváth (1987a) propose nonparametric procedures for the changepoint problem based on the $\beta_n(s, t)$ process. Using a different parameterization of contiguity from that of (3.0.1)–(3.0.3) and a method of proof that differs from ours, Lejpus (1988, 1989) studies the processes $\alpha_n(s, t)$ and $\beta_n(s, t)$ as well as some functionals of $\beta_n(s, t)$ in connection with changepoint problems. The process $\gamma_n(s, t)$ is used by Csörgő and Horváth (1987b) to construct a sequential procedure for detecting a possible changepoint in a random sequence. Here we introduce some weight functions $q(t)$ and prove weak convergence of the weighted empirical processes as above under the null hypothesis H_0 , as well as under the class of contiguous alternatives H_1 (cf. also Szyszkowicz 1991b,c). These results with $q \equiv 1$ constitute an alternative proof of Theorem 1 of Pardzhanadze and Khmaladze (1986).

When testing for the existence of a changepoint in a random sequence, it may be of interest to make use of test statistics based on comparing the empirical distribution function of the first k observations to that of the last $n-k$ observations. For example, Picard (1985), and Deshayes and Picard (1986) consider

$$(3.5.1) \quad \sup_{1 \leq k < n} \sup_{x \in \mathbb{R}} n^{1/2} \left| \frac{1}{k} \sum_{i=1}^k \mathbf{1}\{X_i \leq x\} - \frac{1}{n-k} \sum_{i=k+1}^n \mathbf{1}\{X_i \leq x\} \right|$$

$$= \sup_{1 \leq k < n} \sup_{x \in \mathbb{R}} n^{1/2} \left| \sum_{i=1}^k \mathbf{1}\{X_i \leq x\} - \frac{k}{n} \sum_{i=1}^n \mathbf{1}\{X_i \leq x\} \right| / (k(n-k)/n).$$

Since, given H_0 , as $n \rightarrow \infty$ we have

$$\sup_{1 \leq k < n} \sup_{x \in \mathbb{R}} \left| \sum_{i=1}^k \mathbf{1}\{X_i \leq x\} - \frac{k}{n} \sum_{i=1}^n \mathbf{1}\{X_i \leq x\} \right| / n^{1/2} \frac{k}{n} \left(1 - \frac{k}{n}\right) \xrightarrow{P} \infty$$

as well as

$$\sup_{1 \leq k < n} \sup_{x \in \mathbb{R}} \left| \sum_{i=1}^k \mathbf{1}\{X_i \leq x\} - \frac{k}{n} \sum_{i=1}^n \mathbf{1}\{X_i \leq x\} \right| / n^{1/2} \left(\frac{k}{n} \left(1 - \frac{k}{n}\right) \right)^{1/2} \xrightarrow{P} \infty,$$

in order to have the above test statistics converge in distribution to a nondegenerate random variable, we have to renormalize them somehow. We may, for example, wish to introduce weight functions which are less severe on the tails than $((k/n)(1 - k/n))^{1/2}$ is.

Just like tests based on the classical Kolmogorov-Smirnov statistic, the ones based on

$$(3.5.2) \quad \sup_{1 \leq k < n} \sup_{x \in \mathbb{R}} \frac{k}{n} \left(1 - \frac{k}{n}\right) n^{1/2} \left| \frac{1}{k} \sum_{i=1}^k \mathbf{1}\{X_i \leq x\} - \frac{1}{n-k} \sum_{i=k+1}^n \mathbf{1}\{X_i \leq x\} \right|$$

$$= \sup_{1 \leq k < n} \sup_{x \in \mathbb{R}} n^{-1/2} \left| \sum_{i=1}^k \mathbf{1}\{X_i \leq x\} - \frac{k}{n} \sum_{i=1}^n \mathbf{1}\{X_i \leq x\} \right|$$

should be more powerful for detecting changes that occur in the middle, namely near $n/2$, where $\frac{k}{n} (1 - \frac{k}{n})$ has its maximum, than for noticing the ones occurring

near the endpoints 0 and n . Thus, a weighted version of (3.5.2) should emphasize changes which may have occurred near the endpoints, while retaining sensitivity to possible changes in the middle as well. To see what kind of functions are possible for the statistic in (3.5.1) or (3.5.2), we will study the asymptotic behavior of the process

$$(\alpha_n(s, t) - t\alpha_n(s, 1))/q(t)$$

for a wide class of functions q . In particular, we obtain that

$$(3.5.3) \quad \sup_{0 < t < 1} \sup_{0 \leq s \leq 1} |\hat{\alpha}_n(s, t)| / (t(1-t) \log \log 1/(t(1-t)))^{1/2} \\ \xrightarrow{\mathcal{D}} \sup_{0 < t < 1} \sup_{0 \leq s \leq 1} |\Gamma(s, t)| / (t(1-t) \log \log 1/(t(1-t)))^{1/2},$$

where

$$\hat{\alpha}_n(s, t) = \begin{cases} n^{-1/2} \left(\sum_{i=1}^{[(n+1)t]} 1\{F(X_i) \leq s\} - \frac{[(n+1)t]}{n} \sum_{i=1}^n 1\{F(X_i) \leq s\} \right), & 0 \leq t < 1, \\ & 0 \leq s \leq 1, \\ 0 & t = 1, \end{cases}$$

and $\{\Gamma(s, t); 0 \leq s \leq 1, 0 \leq t \leq 1\}$ denotes a Gaussian process with mean zero and covariance function

$$E\Gamma(s_1, t_1)\Gamma(s_2, t_2) = (s_1 \wedge s_2 - s_1 s_2)(t_1 \wedge t_2 - t_1 t_2).$$

The following test statistics for the changepoint problem were proposed by Csörgö and Horváth (1987a), where they derive their distributions under the null hypothesis H_0 . Let $\Gamma(s, t)$ be a Gaussian process as above and $B(s)$ be a Brownian bridge.

Then, assuming H_0

$$\begin{aligned}
& \sup_{0 \leq t \leq 1} \sup_{0 \leq s \leq 1} |\beta_n(s, t)| \xrightarrow{\mathcal{D}} \sup_{0 \leq t \leq 1} \sup_{0 \leq s \leq 1} |\Gamma(s, t)| \\
& \int_0^1 \int_0^1 \beta_n^2(s, t) ds dt \xrightarrow{\mathcal{D}} \int_0^1 \int_0^1 \Gamma^2(s, t) ds dt \\
& (s_0(1-s_0))^{-1/2} \sup_{0 \leq t \leq 1} |\beta_n(s_0, t)| \xrightarrow{\mathcal{D}} \sup_{0 \leq t \leq 1} |B(t)| \\
& (t_0(1-t_0))^{-1/2} \sup_{0 \leq s \leq 1} |\beta_n(s, t_0)| \xrightarrow{\mathcal{D}} \sup_{0 \leq s \leq 1} |B(s)| \\
& 12 \int_0^1 \int_0^1 \beta_n(s, t) ds dt \xrightarrow{\mathcal{D}} N(0, 1) \\
& \left(\frac{12}{s_0(1-s_0)} \right)^{1/2} \int_0^1 \beta_n(s_0, t) dt \xrightarrow{\mathcal{D}} N(0, 1) \\
& (48)^{1/2} \int_0^1 \left(-\beta_n\left(\frac{1}{2}, t\right) \right) dt \xrightarrow{\mathcal{D}} N(0, 1) \\
& \left(\frac{12}{t_0(1-t_0)} \right)^{1/2} \int_0^1 \beta_n(s, t_0) ds \xrightarrow{\mathcal{D}} N(0, 1),
\end{aligned}$$

where $0 < t_0, s_0 < 1$, and $N(0, 1)$ stands for the standard normal r.v.

Theorem 8.4.1 give the asymptotics of weighted versions of such functionals for $q(t)$ which is positive on $(0, 1)$ and such that

$$(3.5.4) \quad \begin{cases} \lim_{t \downarrow 0} (t(1-t) \log \log 1/(t(1-t)))^{1/2} / q(t) = 0 \\ \text{and} \\ \lim_{t \uparrow 1} (t(1-t) \log \log 1/(t(1-t)))^{1/2} / q(t) = 0. \end{cases}$$

For the sup-functional itself (3.5.3) holds with $\hat{\alpha}_n(s, t)$ replaced by $\hat{\beta}_n(s, t)$, where

$$\hat{\beta}_n(s, t) = \begin{cases} n^{-1/2} \sum_{i=1}^{[(n+1)t]} \left(\mathbf{1}\{R_{in} \leq s\} - \frac{[ns]}{n} \right), & 0 \leq t < 1, \\ & 0 \leq s \leq 1, \\ 0, & t = 1. \end{cases}$$

For other possible weight functions we refer to the discussion at the end of Section 8.2, where we also consider the weighted L_p functionals of $\alpha_n(s, t)$ process. As an immediate consequence we obtain

$$\int_0^1 \int_0^1 |\hat{\beta}_n(s, t)|^p / q(t) ds dt \xrightarrow{\mathcal{D}} \int_0^1 \int_0^1 |\Gamma(s, t)|^p / q(t) ds dt$$

for any q positive on $(0, 1)$ and such that

$$\int_0^1 (t(1-t))^{p/2}/q(t)dt < \infty.$$

Considering asymptotics of $\hat{\beta}_n(s, t)$ under changepoint alternatives H_1^η we obtain

$$\begin{aligned} \hat{\beta}_n(s, t)/q(t) \xrightarrow{\mathcal{D}} & \left(\Gamma(s, t) - (t(1-\eta)1\{t < \eta\} \right. \\ & \left. + (1-t)\eta 1\{t \geq \eta\}) \int_0^s g(u)du \right) / q(t) \end{aligned}$$

for any q positive on $(0, 1)$ and such that (3.5.4) holds. Consequently any continuous functional of $\hat{\beta}_n(s, t)/q(t)$ converges to the same functional of the limiting process. The above result holds also when replacing the process $\hat{\beta}_n(s, t)$ by $\hat{\alpha}_n(s, t)$.

CHAPTER 4

Weighted partial sum processes of sequential ranks

4.0. Introduction

Let $X_1, X_2, \dots$ be independent continuously distributed random variables with unknown distribution functions. For each $n \geq 1$, let $R_{1n}, \dots, R_{nn}$ denote the normalized ranks

$$R_{kn} = n^{-1} \sum_{i=1}^n 1\{X_i \leq X_k\}, \quad k = 1, \dots, n,$$

and let $\xi_1, \dots, \xi_n$ be the normalized sequential ranks

$$\xi_k = k^{-1} \sum_{i=1}^k 1\{X_i < X_k\}, \quad k = 1, \dots, n,$$

of the first n of the random variables $X_1, X_2, \dots$. Thus, the rank of X_k , in our notation nR_{kn} , is the integer among the numbers $1, \dots, n$ which corresponds to the position of X_k in the order statistics $X_{1n} \leq \dots \leq X_{nn}$ of a random sample of size $n \geq 1$, while the sequential rank of X_k , here $k\xi_k$, is the position of X_k in the first $k-1$ ordered observations $X_{0,k-1} \leq \dots \leq X_{k-1,k-1}$, $1 \leq k \leq n$ of a random sample of size $n \geq 1$, where $X_{0,0} \equiv 0$ by convention, and $\xi_1 = 0$ by definition.

When observations are available sequentially and to be kept in their chronological order, then, for computing nR_{kn} , one has to recompute the ranks of all the available observations at each stage. This is often quite tedious as well as time consuming. However, computing only the rank of the last observation at each stage gives the sequential ranks $k\xi_k$ immediately. Hence it is, comparatively, also easy to calculate statistics of sequential ranks. In particular, one can calculate these statistics successively as the sequential ranks are calculated, in contrast to

statistics based on ranks that require the entire sequence $R_{1n}, \dots, R_{nn}$ to be calculated. Moreover, sequential ranks have the important property that, under the null hypothesis of all the independent X_i 's being also identically distributed, they themselves are independent (in contrast to ranks). The following result is due to Barndorff-Nielsen (1963).

Theorem 4.0.A. *If the random variables $X_1, \dots, X_n$ are independent and identically continuously distributed, then the sequential ranks $\xi_1, \dots, \xi_n$ are independent and*

$$(4.0.1) \quad P\left\{\xi_k = \frac{i}{k}\right\} = \frac{1}{k}, \quad i = 0, \dots, k-1; \quad k = 1, \dots, n.$$

Statistical tests based on sequential ranks are frequently used because of these convenient properties. Sequential ranks are ready made, of course for sequential procedures. However, they are also very convenient to use for sequential studies of chronologically ordered observations of finite size samples.

For related works on statistics based on sequential ranks we refer to Parent (1965), Reynolds (1975), Mason (1981), Lombard and Mason (1985). Bhattacharya and Frierson (1981) propose a control chart based on partial sums of sequential ranks for detecting small changes in the distribution of a given random sample. They consider $\max_{1 \leq k \leq n} \sum_{i=1}^k \xi_i / n^{1/2}$ type statistics while Gombay (1991) studies their sequentially weighted versions $\max_{1 \leq k \leq n} \sum_{i=1}^k \xi_i / k^{1/2}$ instead, via a theorem of Darling and Erdős (1956). Csörgő and Horváth (1987b) propose a sequential procedure based on empirical process of sequential ranks for detecting a possible changepoint in a random sequence. Asymptotic properties of CUSUM versions of these procedures are studied by Huse (1989). Test statistics based on sequential ranks under contiguous alternatives and in connection with changepoint problems are considered by Lombard (1981, 1983).

Khmaladze and Parjanadze (1986) study the asymptotic distribution of processes formed by partial sums of sequential ranks. They consider the process

$$(4.0.2) \quad V_n(t) = n^{-1/2} \sum_{i=1}^{[nt]} a(\xi_i), \quad 0 \leq t \leq 1,$$

for $a \in L^2(0,1)$ and such that $\int_0^1 a(u)du = 0$, where $\xi_1, \xi_2, \dots$ are normalized sequential ranks. For suitable a , they derive the limiting distribution of this process under a class of contiguous alternatives as described in Chapter 2 (cf. (2.3.1) – (2.3.3)), which accommodates the occurrence of a changepoint in a series of chronologically ordered measurements.

To make statistical tests more sensitive to detect a possible change which has occurred at the early stage of random sampling, one may wish to introduce weight functions which would emphasize such a change.

In Section 4.1 we prove weighted approximations of the partial sum process V_n as in (4.0.2) by a standard Wiener process in supremum and L_p , $0 < p < \infty$, distances. The obtained results are optimal in terms of moment conditions and the class of weight functions. Since $k\xi_k$ are independent, uniformly distributed random variables over $\{0, 1, \dots, k-1\}$, $1 \leq k \leq n$, (cf. (4.0.1)), we can represent the normalized sequential ranks ξ_k in terms of uniformly distributed random variables. Namely, with independent random variables U_k , uniformly distributed on $(0, 1)$, we have

$$(4.0.3) \quad \xi_k \stackrel{D}{=} \frac{[kU_k]}{k}, \quad k = 1, 2, \dots$$

Such a representation of sequential ranks was used by Csörgő and Horváth (1987b). To prove our results in Section 4.1, instead of $V_n(t)$ we consider the equivalent process

$$\tilde{V}_n(t) = n^{-1/2} \sum_{i=1}^{[nt]} a\left(\frac{[iU_i]}{i}\right), \quad 0 \leq t \leq 1,$$

and show that $\tilde{V}_n$ is close enough to the process V_n to inherit the weighted approximations of the partial sum process of i.i.d.r.v.'s proved in Chapter 1. In Section 4.2 we show that the weak convergence of the process $V_n(t)/q(t)$ continues to hold under the sequence of the contiguous alternatives H_1 (cf. (2.3.1) – (2.3.3)) for the same class of weight functions as in the case of the null assumption of X_i 's being identically distributed. When doing this, we would wish to use Le Cam's third lemma again just like we did in Chapter 2, when studying the asymptotics of weighted partial sums of independent random variables under the sequence of contiguous alternatives H_1 . However, the idea of Section 4.1 cannot be used in this case, since the joint limit of our representation process $\tilde{V}_n$ with the log-likelihood process L_n (cf. (2.3.4)), i.e. that of $(\tilde{V}_n/q, L_n)$, does not give the proper covariance structure. Combining a result of Khmaladze and Parjanadze (1986), where they prove the weak convergence of (V_n, L_n) to the appropriate Gaussian process, with our results of Section 4.1, will allow us to obtain the correct limit distribution of the $V_n(t)/q(t)$ process under the desired contiguous alternatives.

4.1. Null hypothesis

Let $X_1, X_2, \dots$ be independent continuously distributed random variables. In this section we study the asymptotic behaviour of the process

$$n^{-1/2} \sum_{k=1}^{[nt]} a(\xi_k), \quad 0 \leq t \leq 1$$

for $a \in L^2(0, 1)$ and such that $\int_0^1 a(u) du = 0$, where $\xi_1, \xi_2, \dots$ are normalized sequential ranks, in weighted supremum and L_p -distances under the null hypothesis of $X_1, X_2, \dots$ being identically distributed random variables.

We recall that $\mathcal{Q}$ is the class of positive functions i.e., such that $\inf_{\delta \leq t \leq 1} q(t) > 0$ for all $0 < \delta < 1$, which are nondecreasing in a neighbourhood of zero. Let again

$$I(q, c) = \int_0^1 t^{-1} \exp(-ct^{-1}q^2(t))dt$$

and

$$E(q, c) = \int_0^1 t^{-3/2} q(t) \exp(-ct^{-1}q^2(t))dt, \quad 0 < c < \infty.$$

Theorem 4.1.1. *Let H_0 hold and assume that $a \in L^2[0, 1]$, where $\int_0^1 a(u)du = 0$, $\int_0^1 a^2(u)du = 1$ and $a(u)$ has a uniformly bounded derivative on $[0, 1]$.*

Let $q \in \mathcal{Q}$. Then there exists a sequence of Wiener processes $\{W_n(t), 0 \leq t \leq 1\}$ such that

$$(4.1.1) \quad \sup_{0 < t \leq 1} \left| n^{-1/2} \sum_{i=1}^{[nt]} a(\xi_i) - W_n(t) \right| / q(t) = o_P(1)$$

if and only if $I(q, c) < \infty$ for all $c > 0$, and

$$(4.1.2) \quad \sup_{0 < t \leq 1} \left| n^{-1/2} \sum_{i=1}^{[nt]} a(\xi_i) - W_n(t) \right| / q(t) = O_P(1)$$

if and only if $I(q, c) < \infty$ for some $c > 0$.

Corollary 4.1.1. *Assume that H_0 holds and $a(u)$ is a function as in Theorem 4.2.1.*

(a) *Let $q \in \mathcal{Q}$. Then*

$$(4.1.3) \quad n^{-1/2} \sum_{i=1}^{[nt]} a(\xi_i) / q(t) \xrightarrow{D} W(t) / q(t) \quad \text{in } D[0, 1]$$

if and only if $I(q, c) < \infty$ for all $c > 0$, where W is a standard Wiener process.

(b) *Let $q : (0, 1] \rightarrow (0, \infty)$ be such that $\inf_{\delta \leq t \leq 1} q(t) > 0$ for all $\delta \in (0, 1)$. Then (4.1.3) holds in $D[\delta, 1]$ for any $\delta \in (0, 1)$.*

Remark 4.1.1. According to Theorem 1.1.C, the statement (4.1.1) of Theorem 4.1.1 as well as (4.1.3) of Corollary 4.1.1 is also equivalent to

$$E(q, c) < \infty \text{ for all } c > 0 \text{ and } \lim_{t \downarrow 0} q(t)/t^{1/2} = \infty.$$

Similarly, according to Theorem 1.1.B, the statement (4.1.2) of Theorem 4.1.1 is also equivalent to

$$E(q, c) < \infty \text{ for some } c > 0 \text{ and } \lim_{t \downarrow 0} q(t)/t^{1/2} = \infty.$$

In order to derive our results we represent the normalized sequential ranks ξ_i in terms of uniformly distributed random variables. As noted already (cf. Theorem 4.0.A.) ξ_i are independent with distribution

$$P\left\{\xi_i = \frac{k}{i}\right\} = \frac{1}{i}, \quad 0 \leq k < i, \quad 1 \leq i \leq n,$$

Hence, with independent random variables U_i , uniformly distributed on $[0, 1]$, we have

$$(4.1.4) \quad \xi_i \stackrel{D}{=} \frac{[iU_i]}{i}, \quad i = 1, 2, \dots$$

Lemma 4.1.1. *Let $U_1, U_2, \dots$ be independent uniform $[0, 1]$ random variables and assume that $a(u)$ has a uniformly bounded derivative on $[0, 1]$. Then we have, as $n \rightarrow \infty$*

$$(4.1.5) \quad \sup_{0 \leq t \leq 1} \left| n^{-1/2} \left(\sum_{i=1}^{[nt]} a\left(\frac{[iU_i]}{i}\right) - \sum_{i=1}^{[nt]} a(U_i) \right) \right| \stackrel{a.s.}{=} o(1)$$

and

$$(4.1.6) \quad \sup_{1/n \leq t \leq 1} \left| n^{-1/2} \left(\sum_{i=1}^{[nt]} a\left(\frac{[iU_i]}{i}\right) - \sum_{i=1}^{[nt]} a(U_i) \right) \right| / t^{1/2} \stackrel{a.s.}{=} O(1).$$

Proof. First we note that with $[iU_i]/i \wedge U_i \leq \eta_i \leq [iU_i] \vee U_i$, $i = 1, \dots, n$, there is a constant $C > 0$ such that

$$\begin{aligned} & \sup_{0 \leq t \leq 1} \left| \sum_{i \leq nt} a\left(\frac{[iU_i]}{i}\right) - \sum_{i \leq nt} a(U_i) \right| \\ & \leq \sup_{0 \leq t \leq 1} \left| \sum_{i \leq nt} a'(\eta_i) \left(\frac{[iU_i]}{i} - U_i \right) \right| \\ & \leq \sup_{0 \leq t \leq 1} C \sum_{i \leq nt} \frac{|[iU_i] - iU_i|}{i} \\ & \leq C \log n, \end{aligned}$$

which implies (4.1.5), as well as that for $nt \rightarrow \infty$ we have

$$\left| \sum_{i \leq nt} a\left(\frac{[iU_i]}{i}\right) - \sum_{i \leq nt} a(U_i) \right| \stackrel{a.s.}{=} O(\log nt).$$

Hence, as $n \rightarrow \infty$,

$$\sup_{1/n \leq t \leq 1} \left| \sum_{i \leq nt} a\left(\frac{[iU_i]}{i}\right) - \sum_{i \leq nt} a(U_i) \right| / (nt)^{1/2} \stackrel{a.s.}{=} O(1).$$

Remark 4.1.2. We note that, on account of (4.1.4), the two stochastic processes

$$\left\{ \tilde{V}_n(t) = n^{-1/2} \sum_{i=1}^{[nt]} a\left(\frac{[iU_i]}{i}\right), \quad 0 \leq t \leq 1 \right\}$$

and

$$\left\{ V_n(t) = n^{-1/2} \sum_{i=1}^{[nt]} a(\xi_i), \quad 0 \leq t \leq 1 \right\}$$

are equivalent.

Proof of Theorem 4.1.1. Let $q \in \mathcal{Q}$ be such that

$$(4.1.7) \quad \lim_{t \downarrow 0} t^{1/2}/q(t) = 0.$$

We note that $I(q, c) < \infty$ for some $c > 0$ implies (4.1.7) (cf. Theorem 1.1.A). Since q is a positive function, i.e., $\inf_{\delta \leq t \leq 1} q(t) > 0$ for all $0 < \delta < 1$, by (4.1.5) of Lemma 4.1.1 we have, as $n \rightarrow \infty$

$$(4.1.8) \quad \sup_{\delta \leq t \leq 1} \left| n^{-1/2} \sum_{i \leq nt} a\left(\frac{[iU_i]}{i}\right) - n^{-1/2} \sum_{i \leq nt} a(U_i) \right| / q(t) = o(1)$$

for all $\delta \in (0, 1]$. By (4.1.6) of Lemma 4.1.1 we conclude that

$$\begin{aligned} & \sup_{\frac{1}{n} \leq t \leq \delta} \left| n^{-1/2} \sum_{i \leq nt} a\left(\frac{[iU_i]}{i}\right) - n^{-1/2} \sum_{i \leq nt} a(U_i) \right| / q(t), \\ & \leq \sup_{\frac{1}{n} \leq t \leq \delta} \left| n^{-1/2} \sum_{i \leq nt} a\left(\frac{[iU_i]}{i}\right) - n^{1/2} \sum_{i \leq nt} a(U_i) \right| / t^{1/2} \sup_{0 < t \leq \delta} t^{1/2} / q(t) \\ & \leq O(1) \sup_{0 < t \leq \delta} t^{1/2} / q(t) \end{aligned}$$

as $n \rightarrow \infty$ for any $\delta \in (0, 1)$. Since $\delta > 0$ can be taken arbitrarily small, the latter combined with (4.1.8) yields

$$\sup_{0 < t \leq 1} \left| n^{-1/2} \sum_{i=1}^{[nt]} a\left(\frac{[iU_i]}{i}\right) - n^{-1/2} \sum_{i=1}^{[nt]} a(U_i) \right| / q(t) = o(1).$$

Hence, by Theorem 1.2.1, we conclude that one can construct a Wiener process $\{W(t), 0 \leq t < \infty\}$ such that with $q \in \mathcal{Q}$ we have, as $n \rightarrow \infty$,

$$\begin{aligned} & \sup_{0 < t < 1} \left| n^{-1/2} \sum_{i=1}^{[nt]} a\left(\frac{[iU_i]}{i}\right) - n^{-1/2} W(nt) \right| / q(t) \\ & = \begin{cases} o_P(1) & \text{if and only if } I(q, c) < \infty \text{ for all } c > 0 \\ O_P(1) & \text{if and only if } I(q, c) < \infty \text{ for some } c > 0. \end{cases} \end{aligned}$$

The latter implies (4.1.1) and (4.1.2) due to (4.1.4) (cf. Remark 4.1.2) when replacing $n^{-1/2}W(nt)$ by a possibly different sequence of Wiener processes W_n . The general validity of this technique is described in Lemma 4.4.4 of Csörgő and Révész (1981) (cf. Section A.1).

Proof of Corollary 4.1.1. Let $V_n(t) = n^{-1/2} \sum_{i=1}^{[nt]} a(\xi_i)$.

(a) Obviously, with $q \in \mathcal{Q}$ and $I(q, c) < \infty$ for all $c > 0$, by (4.1.1) of Theorem 4.1.1 we have

$$(4.1.9) \quad V_n(t)/q(t) \xrightarrow{\mathcal{D}} W(t)/q(t) \quad \text{in } D[0, 1],$$

where W is a standard Wiener process.

Assuming now (4.1.9) with $q \in \mathcal{Q}$, by the Skorohod-Dudley-Wichura theorem (cf. Theorem A.2.1) there exists V_n^* , $n \geq 1$, and W^* such that

$$V_n^* \stackrel{\mathcal{D}}{=} V_n, \quad n \geq 1, \quad W^* \stackrel{\mathcal{D}}{=} W$$

and

$$\sup_{0 \leq t \leq 1} |V_n^*(t) - W^*(t)|/q(t) = o_P(1).$$

Hence, using the technique of a lemma of Csörgő and Révész (cf. Section A.1), there exists a sequence of Wiener processes $\{W_n(t), 0 \leq t \leq 1\}$ such that

$$\sup_{0 \leq t \leq 1} |V_n(t) - W_n(t)|/q(t) = o_P(1),$$

which, by (4.1.1) of Theorem 4.1.1 implies $I(q, c) < \infty$ for all $c > 0$.

(b) By (4.1.1) of Theorem 4.1.1 with $q \equiv 1$, we have

$$\sup_{0 \leq t \leq 1} |V_n(t) - n^{-1/2}W(nt)| = o_P(1),$$

and hence with any $q : (0, 1] \rightarrow (0, \infty)$ such that $\inf_{\delta \leq t \leq 1} q(t) > 0$ we have

$$\sup_{\delta \leq t \leq 1} |V_n(t) - n^{-1/2}W(nt)|/q(t) = o_P(1)$$

for any $\delta \in (0, 1)$.

For L_p -functionals we obtain the following description.

Theorem 4.1.2. *Let $0 < p < \infty$. We assume that $a \in L^2[0,1]$, where $\int_0^1 a(u)du = 0$, $\int_0^1 a^2(u)du = 1$ and $a(u)$ has a uniformly bounded derivative on $[0,1]$.*

(a) *Let $q \in \mathcal{Q}$. Then there exists a sequence of Wiener processes $\{W_n(t), 0 \leq t \leq 1\}$ such that*

$$(4.1.10) \quad \int_0^1 \left| n^{-1/2} \sum_{i=1}^{[nt]} a(\xi_i) - W_n(t) \right|^p / q(t) dt = o_P(1)$$

if and only if

$$(4.1.11) \quad \int_0^1 t^{1/2} / q(t) dt < \infty.$$

(b) *If $q \in \mathcal{Q}$ then*

$$(4.1.12) \quad \int_0^1 \left| n^{-1/2} \sum_{i=1}^{[nt]} a(\xi_i) \right|^p / q(t) dt \xrightarrow{D} \int_0^1 |W(t)|^p / q(t) dt$$

if and only if (4.1.11) holds.

Proof of Theorem 4.1.2. (a) We assume that (4.1.11) holds. Then

$$\begin{aligned} & \int_0^1 \left| n^{-1/2} \left(\sum_{i=1}^{[nt]} a\left(\frac{[iU_i]}{i}\right) - \sum_{i=1}^{[nt]} a(U_i) \right) \right|^p / q(t) \\ & \leq \sup_{1/n \leq t \leq 1} \left(\left| n^{-1/2} \left(\sum_{i=1}^{[nt]} a\left(\frac{[iU_i]}{i}\right) - \sum_{i=1}^{[nt]} a(U_i) \right) \right| / t^{1/2} \right)^p \int_0^\delta t^{p/2} / q(t) dt \\ & + \int_\delta^1 \left| n^{-1/2} \left(\sum_{i=1}^{[nt]} a\left(\frac{[iU_i]}{i}\right) - \sum_{i=1}^{[nt]} a(U_i) \right) \right|^p / q(t) = o_P(1), \end{aligned}$$

when $n \rightarrow \infty$ and taking $\delta > 0$ arbitrarily small. Let $\{W(t), 0 \leq t < \infty\}$ be the Wiener process as defined in Theorem 1.3.1, where we now put $X_i = a(U_i)$, $i = 1, 2, \dots$. Then we have

$$\begin{aligned} & \int_0^1 \left| n^{-1/2} \left(\sum_{i=1}^{[nt]} a\left(\frac{[iU_i]}{i}\right) - W(nt) \right) \right|^p / q(t) dt \\ & \leq 2^p \int_0^1 \left| n^{-1/2} \left(\sum_{i=1}^{[nt]} a\left(\frac{[iU_i]}{i}\right) - \sum_{i=1}^{[nt]} a(U_i) \right) \right|^p / q(t) dt \\ & + 2^p \int_0^1 \left| n^{-1/2} \left(\sum_{i=1}^{[nt]} a(U_i) - W(nt) \right) \right|^p / q(t) dt \\ & = o_P(1), \end{aligned}$$

which implies (4.1.10) due to (4.1.4) (cf Remark 4.1.2) when replacing $n^{-1/2}W(nt)$ by a possibly different sequence of Wiener processes W_n (cf. Section A.1).

Next we assume (4.1.10) to hold. Then

$$\int_0^{1/n} |W_n(t)|^p / q(t) dt = o_P(1),$$

and we use an argument as in the proof of Theorem 1.4.1 to conclude (4.1.11).

(b) Same as the proof of part (b) of Theorem 1.3.1.

4.2. Contiguous Alternatives

After describing the weighted asymptotics of $V_n(t)$ under the null hypothesis

$$H_0: X_i, 1 \leq i \leq n, \text{ have the same distribution } F,$$

we study the $V_n(t)$ process with weight functions $q(t)$ under the sequence of alternative hypotheses introduced already in Chapter 2, namely

$$H_1: X_i, 1 \leq i \leq n, \text{ have the respective distribution functions } F_{in},$$

where we assume that all F_{in} are absolutely continuous with respect to the distribution function F and

$$\left[\frac{dF_{in}}{dF}(F^{-1}(u)) \right]^{1/2} = 1 + \frac{1}{2\sqrt{n}} g_n(t, u), \quad \frac{i-1}{n} < t \leq \frac{i}{n},$$

where $F^{-1}(u) = \inf\{x : (F(x) \geq u)\}$, $0 < u \leq 1$, $F^{-1}(0) = F^{-1}(0+)$ (cf. (2.3.1)).

We assume also that there exists a function $g \in L^2[0, 1]^2$, such that

$$\int_0^1 g(t, u) du = 0 \quad \text{for almost all } t \in [0, 1]$$

and

$$\int_0^1 \int_0^1 \left[g_n(t, u) - g\left(\frac{[nt]}{n}, u\right) \right]^2 du dt \rightarrow 0, \quad n \rightarrow \infty,$$

(cf. (2.3.1) – (2.3.3)). As noted already, the sequence of direct products $F_{1n} \times \dots \times F_{nn}$, $n = 1, 2, \dots$ is contiguous to the sequence $F \times \dots \times F$. In particular, for the changepoint alternatives H_1^n we assume that there exists $\eta \in (0, 1)$ such that

$$(4.2.1) \quad g(t, u) = 1\{t \geq \eta\}g(u)$$

for some square integrable function g . Such description of contiguous alternatives was used by Khmaladze and Parjanadze (1986) in the context of studying changepoint problems using linear statistics of sequential ranks.

Here we show that the weak convergence of the weighted $V_n(\cdot)$ process under the sequence of contiguous alternatives of H_1 holds for the same optimal class of weight functions as in the case of the null hypothesis H_0 (cf. Corollary 4.1.1).

As before, let $\mathcal{Q}$ be the class of functions q defined on $(0, 1]$ which are nondecreasing near zero and $\inf_{\delta \leq t \leq 1} q(t) > 0$ for all $0 < \delta < 1$ and $\{W(t), 0 \leq t \leq 1\}$ be a standard Wiener process. We recall also the integrals

$$I(q, c) = \int_0^1 t^{-1} \exp(-ct^{-1}q^2(t)) dt$$

and

$$E(q, c) = \int_0^1 t^{-3/2} \exp(-ct^{-1}q^2(t))dt, \quad 0 < c < \infty.$$

Theorem 4.2.1. *Let H_1 hold and assume that $a \in L^2[0, 1]$, where $\int_0^1 a(u)du = 0$, $\int_0^1 a^2(u)du = 1$ and $a(u)$ has a uniformly bounded derivative on $[0, 1]$.*

(a) *Let $q \in \mathcal{Q}$. If $I(q, c) < \infty$ for all $c > 0$, then*

$$(4.2.2) \quad n^{-1/2} \sum_{i=1}^{[nt]} a(\xi_i)/q(t) \xrightarrow{D} (W(t) + R(t))/q(t) \quad \text{in } D[0, 1],$$

where

$$R(t) = - \int_0^t \frac{Q(s)}{s} ds + Q(t)$$

and

$$Q(t) = \int_0^t \int_0^1 g(s, u)a(u)duds.$$

(b) *Let $q : (0, 1] \rightarrow (0, \infty)$ be such that $\inf_{\delta \leq t \leq 1} q(t) > 0$ for all $\delta \in (0, 1)$. Then (4.2.2) holds in $D[\delta, 1]$ for any $\delta \in (0, 1)$.*

Remark 4.2.1. According to Theorem 1.1.C, part (a) Theorem 4.2.1 can be stated equivalently in terms of the integral $E(q, c)$ (cf. Remark 2.5.1).

We noted already that sequential ranks from i.i.d. observations, i.e. $k\xi_k$ in our terminology, are independent, the i th being uniformly distributed on $\{0, \dots, i-1\}$. Thus under the hypothesis H_0 the functional limit theorem for the partial sums process V_n follows immediately from Donsker's theorem. But for the alternatives H_1 the situation is different. For the sake of applying Le Cam's third lemma in this setup, one is interested in the asymptotic distribution of (V_n, L_n) under H_0 , where L_n is the centered log-likelihood ratio process defined by

$$(4.2.3) \quad L_n(t) = \sum_{i=1}^{[nt]} \left\{ \log \frac{dF_{in}}{dF}(F^{-1}(\bar{U}_i)) - E \log \frac{dF_{in}}{dF}(F^{-1}(\bar{U}_i)) \right\},$$

where

$$\bar{U}_i = F(X_i).$$

An analogous situation is provided by the example of studying

$$\left(n^{-1/2} \sum_{i \leq nt} c_i a(R_{in}), L_n(t) \right)$$

is a similar vein (cf. Hájek and Šidák, 1967).

However, as Khmaladze and Parjanadze (1986) point out, the classical method to obtain the joint limit theorem for linear rank statistics $\sum c_i a(R_{in})$ and L_n cannot be applied here. The essence of this method is that the relation

$$n^{-1/2} \sum_{i=1}^n c_i a(R_{in}) - n^{-1/2} \sum_{i=1}^n c_i \varphi(\bar{U}_i) = o_P(1), \quad n \rightarrow \infty$$

is established, where $\sum c_i = 0$, $\bar{U}_i = F(X_i)$ and, hence, it is sufficient to prove the joint limit theorem for the two sums of independent random variables $n^{-1/2} \sum_{i=1}^n c_i \varphi(F(X_i))$ and L_n . However, in the case of sequential ranks Khmaladze and Parjanadze (1986) prove the following lemma.

Lemma 4.2.A. *Let $U_1, U_2, \dots$ be independent uniform $-[0, 1]$ random variables and for each $n \geq 1$, let $\xi_1, \dots, \xi_n$ be the normalized sequential ranks of the random sample $U_1, \dots, U_n$. Then for any $t \in [0, 1]$*

$$n^{-1/2} \sum_{i=1}^{[nt]} a(\xi_i) - n^{-1/2} \sum_{i=1}^{[nt]} \varphi(U_i) \neq o_P(1)$$

for any functions a and φ such that

$$\int_0^1 a(u) du = \int_0^1 \varphi(u) du = 0, \quad 0 < \int_0^1 a^2(u) du < \infty,$$

where a is left continuous and has finite right-hand limits.

Nevertheless, using martingale methods and assuming some mild, appropriate conditions on a , Khmaladze and Parjanadze (1986) obtain the following results. Let

$$(4.2.4) \quad V(t) = - \int_0^t \frac{W(s)}{s} ds + W(t),$$

where W is a standard Wiener process. Then, given H_0 ,

$$(4.2.5) \quad (V_n, L_n) \xrightarrow{D} (V, L) \text{ in } D^2[0, 1],$$

where (V, L) is a two-dimensional Gaussian process, V is a standard Wiener process, L is a Gaussian process with mean zero and covariance function $Q_L(t_1 \wedge t_2)$, where

$$Q_L(t) = \int_0^t \int_0^1 g^2(s, u) du ds,$$

and the covariance function of the process (V, L) is

$$(4.2.6) \quad Q_{V,L}(t_1, t_2) = EV(t_1)L(t_2) = - \int_0^{t_1} \frac{Q(s \wedge t_2)}{s} ds + Q(t_1 \wedge t_2),$$

with

$$(4.2.7) \quad Q(t) = \int_0^t \int_0^1 g(s, u) a(u) du ds.$$

Under the alternatives of H_1 ,

$$(4.2.8) \quad V_n(t) \xrightarrow{D} V(t) + R(t) \text{ in } D[0, 1].$$

where

$$R(t) = - \int_0^t \frac{Q(s)}{s} ds + Q(t).$$

In particular, for the changepoint alternatives (4.2.1)

$$R(t) = \langle g, a \rangle - \eta \log(t/\eta) 1\{t \geq \eta\},$$

where

$$\langle g, a \rangle = \int_0^1 g(u)a(u)du.$$

Proof of Theorem 4.2.1. (a) Let $L_n(\cdot)$ be defined as in (4.2.3) and assume that H_0 holds. By (4.2.5) we have

$$(V_n, L_n) \xrightarrow{D} (V, L) \text{ in } D^2[0, 1],$$

where (V, L) is defined by (4.2.4), (4.2.6) and (4.2.7). We have

$$E \frac{V_n(t_1)}{q(t_1)} L_n(t_2) = \frac{1}{q(t_1)} E V_n(t_1) L_n(t_2) \rightarrow \frac{1}{q(t_1)} E V(t_1) L(t_2).$$

Hence, and using also Corollary 4.1.1, we conclude that under H_0

$$(V_n/q, L_n) \xrightarrow{D} (V/q, L) \text{ in } D^2[0, 1].$$

Now Le Cam's third lemma implies that under the alternatives H_1

$$V_n(t)/q(t) \xrightarrow{D} V(t)/q(t) + R(t)/q(t) \text{ in } D(0, 1],$$

where

$$R(t)/q(t) = \text{Cov}(V(t)/q(t), L(1)).$$

We note that the limiting process is finite almost surely. In order to see this, it suffices to note that since $g \in L^2[0, 1]^2$ and $\int_0^1 a^2(u)du = 1$, we have

$$\begin{aligned} |Q(t)| &= \left| \int_0^t \int_0^1 g(s, u)a(u)duds \right| \\ &\leq \int_0^t \left(\left(\int_0^1 g^2(s, u)du \right)^{1/2} \left(\int_0^1 a^2(u)du \right)^{1/2} \right) ds \\ &= \int_0^t \left(\int_0^1 g^2(s, u)du \right)^{1/2} ds \\ &\leq \left(\int_0^t \int_0^1 g^2(s, u)duds \right)^{1/2} \left(\int_0^t ds \right)^{1/2} \\ &= t^{1/2} \left(\int_0^t \int_0^1 g^2(s, u)duds \right)^{1/2} < \infty. \end{aligned}$$

Also,

$$\left| \int_0^t \frac{Q(s)}{s} ds \right| \leq \int_0^t \left| \frac{Q(s)}{s} \right| ds,$$

and since

$$\left| \frac{Q(s)}{s} \right| \leq s^{-1/2} \left(\int_0^1 \int_0^1 g^2(s, u) du ds \right)^{1/2},$$

we get

$$\left| \int_0^t \frac{Q(s)}{s} ds \right| \leq 2t^{1/2} \left(\int_0^1 \int_0^1 g^2(s, u) du ds \right)^{1/2} < \infty.$$

Since $q \in \mathcal{Q}$, we have $\inf_{\delta \leq t \leq 1} q(t) > 0$ for all $\delta \in (0, 1)$. Due to $I(q, c) < \infty$ for all $c > 0$, we have also $\lim_{t \downarrow 0} t^{1/2}/q(t) = 0$. Hence

$$\sup_{0 < t \leq 1} |R(t)|/q(t) < \infty.$$

(b) By (4.2.8), or part (a) of Theorem 4.2.1 with $q \equiv 1$ we have that under the alternative H_1

$$V_n(t) \xrightarrow{D} V(t) + R(t) \quad \text{in } D[0, 1]$$

Now we use the Skorohod-Dudley-Wichura theorem and the positivity of the function q and conclude the result the very same way as in the proof of Theorem 2.5.1 (b).

Corollary 4.2.1. *Let $a(u)$ be as in Theorem 4.2.1. Then under the change-point alternatives (4.2.1) we have*

(a) *Let $q \in \mathcal{Q}$. If $I(q, c) < \infty$ for all $c > 0$, then*

$$(4.2.9) \quad \frac{n^{-1/2} \sum_{i=1}^{\lfloor nt \rfloor} a(\xi_i)}{q(t)} \xrightarrow{D} \frac{W(t) + \eta \log(t/\eta) 1\{t \geq \eta\} \langle g, a \rangle}{q(t)} \quad \text{in } D[0, 1]$$

where

$$\langle g, a \rangle = \int_0^1 g(u) a(u) du$$

and W is a standard Wiener process.

(b) Let $q : (0, 1] \rightarrow (0, \infty)$ be such that $\inf_{\delta \leq t \leq 1} q(t) > 0$ for all $\delta \in (0, 1)$. Then (4.2.9) holds in $D[\delta, 1]$ for any $\delta \in (0, 1)$.

Proof. Let $g(t, u) = 1\{t \geq \eta\}g(u)$. Then

$$\begin{aligned} Q(t) &= \int_0^t \int_0^1 g(s, u) a(u) du ds \\ &= \int_0^t \int_0^1 1\{t \geq \eta\} g(u) a(u) du ds \\ &= (t - \eta) 1\{t \geq \eta\} \int_0^1 g(u) a(u) du. \end{aligned}$$

Also

$$\begin{aligned} \int_0^t \frac{Q(s)}{s} ds &= \left(\int_0^t 1\{s \geq \eta\} ds - \eta \int_0^t \frac{1}{s} 1\{s \geq \eta\} ds \right) \int_0^1 g(u) a(u) du \\ &= \left((t - \eta) 1\{t \geq \eta\} - \eta \log t / \eta 1\{t \geq \eta\} \right) \int_0^1 g(u) a(u) du. \end{aligned}$$

Hence

$$R(t) = - \int_0^t \frac{Q(s)}{s} ds + Q(t) = \eta \log t / \eta 1\{t \geq \eta\} \int_0^1 g(u) a(u) du.$$

CHAPTER 5

Tied Down Partial Sum Processes

5.0. Introduction

In Chapter 3 we illustrated that a large number of nonparametric as well as parametric modelling of AMOC problems result in the same test statistic, namely in

$$(5.0.1) \quad \max_{1 \leq k < n} \left\{ |S_k - kS_n/n| / (k(1 - k/n))^{1/2} \right\}.$$

Clearly then, when considering changepoint problems, the study of the limiting behaviour of the test statistic (5.0.1) is of interest. Unfortunately, given H_0 , as $n \rightarrow \infty$ we have

$$\max_{1 \leq k < n} \left(S_k - \frac{k}{n} S_n \right) / n^{1/2} \left(\frac{k}{n} \left(1 - \frac{k}{n} \right) \right)^{1/2} \xrightarrow{P} \infty.$$

Hence, in order to have the above test statistic converge in distribution to a nondegenerate random variable, we have to renormalize it, or introduce weight functions which are less severe on the tails than $((k/n)(1 - k/n))^{1/2}$ is.

The asymptotic behaviour of the statistic

$$\max_{1 \leq k < n} |S_k - kS_n/n|$$

is well known. Namely, under the null hypothesis of X_i 's being i.i.d.r.v.'s, Donsker's invariance principle yields

$$(S_{[nt]} - ([nt]/n)S_n) / (n^{1/2}\sigma) \xrightarrow{D} B(t) \quad \text{in } D[0, 1],$$

where $\{B(t), 0 \leq t \leq 1\}$ is a Brownian bridge.

Assuming that $E|X_1|^\nu < \infty$ for some $\nu > 2$, the asymptotic theory of (5.0.1) under the null hypothesis of independent, identically distributed observations was described by Csörgő and Horváth (1988) (cf. their Theorem 2.1 and Corollary 2.1) for the optimal class of weight functions in $\|\cdot/q\|$ -metrics. Here, in Section 5.2, we improve on their result under H_0 , assuming only two moments for X_1 , and extend it as well under the contiguous alternatives as parametrized in Chapter 2. In Section 5.3 we establish optimal approximations in weighted L_p -norms for tied down partial sum processes.

5.1. Local functions of a Brownian bridge

We assume that $\{W(t), 0 \leq t < \infty\}$ is a Wiener process and $\{B(t), 0 \leq t \leq 1\}$ is a Brownian bridge defined by

$$(5.1.1) \quad B(t) = W(t) - tW(1).$$

For each Wiener process, (5.1.1) is always a Brownian bridge and each Brownian bridge can be written in the form of (5.1.1) with a suitably chosen Wiener process. Hence when studying the behaviour of Brownian bridges, it will be enough to study the Wiener process. In Section 1.1 we presented some definitions and results concerning weighted Wiener processes. Here we summarize their extensions and modifications to Brownian bridges.

A function q defined on $(0, 1)$ will be called *positive* if $\inf_{\delta \leq t \leq 1-\delta} q(t) > 0$ for all $\delta \in (0, 1/2)$.

Let $\mathcal{Q}^*$ be the class of those positive functions on $(0, 1)$ which are nondecreasing in a neighbourhood of zero and nonincreasing in a neighbourhood of one.

A function $q \in \mathcal{Q}^*$ will be called a *local function of a Brownian bridge* if $q(t)$ and $q(1-t)$ are local functions of a Wiener process on $(0, 1/2]$ according to the previous definition (1.1.2) in Section 1.1. i.e. if in addition to

$$\limsup_{t \downarrow 0} |W(t)|/q(t) < \infty \quad a.s.$$

we have also

$$\limsup_{t \downarrow 0} |W(t)|/q(1-t) < \infty \quad a.s. .$$

Similarly, we extend the definition of a *Chibisov-O'Reilly function* of a Wiener process to that of a *Brownian bridge* and, consequently, we require

$$\lim_{t \downarrow 0} |W(t)|/q(t) = 0 \quad a.s.,$$

and

$$\lim_{t \downarrow 0} |W(t)|/q(1-t) = 0 \quad a.s.$$

to hold with $q \in \mathcal{Q}^*$.

In this case the integrals $E(q, c)$ and $I(q, c)$ (cf. (1.1.6) and (1.1.7)) will be replaced respectively by

$$E^*(q, c) = \int_0^1 (t(1-t))^{-3/2} q(t) \exp(-c(t(1-t))^{-1} q^2(t)) dt$$

and

$$I^*(q, c) = \int_0^1 (t(1-t))^{-1} \exp(-c(t(1-t))^{-1} q^2(t)) dt.$$

Theorems 1.1.A, 1.1.B and 1.1.C hold in their modified forms where the condition $\lim_{t \downarrow 0} q(t)/t^{1/2} = \infty$ is replaced by

$$(5.1.2) \quad \lim_{t \downarrow 0} q(t)/t^{1/2} = \lim_{t \uparrow 1} q(t)/(1-t)^{1/2} = \infty.$$

Theorem 5.1.A. (i) *Whenever the integral $I^*(q, c) < \infty$ for $q \in \mathcal{Q}^*$, then $E^*(q, c + \epsilon) < \infty$ for every $\epsilon > 0$ and (5.1.2) holds.*

(ii) *Whenever $E^*(q, c) < \infty$ and (5.1.2) holds for $q \in \mathcal{Q}^*$, then $I^*(q, c) < \infty$.*

Theorem 5.1.B. *A function $q \in \mathcal{Q}^*$ is a local function of a Brownian bridge if and only if the integral $I^*(q, c) < \infty$ for some $c > 0$ or, equivalently, if and only if the integral $E^*(q, c) < \infty$ for some $c > 0$ and (5.1.2) holds.*

Theorem 5.1.C. *A function $q \in \mathcal{Q}^*$ is a Chibisov-O'Reilly function of a Brownian bridge if and only if the integral $I^*(q, c) < \infty$ for all $c > 0$ or, equivalently, if and only if the integral $E^*(q, c) < \infty$ for all $c > 0$ and (5.1.2) holds.*

5.2. Supremum norm approximation

Let $X_1, X_2, \dots$ be independent, identically distributed random variables. With EX_1 not necessarily zero, and the X_i now are not to be centered at their unknown mean, we introduce

$$(5.2.1) \quad Z_n(t) = \begin{cases} (S((n+1)t) - [(n+1)t]S(n)/n)/(n^{1/2}\sigma), & 0 \leq t < 1 \\ 0, & t = 1, \end{cases}$$

where $\sigma^2 = E(X_1 - EX_1)^2$, $S((n+1)t) = \sum_{i=1}^{[(n+1)t]} X_i$, $S(n) = \sum_{i=1}^n X_i$ and $S((n+1)t) = 0$ for $0 \leq t < 1/(n+1)$, and note that $EZ_n(t) = 0$.

As in Section 5.1, let $\mathcal{Q}^*$ be the class of functions q defined on $(0, 1)$ which are nondecreasing near zero and nonincreasing near one, and such that

$$(5.2.2) \quad \inf_{\delta \leq t \leq 1-\delta} q(t) > 0 \quad \text{for all } \delta \in (0, 1/2).$$

We recall also the integrals:

$$E^*(q, c) = \int_0^1 (t(1-t))^{-3/2} q(t) \exp(-c(t(1-t))^{-1} q^2(t)) dt$$

and

$$I^*(q, c) = \int_0^1 (t(1-t))^{-1} \exp(-c(t(1-t))^{-1} q^2(t)) dt, \quad 0 < c < \infty.$$

Theorem 5.2.1 *Let $X_1, X_2, \dots$ be i.i.d.r.v.'s such that $EX_1^2 < \infty$. Then there exists a sequence of Brownian bridges $\{B_n(t), 0 \leq t \leq 1\}$ such that the following statements hold true.*

(a) *Let $q \in \mathcal{Q}^*$. Then*

$$\sup_{0 < t < 1} |Z_n(t) - B_n(t)|/q(t) = o_P(1)$$

if and only if $I^(q, c) < \infty$ for all $c > 0$.*

(b) *Let $q \in \mathcal{Q}^*$. Then*

$$\sup_{0 < t < 1} |Z_n(t) - B_n(t)|/q(t) = O_P(1)$$

if and only if $I^(q, c) < \infty$ for some $c > 0$.*

Remark 5.2.1. According to Theorems 5.1.C and 5.1.B, the results of Theorem 5.2.1 can be stated equivalently in terms of the integral $E^*(q, c)$ (cf. also Remark 1.2.1).

In order to present a proof of our theorem, we first state and prove some technical observations which will be seen to be useful also for proving further results.

Observation 5.2.1. We have

$$(5.2.3) \quad \sup_{0 < t < 1} \left| t - \frac{[(n+1)t]}{n} \right| = o(1),$$

$$(5.2.4) \quad \sup_{0 < t < 1/2} \left| t - \frac{[(n+1)t]}{n} \right| / t^{1/2} = O(1)$$

$$(5.2.5) \quad \sup_{1/2 < t < 1} \left| t - \frac{[(n+1)t]}{n} \right| / (1-t)^{1/2} = O(1).$$

Proof. By noting that

$$\sup_{0 < t < 1} \frac{|[(n+1)t] - nt|}{n} = \sup_{0 < t < 1} \frac{|[(n+1)t] - (n+1)t + t|}{n} \leq \frac{2}{n}$$

we have (5.2.3), and (5.2.4) holds since

$$\sup_{0 < t \leq 1/2} \frac{|[(n+1)t] - nt|}{nt^{1/2}} \leq \sup_{0 < t \leq 1/2} \frac{n+1}{n} t^{1/2} + \sup_{0 < t \leq 1/2} t^{1/2} = O(1).$$

In order to prove (5.2.5), we note that

$$(n+1)t - 1 < [(n+1)t]$$

and hence

$$\begin{aligned} \left| t - \frac{[(n+1)t]}{n} \right| &\leq |1-t| + \left| 1 - \frac{[(n+1)t]}{n} \right| \leq |1-t| + \left| \frac{n - (n+1)t + 1}{n} \right| \\ &\leq |1-t| + \left| \frac{(1-t)(n+1)}{n} \right|, \end{aligned}$$

which implies

$$\sup_{1/2 < t < 1} \left| t - \frac{[(n+1)t]}{n} \right| / (1-t)^{1/2} = O(1).$$

Lemma 5.2.1. *Let $X_1, X_2, \dots$ be i.i.d.r.v.'s and $S(n) = X_1 + \dots + X_n$. Let $\{W(t), 0 \leq t < \infty\}$ be a standard Wiener process such that*

$$(5.2.6) \quad n^{-1/2} |S(n) - W(n+1)| = o_P(1).$$

Then we have

$$(5.2.7) \quad \sup_{0 < t \leq 1/2} n^{-1/2} \left| \frac{[(n+1)t]}{n} S(n) - tW(n+1) \right| = o_P(1),$$

$$(5.2.8) \quad \sup_{0 < t \leq 1/2} n^{-1/2} \left| \frac{[(n+1)t]}{n} S(n) - tW(n+1) \right| / t^{1/2} = O_P(1),$$

$$(5.2.9) \quad \sup_{1/2 < t < 1} n^{-1/2} \left| \left(1 - \frac{[(n+1)t]}{n} \right) S(n) - (1-t)W(n+1) \right| = o_P(1),$$

$$(5.2.10) \quad \sup_{1/2 < t < 1} n^{-1/2} \left| \left(1 - \frac{[(n+1)t]}{n} \right) S(n) - (1-t)W(n+1) \right| / (1-t)^{1/2} = O_P(1).$$

Proof. Due to (5.2.6) and (5.2.3) we have

$$\begin{aligned}
& \sup_{0 < t \leq 1/2} |n^{-3/2}[(n+1)t]S(n) - n^{-1/2}tW(n+1)| \\
& \leq \sup_{0 < t \leq 1/2} n^{-1/2}t|S(n) - W(n+1)| \\
& \quad + \sup_{0 < t \leq 1/2} n^{-1/2}|[(n+1)t]/n - t|S(n)| \\
& \leq n^{-1/2}|S(n) - W(n+1)| \sup_{0 < t \leq 1/2} t \\
& \quad + n^{-1/2}|S(n)| \sup_{0 < t \leq 1/2} |[(n+1)t]/n - t| \\
& = o_P(1) + O_P(1)o(1) \\
& = o_P(1).
\end{aligned}$$

Consequently (5.2.7) holds true.

Similarly, using (5.2.6) and (5.2.4) we obtain

$$\begin{aligned}
& \sup_{0 < t \leq 1/2} |n^{-3/2}[(n+1)t]S(n) - n^{-1/2}tW(n+1)|/t^{1/2} \\
& \leq n^{-1/2}|S(n) - W(n+1)| \sup_{0 < t \leq 1/2} t^{1/2} \\
& \quad + n^{-1/2}|S(n)| \sup_{0 < t \leq 1/2} |[(n+1)t]/n - t|/t^{1/2} \\
& = o_P(1) + O_P(1)O(1) \\
& = O_P(1).
\end{aligned}$$

Hence (5.2.8) holds.

Considering (5.2.9), we have

$$\begin{aligned}
& \sup_{1/2 < t < 1} n^{-1/2} |(1 - [(n+1)t]/n)S(n) - (1-t)W(n+1)| \\
& \leq \sup_{1/2 < t < 1} n^{-1/2}(1-t)|S(n) - W(n+1)| \\
& \quad + \sup_{1/2 < t < 1} n^{-1/2} |(t - [(n+1)t]/n)S(n)| \\
& \leq n^{-1/2}|S(n) - W(n+1)| \sup_{1/2 < t < 1} (1-t) \\
& \quad + n^{-1/2}|S(n)| \sup_{1/2 < t < 1} \left| t - \frac{[(n+1)t]}{n} \right| \\
& = o_P(1),
\end{aligned}$$

due to (5.2.6) and (5.2.3).

We have also

$$\begin{aligned}
& \sup_{1/2 < t < 1} n^{-1/2} |(1 - [(n+1)t]/n)S(n) - (1-t)W(n+1)|/(1-t)^{1/2} \\
& \leq \sup_{1/2 < t < 1} n^{-1/2} (1-t)|S(n) - W(n+1)|/(1-t)^{1/2} \\
& + \sup_{1/2 < t < 1} n^{-1/2} |(t - [(n+1)t]/n)S(n)|/(1-t)^{1/2} \\
& \leq n^{-1/2} |S(n) - W(n+1)| \sup_{1/2 < t < 1} (1-t)^{1/2} \\
& + n^{-1/2} |S(n)| \sup_{1/2 < t < 1} |t - [(n+1)t]/n|/(1-t)^{1/2} \\
& = o_P(1) + O_P(1) \\
& = O_P(1),
\end{aligned}$$

due to (5.2.6) and (5.2.5). Hence (5.2.10) holds.

Proof of Theorem 5.2.1. For the sake of proof only, and without loss of generality, we will assume that $EX_1 = 0$ and $\sigma = 1$. We introduce

$$S^{(1)}((n+1)t) = \sum_{i=1}^{[(n+1)t]} X_i, \quad 0 \leq t \leq 1/2,$$

and

$$S^{(2)}((n+1)t) = \sum_{i=[(n+1)t]+1}^n X_i, \quad 1/2 \leq t < 1.$$

Then

(5.2.11)

$$S((n+1)t) = \begin{cases} S^{(1)}((n+1)t), & 0 \leq t \leq 1/2 \\ S^{(1)}\left(\frac{n+1}{2}\right) + S^{(2)}\left(\frac{n+1}{2}\right) - S^{(2)}((n+1)t), & 1/2 < t < 1. \end{cases}$$

First we show that one can construct two independent Wiener processes

$$\{W^{(1)}(t), 0 \leq t < \infty\} \text{ and } \{W^{(2)}(t), 0 \leq t < \infty\}$$

such that with $q \in \mathcal{Q}^*$

$$(5.2.12) \quad \sup_{0 < t \leq 1/2} \left| n^{-1/2} S^{(1)}((n+1)t) - n^{-1/2} W^{(1)}((n+1)t) \right| / q(t) \\ = \begin{cases} o_P(1) & \text{if and only if } I(q(t), c) < \infty \text{ for all } c > 0 \\ O_P(1) & \text{if and only if } I(q(t), c) < \infty \text{ for some } c > 0 \end{cases}$$

and

$$(5.2.13) \quad \sup_{1/2 \leq t < 1} \left| n^{-1/2} S^{(2)}((n+1)t) - n^{-1/2} W^{(2)}((n+1)(1-t)) \right| / q(t) \\ = \sup_{0 \leq t < 1/2} \left| n^{-1/2} S^{(2)}((n+1)(1-t)) - n^{-1/2} W^{(2)}((n+1)t) \right| / q(1-t) = o_P(1) \\ = \begin{cases} o_P(1) & \text{if and only if } I(q(1-t), c) < \infty \text{ for all } c > 0 \\ O_P(1) & \text{if and only if } I(q(1-t), c) < \infty \text{ for some } c > 0, \end{cases}$$

where $I(q, c)$ is defined in (1.1.7). While (5.2.12) holds immediately because of Theorem 1.2.1, in order to see that (5.2.13) is also a consequence of this theorem, we note that we have

$$S^{(2)}((n+1)(1-t)) = \begin{cases} 0 & 0 < t < 1/(n+1) \\ \sum_{i=1}^{[(n+1)t]} X_{n-i+1} & 1/(n+1) \leq t \leq 1/2. \end{cases}$$

We define a Wiener process W by

$$(5.2.14) \quad \{W(nt), 0 \leq t \leq 1\} = \begin{cases} W^{(1)}(nt), & 0 \leq t \leq 1/2 \\ W^{(1)}\left(\frac{n}{2}\right) + W^{(2)}\left(\frac{n}{2}\right) - W^{(2)}(n - nt), & 1/2 < t \leq 1. \end{cases}$$

Then we have

$$\sup_{0 < t \leq 1/2} n^{-1/2} \left| S((n+1)t) - [(n+1)t]S(n)/n - (W((n+1)t) - tW(n+1)) \right| / q(t) \\ \leq \sup_{0 < t \leq 1/2} n^{-1/2} \left| S^{(1)}((n+1)t) - W^{(1)}((n+1)t) \right| / q(t) \\ + \sup_{0 < t \leq 1/2} \left| n^{-3/2} [(n+1)t]S(n) - n^{-1/2} tW(n+1) \right| / q(t) \\ = I_1^{(1)}(n) + I_2^{(1)}(n).$$

We note that by definition of $S^{(1)}$ and $S^{(2)}$ we have for each $n \geq 1$

$$(5.2.15) \quad S^{(1)}\left(\frac{n+1}{2}\right) + S^{(2)}\left(\frac{n+1}{2}\right) = S(n),$$

and by (5.2.14) with $t = 1$

$$(5.2.16) \quad W^{(1)}\left(\frac{n+1}{2}\right) + W^{(2)}\left(\frac{n+1}{2}\right) = W(n+1).$$

Hence for $t \in (1/2, 1)$ we obtain

$$\begin{aligned} & \sup_{1/2 < t < 1} n^{-1/2} \left| S((n+1)t) - [(n+1)t]S(n)/n - (W((n+1)t) - tW(n+1)) \right| / q(t) \\ & \leq \sup_{1/2 < t < 1} n^{-1/2} \left| S^{(2)}((n+1)t) - W^{(2)}((n+1)(1-t)) \right| / q(t) \\ & + \sup_{1/2 < t < 1} \left| n^{-1/2} \left(S^{(1)}\left(\frac{n+1}{2}\right) + S^{(2)}\left(\frac{n+1}{2}\right) - [(n+1)t]S(n)/n \right) \right. \\ & \quad \left. - n^{-1/2} \left(W^{(1)}\left(\frac{n+1}{2}\right) + W^{(2)}\left(\frac{n+1}{2}\right) - tW(n+1) \right) \right| / q(t) \\ & = \sup_{1/2 < t < 1} n^{-1/2} \left| S^{(2)}((n+1)t) - W^{(2)}((n+1)(1-t)) \right| / q(t) \\ & + \sup_{1/2 < t < 1} n^{-1/2} \left| (1 - [(n+1)t]/n)S(n) - (1-t)W(n+1) \right| / q(t) \\ & = I_1^{(2)}(n) + I_2^{(2)}(n). \end{aligned}$$

Let $q \in \mathcal{Q}^*$. By (5.2.12) we have

$$I_1^{(1)}(n) = \begin{cases} o_P(1) & \text{if } I^*(q, c) < \infty \text{ for all } c > 0 \\ O_P(1) & \text{if } I^*(q, c) < \infty \text{ for some } c > 0. \end{cases}$$

Similarly, by (5.2.13)

$$I_1^{(2)}(n) = \begin{cases} o_P(1) & \text{if } I^*(q, c) < \infty \text{ for all } c > 0 \\ O_P(1) & \text{if } I^*(q, c) < \infty \text{ for some } c > 0. \end{cases}$$

Next we show that both terms $I_2^{(1)}(n)$ and $I_2^{(2)}(n)$ are $o_P(1)$ as $n \rightarrow \infty$, for any function q such that (5.1.2) holds, i.e.

$$\lim_{t \downarrow 0} q(t)/t^{1/2} = \lim_{t \uparrow 1} q(t)/(1-t)^{1/2} = \infty.$$

Since $I^*(q, c) < \infty$ for some $c > 0$ with $q \in \mathcal{Q}^*$ implies (5.1.2), this will also complete the proofs of the "if" parts of (a) and (b).

Using (5.2.15) and (5.2.16) we have

$$\begin{aligned}
 (5.2.17) \quad & n^{-1/2} |S(n) - W(n+1)| \\
 &= n^{-1/2} \left| S^{(1)}\left(\frac{n+1}{2}\right) + S^{(2)}\left(\frac{n+1}{2}\right) - \left(W^{(1)}\left(\frac{n+1}{2}\right) + W^{(2)}\left(\frac{n+1}{2}\right) \right) \right| \\
 &\leq n^{-1/2} \left| S^{(1)}\left(\frac{n+1}{2}\right) - W^{(1)}\left(\frac{n+1}{2}\right) \right| \\
 &\quad + n^{-1/2} \left| S^{(2)}\left(\frac{n+1}{2}\right) - W^{(2)}\left(\frac{n+1}{2}\right) \right| \\
 &= o_P(1),
 \end{aligned}$$

due to (5.2.12) and (5.2.13) applied with $q \equiv 1$ and $t = 1/2$.

Let $q \in \mathcal{Q}^*$ be such that (5.1.2) holds. Using (5.2.7), (5.2.8), and the positivity of q , we have for any $0 < \delta < 1/2$

$$\begin{aligned}
 I_2^{(1)}(n) &= \sup_{0 < t \leq 1/2} |n^{-3/2} [(n+1)t] S(n) - n^{-1/2} t W(n+1)| / q(t) \\
 &\leq \sup_{0 < t < \delta} |n^{-3/2} [(n+1)t] S(n) - n^{-1/2} t W(n+1)| / t^{1/2} \sup_{0 < t < \delta} t^{1/2} / q(t) \\
 &\quad + \sup_{\delta \leq t \leq 1/2} |n^{-3/2} [(n+1)t] S(n) - n^{-1/2} t W(n+1)| / q(t) \\
 &= O_P(1) \sup_{0 < t < \delta} t^{1/2} / q(t) + o_P(1).
 \end{aligned}$$

Similarly, using (5.2.9) and (5.2.10), we have

$$\begin{aligned}
 I_2^{(2)}(n) &= \sup_{1/2 < t < 1} n^{-1/2} |(1 - [(n+1)t]/n) S(n) - (1-t) W(n+1)| / q(t) \\
 &\leq \sup_{1/2 < t < 1-\delta} n^{-1/2} |(1 - [(n+1)t]/n) S(n) - (1-t) W(n+1)| / q(t) \\
 &\quad + \sup_{1-\delta < t < 1} n^{-1/2} |(1 - [(n+1)t]/n) S(n) - (1-t) W(n+1)| / (1-t)^{1/2} \\
 &\quad \times \sup_{1-\delta < t < 1} (1-t)^{1/2} / q(t) \\
 &= o_P(1) + O_P(1) \sup_{1-\delta < t < 1} (1-t)^{1/2} / q(t),
 \end{aligned}$$

for any $0 < \delta < 1/2$. Since $\delta > 0$ can be taken arbitrarily small, using (5.1.2) we get

$$I_2^{(1)}(n) = o_P(1) \quad \text{and} \quad I_2^{(2)}(n) = o_P(1).$$

The converse parts of (a) and (b) follow from the fact that

$$\sup_{0 < t < 1} |Z_n(t) - B_n(t)|/q(t) = o_P(1)$$

with $q \in \mathcal{Q}^*$, implies

$$o_P(1) = \sup_{0 < t < 1/(n+1)} |B_n(t)|/q(t) \stackrel{D}{=} \sup_{0 < t < 1/(n+1)} |B(t)|/q(t),$$

where $\{B(t), 0 \leq t \leq 1\}$ is a Brownian bridge, as well as

$$o_P(1) = \sup_{n/(n+1) \leq t < 1} |B_n(t)|/q(t) \stackrel{D}{=} \sup_{n/(n+1) \leq t < 1} |B(t)|/q(t),$$

which by Theorem 5.1.C is equivalent to $I^*(q, c) < \infty$ for all $c > 0$.

Similarly

$$\sup_{0 < t < 1} |Z_n(t) - B_n(t)|/q(t) = O_P(1)$$

with $q \in \mathcal{Q}^*$, implies

$$O_P(1) = \sup_{0 < t < 1/(n+1)} |B_n(t)|/q(t) \stackrel{D}{=} \sup_{0 < t < 1/(n+1)} |B(t)|/q(t),$$

as well as

$$O_P(1) = \sup_{n/(n+1) \leq t < 1} |B_n(t)|/q(t) \stackrel{D}{=} \sup_{n/(n+1) \leq t < 1} |B(t)|/q(t),$$

which by Theorem 5.1.B is equivalent to $I^*(q, c) < \infty$ for some $c > 0$.

Corollary 5.2.1. *Let $X_1, X_2, \dots$ be independent, identically distributed random variables such that $EX_1^2 < \infty$. Let $\{B(t), 0 \leq t \leq 1\}$ be a Brownian bridge and, for each $n \geq 1$, Z_n be defined as in (5.2.1).*

(a) Let $q \in Q^*$. Then

$$Z_n(t)/q(t) \xrightarrow{D} B(t)/q(t) \quad \text{in } D[0, 1]$$

if and only if $I^*(q, c) < \infty$ for all $c > 0$.

(b) Let $q : (0, 1) \rightarrow (0, \infty)$ be such that (5.2.2) holds. Then, as $n \rightarrow \infty$

$$Z_n(t)/q(t) \xrightarrow{D} B(t)/q(t) \quad \text{in } D[\delta, 1 - \delta]$$

for any $\delta \in (0, 1/2)$.

Proof. The proof of Corollary 1.2.1 can be repeated here after replacing $S_n(t)$ by $Z_n(t)$ and $W(t)$ by $B(t)$, respectively.

5.3. L_p -approximations

As before, let $Z_n(\cdot)$ be a sequence of stochastic processes as defined by (5.2.1) and Q^* be the class of functions $q : (0, 1) \rightarrow (0, \infty)$ which are nondecreasing near zero and nonincreasing near one, and such that $\inf_{\delta \leq t \leq 1-\delta} q(t) > 0$ for all $\delta \in (0, 1/2)$.

Theorem 5.3.1. Let $X_1, X_2, \dots$ be independent, identically distributed random variables such that $EX_1^2 < \infty$. Let $0 < p < \infty$ and $q \in Q^*$.

(a) There exists a sequence of Brownian bridges $\{B_n(t), 0 \leq t \leq 1\}$ such that

$$\int_0^1 |Z_n(t) - B_n(t)|^p / q(t) dt = o_P(1)$$

if and only if

$$(5.3.1) \quad \int_0^1 (t(1-t))^{p/2} / q(t) dt < \infty.$$

(b) Let $\{B(t), 0 \leq t \leq 1\}$ be a Brownian bridge. Then

$$\int_0^1 |Z_n(t)|^p / q(t) dt \xrightarrow{D} \int_0^1 |B(t)|^p / q(t) dt$$

if and only if (5.3.1) holds.

In the proof we will use the following result, which is an immediate consequence of Corollary 2.1 in Csörgő, Horváth and Shao (1991) (cf. Theorem 1.3.A).

Lemma 5.3.A. *We assume that $0 < p < \infty$ and q is positive on $(0, 1)$, i.e., $\inf_{\delta \leq t \leq 1-\delta} q(t) > 0$ for all $\delta \in (0, 1/2)$. Then the statements*

$$\int_0^1 (t(1-t))^{p/2} / q(t) dt < \infty$$

and

$$\int_0^1 |B(t)|^p / q(t) dt < \infty \quad a.s.$$

are equivalent.

Proof of Theorem 5.3.1. (a) As in the proof of Theorem 5.2.1, we assume without loss of generality that $EX_1 = 0$ and $\sigma = 1$. We introduce

$$S^{(1)}((n+1)t) = \sum_{i=1}^{[(n+1)t]} X_i, \quad 0 \leq t \leq 1/2,$$

and

$$S^{(2)}((n+1)t) = \sum_{i=[(n+1)t]+1}^n X_i, \quad 1/2 \leq t < 1.$$

Then

(5.3.2)

$$S((n+1)t) = \begin{cases} S^{(1)}((n+1)t), & 0 \leq t \leq 1/2 \\ S^{(1)}\left(\frac{n+1}{2}\right) + S^{(2)}\left(\frac{n+1}{2}\right) - S^{(2)}((n+1)t), & 1/2 < t < 1. \end{cases}$$

Using Theorem 1.4.1, we can construct two independent Wiener processes

$$\{W^{(1)}(t), 0 \leq t < \infty\} \text{ and } \{W^{(2)}(t), 0 \leq t < \infty\}$$

such that with $q \in \mathcal{Q}^*$

(5.3.3)

$$\int_0^{1/2} \left| n^{-1/2} S^{(1)}((n+1)t) - n^{-1/2} W^{(1)}((n+1)t) \right|^p / q(t) dt = o_P(1)$$

if and only if $\int_0^1 t^{p/2} / q(t) dt < \infty$,

and

(5.3.4)

$$\int_{1/2}^1 \left| n^{-1/2} S^{(2)}((n+1)t) - n^{-1/2} W^{(2)}((n+1)(1-t)) \right|^p / q(t) dt$$

$$= \int_0^{1/2} \left| n^{-1/2} S^{(2)}((n+1)(1-t)) - n^{-1/2} W^{(2)}((n+1)t) \right|^p / q(1-t) dt = o_P(1)$$

if and only if $\int_0^1 t^{p/2} / q(1-t) dt < \infty$.

In order to see that (5.3.4) is a consequence of Theorem 1.4.1, we note again that

$$S^{(2)}((n+1)(1-t)) = \begin{cases} 0 & 0 < t < 1/(n+1) \\ \sum_{i=1}^{[(n+1)t]} X_{n-i+1} & 1/(n+1) \leq t \leq 1/2. \end{cases}$$

We define a Wiener process W by

(5.3.5)

$$\{W(nt), 0 \leq t \leq 1\} = \begin{cases} W^{(1)}(nt), & 0 \leq t \leq 1/2 \\ W^{(1)}(\frac{n}{2}) + W^{(2)}(\frac{n}{2}) - W^{(2)}(n-nt), & 1/2 < t \leq 1. \end{cases}$$

Then we have

$$\int_0^{1/2} \left| n^{-1/2} \left(S((n+1)t) - [(n+1)t]S(n)/n - (W((n+1)t) - tW(n+1)) \right) \right|^p / q(t) dt$$

$$\leq 2^p \int_0^{1/2} \left| n^{-1/2} S^{(1)}((n+1)t) - n^{-1/2} W^{(1)}((n+1)t) \right|^p / q(t) dt$$

$$+ 2^p \int_0^{1/2} \left| n^{-3/2} [(n+1)t]S(n) - n^{-1/2} tW(n+1) \right|^p / q(t) dt$$

$$= 2^p J_1^{(1)}(n) + 2^p J_2^{(1)}(n).$$

By definition of $S^{(1)}$ and $S^{(2)}$ we have for each $n \geq 1$

$$(5.3.6) \quad S^{(1)}\left(\frac{n+1}{2}\right) + S^{(2)}\left(\frac{n+1}{2}\right) = S(n).$$

and by (5.3.5) with $t = 1$

$$(5.3.7) \quad W^{(1)}\left(\frac{n+1}{2}\right) + W^{(2)}\left(\frac{n+1}{2}\right) = W(n+1).$$

Hence for $t \in (1/2, 1)$ we obtain

$$\begin{aligned} & \int_{1/2}^1 \left| n^{-1/2} \left(S((n+1)t) - [(n+1)t]S(n)/n - (W((n+1)t) - tW(n+1)) \right) \right|^p / q(t) dt \\ & \leq 2^p \int_{1/2}^1 \left| n^{-1/2} S^{(2)}((n+1)t) - n^{-1/2} W^{(2)}((n+1)(1-t)) \right|^p / q(t) dt \\ & + 2^p \int_{1/2}^1 \left| n^{-1/2} \left(S^{(1)}\left(\frac{n+1}{2}\right) + S^{(2)}\left(\frac{n+1}{2}\right) - [(n+1)t]S(n)/n \right) \right. \\ & \quad \left. - n^{-1/2} \left(W^{(1)}\left(\frac{n+1}{2}\right) + W^{(2)}\left(\frac{n+1}{2}\right) - tW(n+1) \right) \right|^p / q(t) dt \\ & = 2^p \int_{1/2}^1 \left| n^{-1/2} S^{(2)}((n+1)t) - n^{-1/2} W^{(2)}((n+1)(1-t)) \right|^p / q(t) dt \\ & + 2^p \int_{1/2}^1 \left| n^{-1/2} (1 - [(n+1)t]/n) S(n) - n^{-1/2} (1-t) W(n+1) \right|^p / q(t) dt \\ & = 2^p J_1^{(2)}(n) + 2^p J_2^{(2)}(n). \end{aligned}$$

Let $q \in \mathcal{Q}^*$. By (5.3.3) and (5.3.4) we conclude that if

$$\int_0^1 (t(1-t))^{p/2} / q(t) < \infty$$

then

$$J_1^{(1)}(n) = o_P(1) = J_1^{(2)}(n).$$

Next we show that both terms $J_2^{(1)}(n)$ and $J_2^{(2)}(n)$ are $o_P(1)$ as $n \rightarrow \infty$, for any function q such that (5.3.1) holds.

We note (cf. (5.2.17)) that we have

$$n^{-1/2}|S(n) - W(n+1)| = o_P(1).$$

Let $q \in \mathcal{Q}^*$ be such that (5.3.1) holds. Using (5.2.7) and (5.2.8) of Lemma 5.2.1, and the positivity of q , we have for any $0 < \delta < 1/2$

$$\begin{aligned} J_2^{(1)} &= \int_0^{1/2} |n^{-3/2}[(n+1)t]S(n) - n^{-1/2}tW(n+1)|^p/q(t)dt \\ &\leq \left(\sup_{0 < t < \delta} |n^{-3/2}[(n+1)t]S(n) - n^{-1/2}tW(n+1)|/t^{1/2} \right)^p \int_0^\delta t^{p/2}/q(t)dt \\ &\quad + \int_\delta^{1/2} |n^{-3/2}[(n+1)t]S(n) - n^{-1/2}tW(n+1)|^p/q(t)dt \\ &= O_P(1) \int_0^\delta t^{p/2}/q(t)dt + o_P(1). \end{aligned}$$

Similarly, using (5.2.9) and (5.2.10), we have

$$\begin{aligned} J_2^{(2)} &= \int_{1/2}^1 |n^{-1/2}(1 - [(n+1)t]/n)S(n) - n^{-1/2}(1-t)W(n+1)|^p/q(t)dt \\ &\leq \int_{1/2}^{1-\delta} |n^{-1/2}(1 - [(n+1)t]/n)S(n) - n^{-1/2}(1-t)W(n+1)|^p/q(t)dt \\ &\quad + \left(\sup_{1-\delta < t < 1} |n^{-1/2}(1 - [(n+1)t]/n)S(n) - n^{-1/2}(1-t)W(n+1)/(1-t)^{1/2}| \right)^p \\ &\quad \times \int_{1-\delta}^1 (1-t)^{p/2}/q(t)dt \\ &= o_P(1) + O_P(1) \int_{1-\delta}^1 (1-t)^{p/2}/q(t)dt, \end{aligned}$$

for any $0 < \delta < 1/2$. Since $\delta > 0$ can be taken arbitrarily small, using (5.3.1) we get

$$J_2^{(1)} = o_P(1) \quad \text{and} \quad J_2^{(2)} = o_P(1),$$

and the proof of the "if" part of (a) is now complete.

To prove the converse part, we assume that

$$\int_0^1 |Z_n(t) - B_n(t)|^p/q(t)dt = o_P(1).$$

Hence we have

$$o_P(1) = \int_0^{1/(n+1)} |B_n(t)|^p/q(t)dt \stackrel{D}{=} \int_0^{1/(n+1)} |B(t)|^p/q(t)dt$$

and

$$o_P(1) = \int_{n/(n+1)}^1 |B_n(t)|^p/q(t)dt \stackrel{D}{=} \int_{n/(n+1)}^1 |B(t)|^p/q(t)dt,$$

where $\{B(t); 0 \leq t \leq 1\}$ is a Brownian bridge. Thus for n large enough there exist positive constants δ and K such that

$$P\left\{\int_0^\delta |B(t)|^p/q(t)dt > K\right\} \leq 1/2$$

and

$$P\left\{\int_{1-\delta}^1 |B(t)|^p/q(t)dt > K\right\} \leq 1/2,$$

and hence by the 0-1 law for a Brownian bridge, we have

$$\int_0^\delta |B(t)|^p/q(t)dt < \infty \quad a.s.$$

as well as

$$\int_{1-\delta}^1 |B(t)|^p/q(t)dt < \infty \quad a.s.$$

Since

$$\int_0^1 |B(t)|^p dt < \infty \quad a.s.$$

and $q(t)$ is positive, i.e. $\inf_{\delta \leq t \leq 1-\delta} q(t) > 0$ for any $\delta \in (0, 1/2)$, we have also

$$\int_\delta^{1-\delta} |B(t)|^p/q(t)dt < \infty \quad a.s.$$

Consequently

$$\int_0^1 |B(t)|^p/q(t)dt < \infty \quad a.s.$$

Now we use Lemma 5.3.A to conclude the result.

(b) First we assume that (5.3.1) holds, i.e.

$$\int_0^1 (t(1-t))^{p/2}/q(t)dt < \infty.$$

Using for example Theorem 5.2.1 (a) with $q \equiv 1$, by positivity of $q \in \mathcal{Q}^*$, we have

$$\sup_{\delta < t < 1-\delta} |Z_n(t) - B_n(t)|/q(t) = o_P(1)$$

for any $\delta \in (0, 1/2)$, and hence also

$$(5.3.8) \quad \int_{\delta}^{1-\delta} |Z_n(t)|^p/q(t)dt \xrightarrow{P} \int_{\delta}^{1-\delta} |B(t)|^p/q(t)dt,$$

where $\{B(t); 0 \leq t \leq 1\}$ is a Brownian bridge. By Lemma 5.3.A

$$(5.3.9) \quad \lim_{\delta \rightarrow 0} \int_0^{\delta} |B(t)|^p/q(t)dt = 0 \quad a.s.$$

and

$$(5.3.10) \quad \lim_{\delta \rightarrow 0} \int_{1-\delta}^1 |B(t)|^p/q(t)dt = 0 \quad a.s.$$

Since

$$\int_0^{\delta} |Z_n(t)|^p/q(t)dt \leq 2^p \int_0^{\delta} |B_n(t)|^p/q(t)dt + 2^p \int_0^{\delta} |Z_n(t) - B_n(t)|^p/q(t)dt,$$

by (5.3.9) and part (a) of this theorem we have for all $\epsilon > 0$

$$(5.3.11) \quad \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} P \left\{ \int_0^{\delta} |Z_n(t)|^p/q(t)dt > \epsilon \right\} = 0.$$

Similarly, using (5.3.10) instead of (5.3.9), we have also

$$(5.3.12) \quad \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} P \left\{ \int_{1-\delta}^1 |Z_n(t)|^p/q(t)dt > \epsilon \right\} = 0$$

for all $\epsilon > 0$. Combining (5.3.8) - (5.3.12) yield

$$\int_0^1 |Z_n(t)|^p / q(t) dt \xrightarrow{p} \int_0^1 |B(t)|^p / q(t) dt.$$

Conversely, if the above convergence holds, it means that the limit is almost surely finite, i.e.

$$\int_0^1 |B(t)|^p / q(t) dt < \infty \quad a.s. .$$

Hence Lemma 5.3.A implies

$$\int_0^1 (t(1-t))^{p/2} / q(t) dt < \infty.$$

5.4. Contiguous alternatives

Having described the weighted asymptotics of $Z_n(t)$ under the null hypothesis

$$H_0: X_i, \quad 1 \leq i \leq n, \quad \text{have the same distribution } F,$$

we now study the $Z_n(t)$ process with weight functions $q(t)$ under the sequence of alternative hypotheses as introduced in Section 2.3, namely

$$H_1: X_i, \quad 1 \leq i \leq n, \quad \text{have the respective distribution functions } F_{in},$$

where we assume that all F_{in} are absolutely continuous with respect to the distribution function F and

$$(5.4.1) \quad \left[\frac{dF_{in}}{dF}(F^{-1}(u)) \right]^{1/2} = 1 + \frac{1}{2\sqrt{n}} g_n(t, u), \quad \frac{i-1}{n} < t \leq \frac{i}{n},$$

where $F^{-1}(u) = \inf\{x : (F(x) \geq u)\}$, $0 < u \leq 1$, $F^{-1}(0) = F^{-1}(0+)$ (cf. (2.3.1)).

We assume also that there exists a function $g \in L^2[0, 1]^2$, such that

$$(5.4.2) \quad \int_0^1 g(t, u) du = 0 \quad \text{for almost all } t \in [0, 1]$$

and

$$(5.4.3) \quad \int_0^1 \int_0^1 \left[g_n(t, u) - g\left(\frac{[nt]}{n}, u\right) \right]^2 du dt \rightarrow 0, \quad n \rightarrow \infty.$$

The sequence of direct products $F_{1n} \times \dots \times F_{nn}$, $n = 1, 2, \dots$ is contiguous to the sequence $F \times \dots \times F$ (cf. Chapter 2). In particular, for the changepoint alternatives H_1^η we assume that there exists $\eta \in (0, 1)$ such that

$$(5.4.4) \quad g(t, u) = 1\{t \geq \eta\}g(u)$$

for some square integrable function g .

We recall that $\mathcal{Q}^*$ is the class of functions $q : (0, 1) \rightarrow (0, \infty)$ which are nondecreasing near zero and nonincreasing near one, and such that $\inf_{\delta \leq t \leq 1-\delta} q(t) > 0$ for all $\delta \in (0, 1/2)$. As before, let

$$I^*(q, c) = \int_0^1 (t(1-t))^{-1} \exp(-c(t(1-t))^{-1} q^2(t)) dt$$

and

$$E^*(q, c) = \int_0^1 (t(1-t))^{-3/2} q(t) \exp(-c(t(1-t))^{-1} q^2(t)) dt.$$

Theorem 5.4.1. *Let $X_1, X_2, \dots$ be independent random variables. We assume that under H_0 $\int x dF = 0$ and $\int x^2 dF = 1$. Then, as $n \rightarrow \infty$, under H_1 the following hold.*

(a) *Let $q \in \mathcal{Q}^*$. If $I^*(q, c) < \infty$ for all $c > 0$, then*

$$Z_n(t)/q(t) \xrightarrow{D} (B(t) + b(t))/q(t) \text{ in } D(0, 1).$$

where

$$b(t) = \int_0^t \int_0^1 g(s, u) F^{-1}(u) du ds - t \int_0^1 \int_0^1 g(s, u) F^{-1}(u) du ds,$$

and $\{B(t); 0 \leq t \leq 1\}$ is a Brownian bridge.

(b) Let $q : (0, 1) \rightarrow (0, \infty)$ be such that (5.2.2) holds. Then

$$Z_n(t)/q(t) \xrightarrow{\mathcal{D}} (B(t) + b(t))/q(t) \quad \text{in } D[\delta, 1 - \delta]$$

for any $\delta \in (0, 1/2)$.

Remark 5.4.1. (a) According to Theorem 5.1.C, in part (a) of Theorem 5.4.1 we can equivalently assume $E^*(q, c) < \infty$ for all $c > 0$ and (5.1.2), instead of $I^*(q, c) < \infty$ for all $c > 0$.

Proof. As in Chapter 2, we define the centered log-likelihood ratio process $L_n(\cdot)$ by

$$L_n(t) = \sum_{i=1}^{[nt]} \left\{ \log \frac{dF_{in}}{dF}(F^{-1}(\bar{U}_i)) - E \log \frac{dF_{in}}{dF}(F^{-1}(\bar{U}_i)) \right\},$$

(cf. (2.3.4)).

Let (B, L) be the two-dimensional Gaussian process, where B is a Brownian bridge, L is a Gaussian process with mean zero and covariance function $Q_L(t_1 \wedge t_2)$, where

$$Q_L(t) = \int_0^t \int_0^1 g^2(s, u) du ds,$$

and the covariance function of B and L is

$$Q_{B,L}(t_1, t_2) = EB(t_1)L(t_2) = C(t_1 \wedge t_2) - t_1 C(t_2)$$

with

$$C(t) = \int_0^t \int_0^1 g(s, u) F^{-1}(u) du ds.$$

It follows from Lemma 2.4.1 that assuming H_0

$$(Z_n(t), L_n(t)) \xrightarrow{\mathcal{D}} (B(t), L(t)) \quad \text{in } D^2[0, 1].$$

Let $q \in \mathcal{Q}^*$ and $I^*(q, c) < \infty$ for all $c > 0$. Using Corollary 5.2.1 (a) and the fact that

$$E \frac{Z_n(t_1)}{q(t_1)} L_n(t_2) = \frac{1}{q(t_1)} E Z_n(t_1) L_n(t_2) \rightarrow \frac{1}{q(t_1)} Q_{B,L}(t_1, t_2),$$

we conclude that

$$(Z_n(t)/q(t), L_n(t)) \xrightarrow{D} (B(t)/q(t), L(t)) \text{ in } D^2[0, 1],$$

and note also that, due to $g \in L^2[0, 1]^2$ and $\int x^2 dF = 1$, the covariance function of B/q and L is finite. In order to see this, we consider

$$\begin{aligned} & \left| \int_0^t \int_0^1 g(s, u) F^{-1}(u) du ds \right| \\ & \leq \int_0^t \left(\int_0^1 g^2(s, u) du \right)^{1/2} \left(\int_0^1 (F^{-1}(u))^2 du \right)^{1/2} ds \\ & \leq \left(\int_0^t \int_0^1 g^2(s, u) du ds \right)^{1/2} \left(\int_0^t ds \right)^{1/2} \\ & = \left(\int_0^t \int_0^1 g^2(s, u) du ds \right)^{1/2} t^{1/2}. \end{aligned}$$

Similarly

$$\begin{aligned} & \left| \int_t^1 \int_0^1 g(s, u) F^{-1}(u) du ds \right| \\ & \leq \left(\int_t^1 \int_0^1 g^2(s, u) du ds \right)^{1/2} (1-t)^{1/2}. \end{aligned}$$

Since

$$\begin{aligned} b(t) &= \int_0^t \int_0^1 g(s, u) F^{-1}(u) du ds - t \int_0^1 \int_0^1 g(s, u) F^{-1}(u) du ds \\ &= (1-t) \int_0^t \int_0^1 g(s, u) F^{-1}(u) du ds - t \int_t^1 \int_0^1 g(s, u) F^{-1}(u) du ds, \end{aligned}$$

we have

$$\begin{aligned} \sup_{0 < t < 1} |b(t)|/q(t) &\leq \sup_{0 < t < 1} \left((1-t)t^{1/2} \left(\int_0^t \int_0^1 g^2(s, u) du ds \right)^{1/2} \right. \\ &\quad \left. - (1-t)^{1/2}t \left(\int_t^1 \int_0^1 g^2(s, u) du ds \right)^{1/2} \right) / q(t) \\ &\leq \sup_{0 < t < 1} \left((1-t)^{1/2} \left(\int_0^t \int_0^1 g^2(s, u) du ds \right)^{1/2} - t^{1/2} \left(\int_0^t \int_0^1 g^2(s, u) du ds \right)^{1/2} \right) \\ &\quad \times \sup_{0 < t < 1} (t(1-t))^{1/2} / q(t) < \infty. \end{aligned}$$

Now Le Cam's third lemma implies that under H_1

$$Z_n(t)/q(t) \xrightarrow{D} (B(t) + b(t))/q(t),$$

where $b(t) = \text{Cov}(B(t), L(1))$. Hence the result.

(b) Using part (a) with $q \equiv 1$, we have that under H_1

$$Z_n(t) \xrightarrow{D} B(t) + b(t) \quad \text{in } D[0, 1].$$

From here on the proof is the same as that of part (b) of Theorem 2.5.1.

Corollary 5.4.1. *Let $X_1, X_2, \dots$ be independent random variables. We assume that under H_0 $\int x dF = 0$ and $\int x^2 dF = 1$. Then, as $n \rightarrow \infty$, under H_1^η the following hold.*

(a) *Let $q \in \mathcal{Q}^*$. If $I^*(q, c) < \infty$ for all $c > 0$, then*

$$(5.4.5) \quad Z_n(t)/q(t) \xrightarrow{D} (B(t) + b^\eta(t))/q(t)$$

in $D(0, 1)$, where

$$b^\eta(t) = (-\eta(1-t)\mathbf{1}\{t \geq \eta\} - (1-\eta)t\mathbf{1}\{t < \eta\}) \int_0^1 g(u)F^{-1}(u)du.$$

(b) *Let $q(t) : (0, 1) \rightarrow (0, \infty)$ be such that (5.2.2) holds. Then (5.4.5) holds in $D[\delta, 1-\delta]$ for any $\delta \in (0, 1/2)$.*

Proof. We have

$$\begin{aligned} & \int_0^t \int_0^1 \mathbf{1}\{s \geq \eta\} g(u)F^{-1}(u)du ds - t \int_0^1 \int_0^1 \mathbf{1}\{s \geq \eta\} g(u)F^{-1}(u)du ds \\ &= ((t-\eta)\mathbf{1}\{t \geq \eta\} - t(1-\eta)) \int_0^1 g(u)F^{-1}(u)du \\ &= (-\eta(1-t)\mathbf{1}\{t \geq \eta\} - (1-\eta)t\mathbf{1}\{t < \eta\}) \int_0^1 g(u)F^{-1}(u)du, \end{aligned}$$

and hence also the results (a) and (b) by those of Theorem 5.4.1 respectively.

CHAPTER 6

U-statistic Type Processes

6.0. Introduction

For the problem of detecting change in distribution, Csörgö and Horváth (1988a) study tests based on processes of U -statistics which are generalizations of Wilcoxon-Mann-Whitney type statistics. They consider

$$(6.0.1) \quad Z_k = \sum_{1 \leq i \leq k} \sum_{k+1 \leq j \leq n} h(X_i, X_j), \quad 1 \leq k < n,$$

where $h(x, y)$ is a symmetric or an antisymmetric kernel. In order to investigate Z_k which itself is not a U -statistic, Csörgö and Horváth (1988a) express it as a linear combination of U -statistics. Here, following Serfling (1980, Chapter 5), we give a short summary of properties of U -statistics.

The class of U -statistics was introduced in a fundamental paper by Hoeffding (1948). This class of statistics arises as a generalization of the sample mean, that is, as a generalization of the notion of forming an average.

Let $X_1, X_2, \dots$ be independent observations on a distribution F . Consider a "parametric function" $\theta = \theta(F)$ for which there is an unbiased estimator. That is, $\theta(F)$ may be represented as

$$\theta(F) = E_F\{h(X_1, \dots, X_m)\} = \int \dots \int h(x_1, \dots, x_m) dF(x_1) \dots dF(x_m),$$

for some function $h = h(x_1, \dots, x_m)$, called a "kernel". For any kernel, the corresponding U - statistic for estimation of θ on the basis of a sample $X_1, \dots, X_n$ of size $n \geq m$ is obtained by averaging the kernel h symmetrically over the observations:

$$U_n = U(X_1, \dots, X_n) = \frac{1}{\binom{n}{m}} \sum_c h(X_{i_1}, \dots, X_{i_m}),$$

where $\sum_r$ denotes summation over the $\binom{n}{m}$ combinations of m distinct elements $\{i_1, \dots, i_m\}$ from $\{1, \dots, n\}$.

Examples: (i) *The sample mean.* $\theta(F) = \mu(F) = \int x dF(x)$. For the kernel $h(x) = x$, the corresponding U -statistic is

$$U(X_1, \dots, X_n) = \frac{1}{n} \sum_{i=1}^n X_i.$$

(ii) *The sample variance.* $\theta(F) = \sigma^2(F) = \int (x - \mu)^2 dF(x)$. For the kernel $h(x_1, x_2) = \frac{1}{2}(x_1 - x_2)^2$, the corresponding U -statistic is

$$\begin{aligned} U(X_1, \dots, X_n) &= \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} h(X_i, X_j) \\ &= \frac{1}{n-1} \left(\sum_{i=1}^n X_i^2 - n\bar{X}^2 \right). \end{aligned}$$

(iii) *"Gini's mean difference".* $\theta(F) = E_F|X_1 - X_2|$, a measure of concentration. For the kernel $h(x_1, x_2) = |x_1 - x_2|$, the corresponding U -statistic is

$$U(X_1, \dots, X_n) = \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} |X_i - X_j|.$$

(iv) *The Wilcoxon one-sample statistic.* For estimation of $\theta(F) = P_F(X_1 + X_2 \leq 0)$, the kernel is given by $h(x_1, x_2) = 1\{x_1 + x_2 \leq 0\}$ and the corresponding U -statistic is

$$U(X_1, \dots, X_n) = \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} 1\{X_i + X_j \leq 0\}.$$

An appealing feature of a U -statistic is its simple structure as a sum of identically distributed random variables. However, except in the case of a kernel of dimension $m = 1$, the summands are not all independent. Hence a direct application of the theory for sums of independent random variables is not possible. An important tool in deriving the asymptotic theory of U -statistics is the "projection"

of a U -statistic on the basic observations of the sample. Using such device, a U -statistic can be approximated, within a sufficient degree of accuracy, by a sum of i.i.d.r.v.'s. Assume $E_F|h| < \infty$. The projection $\hat{U}_n$ of the U -statistic U_n is defined by

$$\hat{U}_n = \sum_{i=1}^n E_F\{U_n|X_i\} - (n-1)\theta,$$

which is a sum of i.i.d.r.v.'s. It can be written as

$$\hat{U}_n - \theta = \frac{1}{n} \sum_{i=1}^n \tilde{h}_1(X_i),$$

where $\tilde{h}_1 = h_1 - E_F\{h(X_1, \dots, X_m)\}$ and $h_1(x_1) = E_F\{h(x_1, X_2, \dots, X_m)\}$ since $E_F\{h(x_1, X_2, \dots, X_m)\} = E_F\{h(X_1, \dots, X_m)|X_1 = x_1\}$. Assuming $E_F h^2 < \infty$, we have $E_F(U_n - \hat{U}_n)^2 = O(n^{-2})$, $n \rightarrow \infty$. This was established by Hoeffding (1948) and applied to obtain the central limit theorem for U -statistics, while Hoeffding (1961) proved the classical strong law of large numbers, namely that assuming $E_F|h| < \infty$ we have $U_n \rightarrow \theta$ a.s. For more details and further results along these lines we refer to Serfling (1980, Chapter 5).

A simple example of a symmetric kernel in (6.0.1) is $h(x, y) = x + y$, which gives Z_k equal to $(n-k) \sum_{i=1}^k X_i + k \sum_{i=k+1}^n X_i$, or, equivalently, to

$$(6.0.2) \quad n \left(\left(1 - \frac{k}{n}\right) S_k + \frac{k}{n} (S_n - S_k) \right), \quad 1 \leq k < n,$$

where $S_k = \sum_{i=1}^k X_i$. Clearly, if $X_1, X_2, \dots$ are i.i.d.r.v.'s with $EX_1 = 0$ and $EX_1^2 = 1$, then Donsker's invariance principle yields

$$n^{-1/2} \left(\left(1 - \frac{[nt]}{n}\right) S_{[nt]} + \frac{[nt]}{n} (S_n - S_{[nt]}) \right) \xrightarrow{D} (1-t)W(t) + t(W(1) - W(t)),$$

where $\{W(t), 0 \leq t \leq 1\}$ is a standard Wiener process.

On the other hand, if we take $h(x, y) = x - y$, which is an example of an antisymmetric kernel, then Z_k becomes equal to $k(n-k) \left(\frac{1}{k} \sum_{i=1}^k X_i - \frac{1}{n-k} \sum_{i=k+1}^n X_i \right)$ or, equivalently, to

$$(6.0.3) \quad n \left(S_k - \frac{k}{n} S_n \right), \quad 1 \leq k < n,$$

and hence, on assuming $EX_1^2 = 1$, by the same argument as above, we have

$$n^{-1/2} \left(S_{[nt]} - \frac{[nt]}{n} S_n \right) \xrightarrow{D} W(t) - tW(1),$$

where $\{W(t), 0 \leq t \leq 1\}$ is a standard Wiener process. Thus, with $h(x, y) = x - y$, the limiting process is a Brownian bridge, while in the case of $h(x, y) = x + y$, the limiting process is a different Gaussian process, though it has the variance of a Brownian bridge.

General results for Z_k based on i.i.d.r.v.'s are due to Csörgő and Horváth (1988a). Assuming first that $h(x, y)$ is a symmetric kernel, they obtain sup-metric approximations of the properly normalized sequence of $Z_{[nt]}$ processes by a sequence of Gaussian processes $\{\Gamma_n(t), 0 \leq t \leq 1\}$, where for each $n \geq 1$

$$\{\Gamma_n(t), 0 \leq t \leq 1\} \xrightarrow{D} \{(1-t)W(t) + t(W(1) - W(t)), 0 \leq t \leq 1\}$$

(cf. Theorem 6.1.A). They prove also that in the case of $h(x, y)$ being an antisymmetric kernel, the sequence of processes $Z_{[nt]}$ (properly normalized) is to be approximated by a sequence of Brownian bridges $B_n(t)$ (cf. Theorem 6.2.A). As discussed above, in the case of an antisymmetric kernel, the asymptotic results for the $Z_{[nt]}$ processes coincide with those for the already considered tied down partial sum processes $Z_n(t)$. Even though results for the particular case follow from the more general one, we preferred to prove all our results directly for the

$Z_n(t)$ processes in Chapter 5, because they actually form the bases of our proofs when studying the more general $\mathcal{Z}_{[n,t]}$ processes with any antisymmetric kernel.

Another example of antisymmetric kernel, which has gained special attention in the literature, is $h(x, y) = \text{sign}(x - y)$. Pettitt (1979) considers statistics based on

$$U_{k,n} = \sum_{i=1}^k \sum_{j=k+1}^n \text{sign}(X_i - X_j),$$

where the X_i are continuously distributed random variables (cumulative rank tests).

In order to make their statistical tests based on $\mathcal{Z}_k$ more sensitive on the tails, Csörgő and Horváth (1988a) introduce weight functions. They study the asymptotic behaviour of $\mathcal{Z}_k$ in $\|\cdot/q\|$ metrics under the null hypothesis H_0 assuming $E|h(X_1, X_2)|^\nu < \infty$ for some $\nu > 2$. Here we improve on these results by assuming two moments only. We study also L_p -functionals of $\mathcal{Z}_k$ processes and consider such processes under the class of contiguous alternatives H_1 (cf. also Szyszkowicz, 1991a, d, 1992a,b,c).

6.1. Symmetric kernel

Let $X_1, X_2, \dots$ be independent identically distributed random variables. We assume that h is a symmetric function, i.e., $h(x, y) = h(y, x)$,

$$(6.1.1) \quad Eh^2(X_1, X_2) < \infty$$

and let $Eh(X_1, X_2) = \theta$, $\tilde{h}(t) = E\{h(X_1, t) - \theta\}$. Condition (6.1.1) implies that $E\tilde{h}^2(X_1) < \infty$ and we assume

$$(6.1.2) \quad 0 < \sigma^2 = E\tilde{h}^2(X_1).$$

We consider the following process

$$\mathcal{U}_k = Z_k - k(n-k)\theta, \quad 1 \leq k < n,$$

introduced by Csörgő and Horváth (1988a), which they expressed as a linear combination of $\mathcal{U}^*$ -statistics, namely

$$(6.1.3) \quad \mathcal{U}_k = \mathcal{U}_n^{(3)} - \{\mathcal{U}_k^{(1)} + \mathcal{U}_k^{(2)}\},$$

where

$$\begin{aligned} \mathcal{U}_k^{(1)} &= \sum_{1 \leq i < j \leq k} h(X_i, X_j) - \binom{k}{2} \theta, \\ \mathcal{U}_k^{(2)} &= \sum_{k+1 \leq i < j \leq n} h(X_i, X_j) - \binom{n-k}{2} \theta, \end{aligned}$$

and

$$\mathcal{U}_n^{(3)} = \sum_{1 \leq i < j \leq n} h(X_i, X_j) - \binom{n}{2} \theta.$$

We define the Gaussian process Γ by

$$(6.1.4) \quad \Gamma(t) = (1-t)W(t) + t\{W(1) - W(t)\}, \quad 0 \leq t \leq 1,$$

where $\{W(t), 0 \leq t < \infty\}$ is a Wiener process. We note that $E\Gamma(t) = 0$ for all $t \in [0, 1]$ and the covariance function of $\Gamma(t)$ is

$$\text{Cov}(\Gamma(t_1), \Gamma(t_2)) = (1-2t_1)(1-2t_2)(t_1 \wedge t_2) + (3-2t_1-2t_2)t_1t_2,$$

and hence $E\Gamma^2(t) = t(1-t)$, i.e., its variance is like that of a Brownian bridge.

Csörgő and Horváth (1988a) obtained the following approximation of U_k process.

Theorem 6.1.A. *Let $X_1, X_2, \dots$ be i.i.d.r.v.'s and $h(x, y) = h(y, x)$. We assume that (6.1.1) and (6.1.2) hold, i.e.*

$$Eh^2(X_1, X_2) < \infty \text{ and } 0 < \sigma^2 = Eh^2(X_1).$$

Then we can define a sequence of Gaussian processes $\{\Gamma_n(t), 0 \leq t \leq 1\}$ such that, as $n \rightarrow \infty$

$$\sup_{0 \leq t \leq 1} \left| \frac{n^{-3/2}}{\sigma} \mathcal{U}_{[(n+1)t]} - \Gamma_n(t) \right| = o_P(1),$$

where for each $n \geq 1$

$$\{\Gamma_n(t); 0 \leq t \leq 1\} \stackrel{D}{=} \{\Gamma(t), 0 \leq t \leq 1\}.$$

Weighted approximation results for $\mathcal{U}_k$ process in supremum norm for the optimal class of weight functions were proved by Csörgő and Horváth (1988a), assuming $E|h(X_1, X_2)|^\nu < \infty$ for some $\nu > 2$. For the same class of weight functions we prove weighted approximations under the above initial assumption of two moments only.

As in Chapter 5, let $\mathcal{Q}^*$ be the class of positive functions on $(0, 1)$ i.e.

$$(6.1.5) \quad \inf_{\delta \leq t \leq 1-\delta} q(t) > 0 \text{ for all } \delta \in (0, 1/2),$$

which are nondecreasing near zero and nonincreasing near one. We recall also the integrals

$$E^*(q, c) = \int_0^1 (t(1-t))^{-3/2} q(t) \exp(-c(t(1-t))^{-1} q^2(t)) dt$$

and

$$I^*(q, c) = \int_0^1 (t(1-t))^{-1} \exp(-cq^2(t)/(t(1-t))) dt, \quad c > 0.$$

Remark 6.1.1. It follows from the definition of the process $\{\Gamma(t); 0 \leq t \leq 1\}$ (cf. (6.1.4)) that $q \in \mathcal{Q}^*$ is a local function of Γ if and only if it is a local function of a Brownian bridge (cf. Section 5.1). Similarly, $q \in \mathcal{Q}^*$ is a Chibisov-O'Reilly function of Γ if and only if it is a Chibisov-O'Reilly function of a Brownian bridge.

Namely, if $q \in Q^*$ is a Chibisov-O'Reilly function of a Brownian bridge, i.e., almost surely

$$\lim_{t \downarrow 0} |W(t)|/q(t) = 0$$

and

$$\lim_{t \downarrow 0} |W(t)|/q(1-t) = 0 \text{ or, equivalently, } \lim_{t \uparrow 1} |W(1) - W(t)|/q(t) = 0,$$

then by Theorem 5.1.C we have $I^*(q, c) < \infty$ for all $c > 0$, which implies

$$\lim_{t \downarrow 0} t^{1/2}/q(t) = \lim_{t \uparrow 1} (1-t)^{1/2}/q(t) = 0.$$

Consequently, since $\Gamma(t) = (1-t)W(t) + t(W(1) - W(t))$, we have also

$$\lim_{t \downarrow 0} |\Gamma(t)|/q(t) = 0 = \lim_{t \uparrow 1} |\Gamma(t)|/q(t),$$

almost surely. Conversely now, if we assume that $q \in Q^*$ is A Chibisov-O'Reilly function of Γ , i.e. that the latter statement holds true, then we must have

$$\lim_{t \downarrow 0} E\Gamma^2(t)/q^2(t) = \lim_{t \uparrow 1} E\Gamma^2(t)/q^2(t) = 0,$$

i.e.,

$$\lim_{t \downarrow 0} (t(1-t))/q^2(t) = \lim_{t \uparrow 1} (t(1-t))/q^2(t) = 0,$$

and hence also

$$\lim_{t \downarrow 0} t/q(t) = \lim_{t \uparrow 1} (1-t)/q(t) = 0.$$

This, in turn, due to the definition of Γ , implies that we have

$$\lim_{t \downarrow 0} |W(t)|/q(t) = 0.$$

as well as

$$\lim_{t \uparrow 1} |W(1) - W(t)|/q(t) = 0.$$

almost surely, i.e. q is a Chibisov-O'Reilly function of a Brownian bridge. The proof for local functions of Γ is similar, only we have to use Theorem 5.1.B instead of Theorem 5.1.C.

Theorem 6.1.1. *Let $X_1, X_2, \dots$ be i.i.d.r.v.'s and $h(x, y) = h(y, x)$. We assume that (6.1.1) and (6.1.2) are satisfied, i.e.*

$$Eh^2(X_1, X_2) < \infty \quad \text{and} \quad 0 < \sigma^2 = E\tilde{h}^2(X_1).$$

Then we can define a sequence of Gaussian processes $\{\Gamma_n(t), 0 \leq t \leq 1\}$ such that for each $n \geq 1$

$$\{\Gamma_n(t), 0 \leq t \leq 1\} \stackrel{D}{=} \{\Gamma(t), 0 \leq t \leq 1\}$$

and with $q \in \mathcal{Q}^$ we have*

$$\sup_{0 < t < 1} \left| \frac{n^{-3/2}}{\sigma} \mathcal{U}_{[(n+1)t]} - \Gamma_n(t) \right| / q(t) = o_P(1)$$

if and only if $I^(q, c) < \infty$ for all $c > 0$, and*

$$\sup_{0 < t < 1} \left| \frac{n^{-3/2}}{\sigma} \mathcal{U}_{[(n+1)t]} - \Gamma_n(t) \right| / q(t) = O_P(1)$$

if and only if $I^(q, c) < \infty$ for some $c > 0$.*

Remark 6.1.2. According to Theorems 5.1.C and 5.1.B and using Remark 6.1.1, the results of Theorem 6.1.1 can be stated equivalently in terms of the integral $E^*(q, c)$.

Proof. We assume that $q \in \mathcal{Q}^*$ and is such that

$$(6.1.6) \quad \lim_{t \downarrow 0} q(t)/t^{1/2} = \lim_{t \uparrow 1} q(t)/(1-t)^{1/2} = \infty.$$

When proving Theorem 6.1.A, Csörgő and Horváth (1988a) use Theorem 1 of Hall (1979) and obtain

$$\max_{1 \leq k \leq n} \left| \mathcal{U}_k - \left\{ (n-k) \sum_{i=1}^k \tilde{h}(X_i) + k \left(\sum_{i=1}^n \tilde{h}(X_i) - \sum_{i=1}^k \tilde{h}(X_i) \right) \right\} \right| = O_P(n).$$

Hence

$$\sup_{0 < t < 1} \left| n^{-3/2} \mathcal{U}_{[(n+1)t]} - \bar{U}_n(t) \right| = o_P(1)$$

as well as

$$\begin{aligned} & \sup_{1/(n+1) \leq t \leq n/(n+1)} \left| n^{-3/2} \mathcal{U}_{[(n+1)t]} - \bar{U}_n(t) \right| / (t(1-t))^{1/2} \\ & \leq O_P(1) \sup_{1/(n+1) \leq t \leq n/(n+1)} 1/(nt(1-t))^{1/2} \\ & = O_P(1), \end{aligned}$$

with

$$\bar{U}_n(t) = n^{-1/2} \left\{ \left(1 - \frac{[(n+1)t]}{n} \right) \sum_{i=1}^{[(n+1)t]} \tilde{h}(X_i) + \frac{[(n+1)t]}{n} \left(\sum_{i=1}^n \tilde{h}(X_i) - \sum_{i=1}^{[(n+1)t]} \tilde{h}(X_i) \right) \right\}.$$

Consequently,

$$\sup_{\delta \leq t \leq 1-\delta} \left| n^{-3/2} \mathcal{U}_{[(n+1)t]} - \bar{U}_n(t) \right| / q(t) = o_P(1)$$

for any $\delta \in (0, 1/2)$, due to (6.5.1), and we have also

$$\begin{aligned} & \sup_{1/(n+1) \leq t < \delta} \left| n^{-3/2} \mathcal{U}_{[(n+1)t]} - \bar{U}_n(t) \right| / q(t) \\ & \leq O_P(1) \sup_{0 < t < \delta} t^{1/2} / q(t) \end{aligned}$$

and

$$\begin{aligned} & \sup_{1-\delta < t < n/(n+1)} \left| n^{-3/2} \mathcal{U}_{[(n+1)t]} - \bar{U}_n(t) \right| / q(t) \\ & \leq O_P(1) \sup_{1-\delta < t < 1} (1-t)^{1/2} / q(t) \end{aligned}$$

for any $\delta \in (0, 1/2)$. Taking δ arbitrarily small and using (6.1.6), we obtain

$$(6.1.7) \quad \sup_{0 < t < 1} \left| n^{-1/2} \mathcal{U}_{[(n+1)t]} - n^{-1/2} \left\{ \left(1 - \frac{[(n+1)t]}{n} \right) \sum_{i=1}^{[(n+1)t]} \dot{h}(X_i) + \frac{[(n+1)t]}{n} \left(\sum_{i=1}^n \dot{h}(X_i) - \sum_{i=[(n+1)t]+1}^n \dot{h}(X_i) \right) \right\} \right| / q(t) dt = o_P(1).$$

Since $I^*(q, c) < \infty$ for some $c > 0$ implies (6.1.6), by (6.1.7) it is enough to verify the statements of theorem for $\sigma^{-1} \mathcal{U}_n$ instead of $n^{-3/2} \sigma^{-1} \mathcal{U}_{[(n+1)t]}$.

We introduce

$$\begin{aligned} S_n &= \sigma^{-1} \sum_{i=1}^n \dot{h}(X_i), \\ S_{[(n+1)t]}^{(1)} &= \sigma^{-1} \sum_{i=1}^{[(n+1)t]} \dot{h}(X_i), \quad 0 \leq t \leq 1/2 \\ S_{[(n+1)t]}^{(2)} &= \sigma^{-1} \sum_{i=[(n+1)t]+1}^n \dot{h}(X_i), \quad 1/2 \leq t < 1. \end{aligned}$$

By Theorem 1.2.1 there exist two independent Wiener processes $\{W^{(1)}(t), 0 \leq t < \infty\}$ and $\{W^{(2)}(t), 0 \leq t < \infty\}$ such that with $q \in \mathcal{Q}^*$

$$(6.1.8) \quad \sup_{0 < t < 1/2} \left| n^{-1/2} \left(S_{[(n+1)t]}^{(1)} - W^{(1)}((n+1)t) \right) \right| / q(t) = \begin{cases} o_P(1) & \text{if and only if } I(q(t), c) < \infty \text{ for all } c > 0 \\ O_P(1) & \text{if and only if } I(q(t), c) < \infty \text{ for some } c > 0 \end{cases}$$

and

$$(6.1.9) \quad \sup_{0 < t < 1/2} \left| n^{-1/2} \left(S_{[(n+1)(1-t)]}^{(2)} - W^{(2)}((n+1)t) \right) \right| / q(1-t) = \begin{cases} o_P(1) & \text{if and only if } I(q(1-t), c) < \infty \text{ for all } c > 0 \\ O_P(1) & \text{if and only if } I(q(1-t), c) < \infty \text{ for some } c > 0. \end{cases}$$

We define a Wiener process W by

$$(6.1.10) \quad W(nt) = \begin{cases} W^{(1)}(nt), & 0 \leq t \leq 1/2 \\ W^{(1)}\left(\frac{n}{2}\right) + W^{(2)}\left(\frac{n}{2}\right) - W^{(2)}(n - nt), & 1/2 < t \leq 1. \end{cases}$$

Then we have

$$\begin{aligned}
& \sup_{0 < t \leq 1/2} \left| n^{-1/2} \sigma^{-1} \left\{ \left(1 - \frac{[(n+1)t]}{n} \right) \sum_{i=1}^{[(n+1)t]} \tilde{h}(X_i) + \frac{[(n+1)t]}{n} \left(\sum_{i=1}^n \tilde{h}(X_i) - \sum_{i=1}^{[(n+1)t]} \tilde{h}(X_i) \right) \right. \right. \\
& \quad \left. \left. - n^{-1/2} \left\{ (1-t)W((n+1)t) + t(W(n+1) - W((n+1)t)) \right\} \right\} \right| / q(t) \\
& \leq \sup_{0 < t \leq 1/2} n^{-1/2} \left| \left(1 - \frac{[(n+1)t]}{n} \right) S_{[(n+1)t]}^{(1)} - (1-t)W^{(1)}((n+1)t) \right| / q(t) \\
& + \sup_{0 < t \leq 1/2} n^{-1/2} \left| \frac{[(n+1)t]}{n} S_n - tW(n+1) \right| / q(t) \\
& + \sup_{0 < t \leq 1/2} n^{-1/2} \left| \frac{[(n+1)t]}{n} S_{[(n+1)t]}^{(1)} - tW^{(1)}((n+1)t) \right| / q(t) \\
& = I_1^{(1)}(n) + I_2^{(1)}(n) + I_3^{(1)}(n).
\end{aligned}$$

Since $q \in \mathcal{Q}^*$ is positive (cf. (6.1.5)), assuming that (6.1.6) holds, we have

$$(6.1.11) \quad \sup_{0 < t \leq 1/2} t/q(t) < \infty$$

and

$$(6.1.12) \quad \sup_{1/2 < t < 1} (1-t)/q(t) < \infty.$$

We have also

$$(6.1.13) \quad \sup_{0 < t < 1} \left| t - \frac{[(n+1)t]}{n} \right| / q(t) = o(1),$$

which is easy to see by Observation 5.2.1. Let $q \in \mathcal{Q}^*$ and we recall that $I^*(q, c) < \infty$ for some $c > 0$ implies (6.1.6). Using (6.1.13) we have

$$\begin{aligned}
I_1^{(1)}(n) &= \sup_{0 < t \leq 1/2} n^{-1/2} \left| \left(1 - \frac{[(n+1)t]}{n} \right) S_{[(n+1)t]}^{(1)} - (1-t)W^{(1)}((n+1)t) \right| / q(t) \\
&\leq \sup_{0 < t \leq 1/2} n^{-1/2} \left| S_{[(n+1)t]}^{(1)} - W^{(1)}((n+1)t) \right| / q(t) \sup_{0 < t \leq 1/2} (1-t) \\
&+ \sup_{0 \leq t \leq 1/2} \left| n^{-1/2} S_{[(n+1)t]}^{(1)} \right| \sup_{0 < t \leq 1/2} \left| t - \frac{[(n+1)t]}{n} \right| / q(t) \\
&= \sup_{0 < t \leq 1/2} n^{-1/2} \left| S_{[(n+1)t]}^{(1)} - W^{(1)}((n+1)t) \right| / q(t) \\
&+ O_P(1) o(1).
\end{aligned}$$

Hence, by (6.1.8) we conclude

$$(6.1.14) \quad I_1^{(1)}(n) = \begin{cases} o_P(1) & \text{if } I^*(q, c) < \infty \text{ for all } c > 0 \\ O_P(1) & \text{if } I^*(q, c) < \infty \text{ for some } c > 0. \end{cases}$$

Since for each $n \geq 1$

$$S_{[(n+1)/2]}^{(1)} + S_{[(n+1)/2]}^{(2)} = S_n$$

and

$$W^{(1)}\left(\frac{(n+1)}{2}\right) + W^{(2)}\left(\frac{(n+1)}{2}\right) = W(n+1),$$

using (6.1.8) and (6.1.9) with $q \equiv 1$, we obtain

$$(6.1.15) \quad n^{-1/2} |S_n - W(n+1)| = o_P(1).$$

Hence,

$$\begin{aligned} I_2^{(1)}(n) &= \sup_{0 < t \leq 1/2} n^{-1/2} \left| \frac{[(n+1)t]}{n} S_n - tW(n+1) \right| / q(t) \\ &\leq n^{-1/2} |S_n - W(n+1)| \sup_{0 < t \leq 1/2} t / q(t) \\ &\quad + \left| n^{-1/2} S_n \right| \sup_{0 < t \leq 1/2} \left| t - \frac{[(n+1)t]}{n} \right| / q(t), \\ &= o_P(1) + O_P(1)o_P(1), \end{aligned}$$

by combining (6.1.15), (6.1.11) and (6.1.13). Consequently

$$(6.1.16) \quad I_2^{(1)}(n) = o_P(1)$$

for any $q \in \mathcal{Q}^*$ and such that (6.1.6) holds.

Next we have

$$\begin{aligned} I_3^{(1)}(n) &= \sup_{0 < t \leq 1/2} n^{-1/2} \left| \frac{[(n+1)t]}{n} S_{[(n+1)t]}^{(1)} - tW^{(1)}((n+1)t) \right| / q(t) \\ &\leq \sup_{0 < t \leq 1/2} n^{-1/2} \left| S_{[(n+1)t]}^{(1)} - W^{(1)}((n+1)t) \right| \sup_{0 < t \leq 1/2} t / q(t), \\ &\quad + \sup_{0 < t \leq 1/2} \left| n^{-1/2} S_{[(n+1)t]}^{(1)} \right| \sup_{0 < t \leq 1/2} \left| t - \frac{[(n+1)t]}{n} \right| / q(t). \end{aligned}$$

which, by (6.1.8) with $q \equiv 1$, (6.1.11) and (6.1.13), gives

$$(6.1.17) \quad I_3^{(1)}(n) = o_P(1),$$

for any $q \in \mathcal{Q}^*$ and such that (6.1.6) holds.

For $t \in (1/2, 1)$ we obtain

$$\begin{aligned} & \sup_{1/2 \leq t < 1} \left| n^{-1/2} \sigma^{-1} \left\{ \left(1 - \frac{[(n+1)t]}{n} \right) \sum_{i=1}^{[(n+1)t]} \tilde{h}(X_i) + \frac{[(n+1)t]}{n} \left(\sum_{i=1}^n \tilde{h}(X_i) - \sum_{i=1}^{[(n+1)t]} \tilde{h}(X_i) \right) \right\} \right. \\ & \quad \left. - n^{1/2} \left\{ (1-t)W((n+1)t) + t(W(n+1) - W((n+1)t)) \right\} \right| / q(t) \\ & \leq \sup_{1/2 \leq t < 1} \left| n^{-1/2} \left\{ \left(1 - \frac{[(n+1)t]}{n} \right) (S_{[(n+1)/2]}^{(1)} + S_{[(n+1)/2]}^{(2)} - S_{[(n+1)t]}^{(2)}) \right. \right. \\ & \quad \left. \left. - (1-t) \left(W^{(1)}\left(\frac{n+1}{2}\right) + W^{(2)}\left(\frac{n+1}{2}\right) - W^{(2)}((n+1)(1-t)) \right) \right\} \right| / q(t) \\ & \quad + \sup_{1/2 \leq t < 1} \left| n^{-1/2} \left\{ \frac{[(n+1)t]}{n} \left(S_n - (S_{[(n+1)/2]}^{(1)} + S_{[(n+1)/2]}^{(2)} - S_{[(n+1)t]}^{(2)}) \right) \right. \right. \\ & \quad \left. \left. - t \left(W(n+1) - \left(W^{(1)}\left(\frac{n+1}{2}\right) + W^{(2)}\left(\frac{n+1}{2}\right) - W^{(2)}((n+1)(1-t)) \right) \right) \right\} \right| / q(t) \\ & \leq \sup_{1/2 \leq t < 1} n^{-1/2} \left| \left(1 - \frac{[(n+1)t]}{n} \right) S_{[(n+1)t]}^{(2)} - (1-t)W^{(2)}((n+1)(1-t)) \right| / q(t) \\ & \quad + \sup_{1/2 \leq t < 1} n^{-1/2} \left| \left(1 - \frac{[(n+1)t]}{n} \right) S_n - (1-t)W(n+1) \right| / q(t) \\ & \quad + \sup_{1/2 \leq t < 1} n^{-1/2} \left| \frac{[(n+1)t]}{n} S_{[(n+1)t]}^{(2)} - tW^{(2)}((n+1)(1-t)) \right| / q(t) \\ & = I_1^{(2)}(n) + I_2^{(2)}(n) + I_3^{(2)}(n). \end{aligned}$$

We have

$$\begin{aligned} & I_1^{(2)}(n) \\ & = \sup_{1/2 \leq t < 1} n^{-1/2} \left| \left(1 - \frac{[(n+1)t]}{n} \right) S_{[(n+1)t]}^{(2)} - (1-t)W^{(2)}((n+1)(1-t)) \right| / q(t) \\ & \leq \sup_{1/2 \leq t < 1} n^{-1/2} \left| S_{[(n+1)t]}^{(2)} - W^{(2)}((n+1)(1-t)) \right| \sup_{1/2 \leq t < 1} (1-t) / q(t) \\ & \quad + \sup_{1/2 \leq t < 1} \left| n^{-1/2} S_{[(n+1)t]}^{(2)} \right| \sup_{1/2 \leq t < 1} \left| t - \frac{[(n+1)t]}{n} \right| / q(t). \end{aligned}$$

which, by (6.1.9) with $q \equiv 1$, (6.1.12) and (6.1.13), gives

$$(6.1.18) \quad I_1^{(2)}(n) = o_P(1)$$

for any $q \in \mathcal{Q}^*$ for which (6.1.6) holds.

Considering $I_2^{(2)}(n)$, we have

$$\begin{aligned} I_n^{(2)}(n) &= \sup_{1/2 \leq t < 1} n^{-1/2} \left| \left(1 - \frac{[(n+1)t]}{n} \right) S_n - (1-t)W(n+1) \right| / q(t) \\ &\leq n^{-1/2} \left| S_n - W(n+1) \right| \sup_{1/2 \leq t < 1} (1-t) / q(t) \\ &\quad + \left| n^{-1/2} S_n \right| \sup_{1/2 \leq t < 1} \left| t - \frac{[(n+1)t]}{n} \right| / q(t) \\ &\quad + o_P(1) + O_P(1)o_P(1), \end{aligned}$$

by combining (6.1.15), (6.1.12) and (6.1.13). Consequently

$$(6.1.19) \quad I_2^{(2)}(n) = o_P(1),$$

for any $q \in \mathcal{Q}^*$ for which (6.1.6) holds.

Finally, we have

$$\begin{aligned} I_3^{(2)}(n) &= \sup_{1/2 \leq t < 1} n^{-1/2} \left| \frac{[(n+1)t]}{n} S_{[(n+1)t]}^{(2)} - tW^{(2)}((n+1)(1-t)) \right| / q(t) \\ &\leq \sup_{1/2 \leq t < 1} n^{-1/2} \left| S_{[(n+1)t]}^{(2)} - W^{(2)}((n+1)(1-t)) \right| / q(t) \sup_{1/2 < t < 1} t \\ &\quad + \sup_{1/2 \leq t < 1} \left| n^{-1/2} S_{[(n+1)t]}^{(2)} \right| \sup_{1/2 \leq t < 1} \left| t - \frac{[(n+1)t]}{n} \right| / q(t) \\ &= \sup_{0 < t \leq 1/2} n^{-1/2} \left| S_{[(n+1)(1-t)]}^{(2)} - W^{(2)}((n+1)t) \right| / q(1-t) \\ &\quad + O_P(1)o_P(1), \end{aligned}$$

due to (6.1.13). Hence, using (6.1.9), we conclude with $q \in \mathcal{Q}^*$

$$(6.1.20) \quad I_1^{(2)}(n) = \begin{cases} o_P(1) & \text{if } I^*(q, c) < \infty \text{ for all } c > 0 \\ O_P(1) & \text{if } I^*(q, c) < \infty \text{ for some } c > 0. \end{cases}$$

Let $q \in \mathcal{Q}^*$. Combining (6.1.14) - (6.1.20) gives

$$(6.1.21) \quad \sup_{0 < t < 1} \left| n^{-1/2} \left\{ \left(1 - \frac{[(n+1)t]}{n} \right) \sum_{i=1}^{[(n+1)t]} \tilde{h}(X_i) + \frac{[(n+1)t]}{n} \left(\sum_{i=1}^n \tilde{h}(X_i) - \sum_{i=1}^{[(n+1)t]} \tilde{h}(X_i) \right) \right\} \right. \\ \left. - n^{-1/2} \sigma \{ (1-t)W((n+1)t) + t(W(n+1) - W((n+1)t)) \} \right| / q(t) \\ = \begin{cases} o_P(1) & \text{if } I^*(q, c) < \infty \text{ for all } c > 0 \\ O_P(1) & \text{if } I^*(q, c) < \infty \text{ for some } c > 0. \end{cases}$$

Now (6.1.7), together with (6.1.21), implies the "if" parts of (a) and (b).

The converse parts of (a) and (b) follow from the fact that

$$\sup_{0 < t < 1} \left| \frac{n^{-3/2}}{\sigma} \mathcal{U}_{[(n+1)t]} - \Gamma_n(t) \right| / q(t) = o_P(1)$$

with $q \in \mathcal{Q}^*$ implies

$$o_P(1) = \sup_{0 < t < 1/(n+1)} |\Gamma_n(t)|/q(t) \stackrel{D}{=} \sup_{0 < t < 1/(n+1)} |\Gamma(t)|/q(t),$$

as well as

$$o_P(1) = \sup_{n/(n+1) \leq t < 1} |\Gamma_n(t)|/q(t) \stackrel{D}{=} \sup_{n/(n+1) \leq t < 1} |\Gamma(t)|/q(t),$$

which by Theorem 5.1.C and Remark 6.1.1 is equivalent to $I^*(q, c) < \infty$ for all $c > 0$ (recall that $\{\Gamma(t), 0 \leq t \leq 1\}$ is a Gaussian process defined by (6.1.4)).

Similarly

$$\sup_{0 < t < 1} \left| \frac{n^{-3/2}}{\sigma} \mathcal{U}_{[(n+1)t]} - \Gamma_n(t) \right| / q(t) = O_P(1)$$

with $q \in \mathcal{Q}^*$ implies

$$O_P(1) = \sup_{0 < t < 1/(n+1)} |\Gamma_n(t)|/q(t) \stackrel{D}{=} \sup_{0 < t < 1/(n+1)} |\Gamma(t)|/q(t),$$

as well as

$$O_P(1) = \sup_{n/(n+1) < t < 1} |\Gamma_n(t)|/q(t) \stackrel{D}{=} \sup_{n/(n+1) < t < 1} |\Gamma(t)|/q(t),$$

which by Theorem 5.1.B and Remark 6.1.1 is equivalent to $I^*(q, c) < \infty$ for some $c > 0$.

Theorem 6.1.2. *Let $0 < p < \infty$ and $h(x, y) = h(y, x)$. We assume that H_0 holds and (6.1.1) and (6.1.2) are satisfied.*

(a) *We can define a sequence of Gaussian processes $\{\Gamma_n(t), 0 \leq t \leq 1\}$ such that (6.1.3) holds and with $q \in \mathcal{Q}^*$ we have*

$$(6.1.22) \quad \int_0^1 \left| \frac{n^{-3/2}}{\sigma} \mathcal{U}_{[(n+1)t]} - \Gamma_n(t) \right|^p / q(t) dt = o_P(1)$$

if and only if

$$(6.1.23) \quad \int_0^1 (t(1-t))^{p/2} / q(t) dt < \infty.$$

(b) *If $q \in \mathcal{Q}^*$, then*

$$(6.1.24) \quad \int_0^1 \left| \frac{n^{-3/2}}{\sigma} \mathcal{U}_{[(n+1)t]} \right|^p / q(t) dt \xrightarrow{D} \int_0^1 |\Gamma(t)|^p / q(t) dt$$

if and only if (6.1.23) holds.

Proof. (a) Let $q \in \mathcal{Q}^*$ and assume that (6.1.23) holds. Using the same arguments as at the beginning of the proof of Theorem 6.1.1, we obtain for any $\delta \in (0, 1/2)$

$$\begin{aligned} & \int_{\delta}^{1-\delta} \left| n^{-3/2} \mathcal{U}_{[(n+1)t]} - \bar{U}_n(t) \right|^p / q(t) dt \\ & \leq \sup_{\delta \leq t \leq 1-\delta} \left(\left| n^{-3/2} \mathcal{U}_{[(n+1)t]} - \bar{U}_n(t) \right|^p \right) \int_{\delta}^{1-\delta} 1/q(t) dt \\ & = o_P(1) O(1) = o_P(1), \end{aligned}$$

as well as

$$\begin{aligned} & \int_{1/(n+1)}^{\delta} \left| n^{-3/2} \mathcal{U}_{[(n+1)t]} - \bar{U}_n(t) \right|^p / q(t) dt \\ & \leq \sup_{1/(n+1) \leq t \leq \delta} \left(\left| n^{-3/2} \mathcal{U}_{[(n+1)t]} - \bar{U}_n(t) \right| / t^{1/2} \right)^p \int_0^{\delta} t^{p/2} / q(t) dt \\ & = O_P(1) \int_0^{\delta} t^{p/2} / q(t) dt \end{aligned}$$

and

$$\begin{aligned} & \int_{1-\delta}^{n/(n+1)} \left| n^{-3/2} \mathcal{U}_{[(n+1)t]} - \bar{U}_n(t) \right|^p / q(t) dt \\ & \leq O_P(1) \int_{1-\delta}^1 (1-t)^{p/2} / q(t) dt, \end{aligned}$$

with

$$\bar{U}_n(t) = n^{-1/2} \left\{ \left(1 - \frac{[(n+1)t]}{n} \right) \sum_{i=1}^{[(n+1)t]} \tilde{h}(X_i) + \frac{[(n+1)t]}{n} \left(\sum_{i=1}^n \tilde{h}(X_i) - \sum_{i=1}^{[(n+1)t]} \tilde{h}(X_i) \right) \right\}.$$

Taking $\delta > 0$ arbitrarily small and using (6.1.23), we conclude

$$\begin{aligned} & \int_0^1 \left| n^{-3/2} \mathcal{U}_{[(n+1)t]} - n^{-1/2} \left\{ \left(1 - \frac{[(n+1)t]}{n} \right) \sum_{i=1}^{[(n+1)t]} \tilde{h}(X_i) + \frac{[(n+1)t]}{n} \left(\sum_{i=1}^n \tilde{h}(X_i) \right. \right. \right. \\ (6.1.25) \quad & \left. \left. \left. - \sum_{i=1}^{[(n+1)t]} \tilde{h}(X_i) \right) \right\} \right|^p / q(t) dt = o_P(1). \end{aligned}$$

We introduce

$$\begin{aligned} S_{[(n+1)t]}^{(1)} &= \sigma^{-1} \sum_{i=1}^{[(n+1)t]} \tilde{h}(X_i), \quad 0 \leq t \leq 1/2 \\ S_{[(n+1)t]}^{(2)} &= \sigma^{-1} \sum_{i=[(n+1)t]+1}^n \tilde{h}(X_i), \quad 1/2 \leq t < 1. \end{aligned}$$

By Theorem 1.4.1 there exists two independent Wiener processes $\{W^{(1)}(t), 0 \leq t < \infty\}$ and $\{W^{(2)}(t), 0 \leq t < \infty\}$ such that

$$(6.1.26) \quad \int_0^{1/2} \left| n^{-1/2} \left(S_{[(n+1)t]}^{(1)} - W^{(1)}((n+1)t) \right) \right|^p / q(t) dt = o_P(1)$$

and

$$(6.1.27) \quad \int_0^{1/2} \left| n^{-1/2} \left(S_{[(n+1)(1-t)]}^{(2)} - W^{(2)}((n+1)t) \right) \right|^p / q(t) dt = o_P(1).$$

We define a Wiener process W by

$$W(nt) = \begin{cases} W^{(1)}(nt), & 0 \leq t \leq 1/2 \\ W^{(1)}\left(\frac{n}{2}\right) + W^{(2)}\left(\frac{n}{2}\right) - W^{(2)}(n - nt), & 1/2 < t \leq 1, \end{cases}$$

and conclude from (6.1.26) and (6.1.27) that

$$(6.1.28) \quad \int_0^1 \left| n^{-1/2} \left\{ \left(1 - \frac{[(n+1)t]}{n} \right) \sum_{i=1}^{[(n+1)t]} \dot{h}(X_i) + \frac{[(n+1)t]}{n} \left(\sum_{i=1}^n \dot{h}(X_i) - \sum_{i=1}^{[(n+1)t]} \dot{h}(X_i) \right) \right\} \right. \\ \left. - n^{-1/2} \sigma \left\{ (1-t)W((n+1)t) + t(W(n+1) - W((n+1)t)) \right\} \right|^p / q(t) dt \\ = o_P(1).$$

Now (6.1.25), together with (6.1.28), implies (6.1.22).

Next we assume that (6.1.22) holds. Then we have

$$\int_0^{1/(n+1)} |\Gamma(t)|^p / q(t) dt = o_P(1) = \int_{n/(n+1)}^1 |\Gamma(t)|^p / q(t) dt.$$

Hence there exists K such that

$$P \left\{ \int_0^1 |\Gamma(t)|^p / q(t) dt > K \right\} < 1/2,$$

and, along the lines of the proof of Theorem 5.3.1, it follows from Corollary 2.1 of Csörgő, Horváth and Shao (1991) (cf. Theorem 1.3.A) that

$$\int_0^1 |\Gamma(t)|^p / q(t) dt < \infty \text{ a.s.}$$

if and only if (6.1.23) holds. Hence the result.

(b) The proof can be done similarly to that of part (b) of Theorem 5.3.1.

6.2. Antisymmetric kernel

When considering an antisymmetric kernel, i.e., $h(x, y) = -h(y, x)$, similar type of results hold. However, the limiting process is different. In this case $Eh(X_1, X_2) = 0$ and, similarly to the symmetric case, we let $\tilde{h}(t) = Eh(t, X_1)$. We assume

$$(6.2.1) \quad Eh^2(X_1, X_2) < \infty \quad \text{and} \quad 0 < \sigma^2 = E\tilde{h}^2(X_1).$$

Now we have $\mathcal{U}_k = \mathcal{Z}_k$, where $\mathcal{Z}_k$ is defined by (6.0.1). We note that (6.1.3) remains true in the case of an antisymmetric kernel with θ taken to be zero.

The following approximation of $\mathcal{U}_k$ process by a sequence of Brownian bridges was obtained by Csörgő and Horváth (1988a).

Theorem 6.2.A. *Let $X_1, X_2, \dots$ are i.i.d.r.v.'s and $h(x, y) = -h(y, x)$. We assume that (6.2.1) holds. Then we can define a sequence of Brownian bridges $\{B_n(t); 0 \leq t \leq 1\}$ such that as $n \rightarrow \infty$,*

$$\sup_{0 \leq t \leq 1} \left| \frac{n^{-3/2}}{\sigma} \mathcal{U}_{[(n+1)t]} - B_n(t) \right| = o_P(1).$$

Considering now the problem of weighted approximation for $\mathcal{U}_k$, we prove the following analog of Theorem 6.1.1.

Theorem 6.2.1. *Let $X_1, X_2, \dots$ be i.i.d.r.v.'s and $h(x, y) = -h(y, x)$. We assume that (6.2.1) is satisfied. Then we can define a sequence of Brownian bridges $\{B_n(t), 0 \leq t \leq 1\}$ such that with $q \in \mathcal{Q}^*$ we have*

$$\sup_{0 < t < 1} \left| \frac{n^{-3/2}}{\sigma} \mathcal{U}_{[(n+1)t]} - B_n(t) \right| / q(t) = o_P(1)$$

if and only if $I^*(q, c) < \infty$ for all $c > 0$, and

$$\sup_{0 < t < 1} \left| \frac{n^{-3/2}}{\sigma} \mathcal{U}_{[(n+1)t]} - B_n(t) \right| / q(t) = O_P(1)$$

if and only if $I^*(q, c) < \infty$ for some $c > 0$.

A remark, similar to Remark 6.1.2, is also valid here.

Proof. We assume that $q \in \mathcal{Q}^*$ and is such that

$$(6.2.2) \quad \lim_{t \downarrow 0} q(t)/t^{1/2} = \lim_{t \uparrow 1} q(t)/(1-t)^{1/2} = \infty.$$

When proving Theorem 6.2.A, Csörgö and Horváth (1988a) use Theorem 2.1 of Janson and Wichura (1983) and obtain

$$\max_{1 \leq k \leq n} \left| \mathcal{U}_k - \left\{ n \sum_{i=1}^k \tilde{h}(X_i) - k \sum_{i=1}^n \tilde{h}(X_i) \right\} \right| = O_P(n).$$

Hence

$$\sup_{0 \leq t < 1} \left| n^{-3/2} \mathcal{U}_{[(n+1)t]} - \bar{\bar{\mathcal{U}}}_n(t) \right| = o_P(1)$$

as well as

$$\begin{aligned} & \sup_{1/(n+1) \leq t \leq n/(n+1)} \left| n^{-3/2} \mathcal{U}_{[(n+1)t]} - \bar{\bar{\mathcal{U}}}_n(t) \right| / (t(1-t))^{1/2} \\ & \leq O_P(1) \sup_{1/(n+1) \leq t \leq n/(n+1)} 1/(nt(1-t))^{1/2} \\ & = O_P(1), \end{aligned}$$

with

$$\bar{\bar{\mathcal{U}}}_n(t) = n^{-1/2} \left(\sum_{i=1}^{[(n+1)t]} \tilde{h}(X_i) - \frac{[(n+1)t]}{n} \sum_{i=1}^n \tilde{h}(X_i) \right).$$

Consequently, for any $\delta \in (0, 1/2)$ we have

$$\sup_{\delta \leq t \leq 1-\delta} \left| n^{-3/2} \mathcal{U}_{[(n+1)t]} - \bar{\bar{\mathcal{U}}}_n(t) \right| / q(t) = o_P(1)$$

due to (6.5.1), and also

$$\begin{aligned} & \sup_{1/(n+1) \leq t < \delta} \left| n^{-3/2} \mathcal{U}_{[(n+1)t]} - \bar{\bar{\mathcal{U}}}_n(t) \right| / q(t) \\ & \leq O_P(1) \sup_{0 < t < \delta} t^{1/2} / q(t) \end{aligned}$$

as well as

$$\begin{aligned} & \sup_{1-\delta < t < n/(n+1)} \left| n^{-3/2} \mathcal{U}_{[(n+1)t]} - \bar{\bar{\mathcal{U}}}_n(t) \right| / q(t) \\ & \leq O_P(1) \sup_{1-\delta < t < 1} (1-t)^{1/2} / q(t). \end{aligned}$$

with

$$\bar{\mathcal{U}}_n(t) = n^{-1/2} \left(\sum_{i=1}^{[(n+1)t]} \tilde{h}(X_i) - \frac{[(n+1)t]}{n} \sum_{i=1}^n \tilde{h}(X_i) \right).$$

Taking δ arbitrarily small and using (6.2.2), we obtain

$$\sup_{0 < t < 1} \left| n^{-3/2} \mathcal{U}_{[(n+1)t]} - n^{-1/2} \left(\sum_{i=1}^{[(n+1)t]} \tilde{h}(X_i) - \frac{[(n+1)t]}{n} \sum_{i=1}^n \tilde{h}(X_i) \right) \right| / q(t) = o_P(1).$$

Since $I^*(q, c) < \infty$ for some $c > 0$ implies (6.2.2) (cf. Theorem 5.1.A), it is enough to show that there exists a sequence of Brownian bridges $\{B_n(t); 0 \leq t \leq 1\}$ such that

$$\sup_{0 < t < 1} \left| n^{-1/2} \sigma^{-1} \left(\sum_{i=1}^{[(n+1)t]} \tilde{h}(X_i) - \frac{[(n+1)t]}{n} \sum_{i=1}^n \tilde{h}(X_i) \right) - B_n(t) \right| / q(t) = o_P(1)$$

if and only if $I^*(q, c) < \infty$ for all $c > 0$, and

$$\sup_{0 < t < 1} \left| n^{-1/2} \sigma^{-1} \left(\sum_{i=1}^{[(n+1)t]} \tilde{h}(X_i) - \frac{[(n+1)t]}{n} \sum_{i=1}^n \tilde{h}(X_i) \right) - B_n(t) \right| / q(t) = O_P(1)$$

if and only if $I^*(q, c) < \infty$ for some $c > 0$.

Since $\tilde{h}(X_i)$, $i \geq 1$ are i.i.d.r.v.'s with two moments being finite, Theorem 5.2.1 yields the result.

For the L_p -functionals we have the following result.

Theorem 6.2.2. Let $X_1, X_2, \dots$ be i.i.d.r.v.'s and $h(x, y) = -h(y, x)$. We assume that $0 < p < \infty$ and (6.2.1) holds, i.e.

$$Eh^2(X_1, X_2) < \infty \quad \text{and} \quad 0 < \sigma^2 = Eh(X_1).$$

(a) We can define a sequence of Brownian bridges $\{B_n(t), 0 \leq t \leq 1\}$ such that with $q \in \mathcal{Q}^*$ we have

$$\int_0^1 \left| \frac{n^{-3/2}}{\sigma} \mathcal{U}_{[(n+1)t]} - B_n(t) \right|^p / q(t) dt = o_P(1)$$

if and only if

$$(6.2.3) \quad \int_0^1 (t(1-t))^{p/2} / q(t) dt < \infty.$$

(b) If $q \in \mathcal{Q}^*$, then

$$\int_0^1 \left| \frac{n^{-3/2}}{\sigma} \mathcal{U}_{[(n+1)t]} \right|^p / q(t) dt \xrightarrow{P} \int_0^1 |B(t)|^p / q(t) dt$$

if and only if (6.2.3) holds.

Proof. Let $q \in \mathcal{Q}^*$ and assume that (6.2.3) holds. By the same arguments as at the beginning of the proof of Theorem 6.2.1 we obtain for any $\delta \in (0, 1/2)$

$$\begin{aligned} & \int_{\delta}^{1-\delta} \left| n^{-3/2} \mathcal{U}_{[(n+1)t]} - \bar{\bar{U}}_n(t) \right|^p / q(t) dt \\ & \leq \sup_{\delta \leq t \leq 1-\delta} \left(\left| n^{-3/2} \mathcal{U}_{[(n+1)t]} - \bar{\bar{U}}_n(t) \right|^p \right) \int_{\delta}^{1-\delta} 1/q(t) dt \\ & = o_P(1) O(1) = o_P(1), \end{aligned}$$

as well as

$$\begin{aligned} & \int_{1/(n+1)}^{\delta} \left| n^{-3/2} \mathcal{U}_{[(n+1)t]} - \bar{\bar{U}}_n(t) \right|^p / q(t) dt \\ & \leq \sup_{1/(n+1) \leq t \leq \delta} \left(\left| n^{-3/2} \mathcal{U}_{[(n+1)t]} - \bar{\bar{U}}_n(t) \right| / t^{1/2} \right)^p \int_0^{\delta} t^{p/2} / q(t) dt \\ & = O_P(1) \int_0^{\delta} t^{p/2} / q(t) dt \end{aligned}$$

and

$$\begin{aligned} & \int_{1-\delta}^{n/(n+1)} \left| n^{-1/2} \mathcal{U}_{[(n+1)t]} - \bar{\bar{U}}_n(t) \right|^p / q(t) dt \\ & \leq O_P(1) \int_{1-\delta}^1 (1-t)^{p/2} / q(t) dt. \end{aligned}$$

with

$$\bar{\bar{U}}_n(t) = n^{-1/2} \left(\sum_{i=1}^{[(n+1)t]} \tilde{h}(X_i) - \frac{[(n+1)t]}{n} \sum_{i=1}^n \tilde{h}(X_i) \right).$$

Taking $\delta > 0$ arbitrarily small and using (6.2.3), we conclude

$$\int_0^1 \left| n^{-1/2} \mathcal{U}_{[(n+1)t]} - n^{-1/2} \left(\sum_{i=1}^{[(n+1)t]} \tilde{h}(X_i) - \frac{[(n+1)t]}{n} \sum_{i=1}^n \tilde{h}(X_i) \right) \right|^p / q(t) dt = o_P(1).$$

Now, by arguments as in the proof of Theorem 6.2.1, the results follow by Theorem 5.3.1.

6.3. Contiguous alternatives

Let $X_1, X_2, \dots$ be independent random variables and $\mathcal{Z}_k$ be the process defined by (6.0.1). In Sections 6.1 and 6.2 we proved approximations of $\mathcal{Z}_k$ for the optimal class of weight functions under the null hypothesis

$$H_0 : X_i, 1 \leq i \leq n, \text{ have the same distribution } F,$$

assuming the existence of two moments only. Here we derive asymptotic distributions of $\mathcal{Z}_k$ under the alternative hypothesis introduced in Chapter 2 (cf. Section 2.3), namely

$$H_1 : X_i, 1 \leq i \leq n, \text{ have the respective distribution functions } F_{in},$$

where all F_{in} are assumed to be absolutely continuous with respect to the distribution function F and

$$(6.3.1) \quad \left[\frac{dF_{in}}{dF}(F^{-1}(u)) \right]^{1/2} = 1 + \frac{1}{2n^{1/2}} g_n(t, u), \quad \frac{i-1}{n} < t \leq \frac{i}{n},$$

with $F^{-1}(u) = \inf\{x : F(x) \geq u\}$, $0 < u \leq 1$, $F^{-1}(0) = F^{-1}(0+)$ (cf. (2.3.1)).

We assume also that there exists a function $g \in L^2[0, 1]^2$ such that

$$(6.3.2) \quad \int_0^1 g(t, u) du = 0 \quad \text{for almost all } t \in [0, 1]$$

and

$$(6.3.3) \quad \int_0^1 \int_0^1 \left[g_n(t, u) - g\left(\frac{[nt]}{n}, u\right) \right]^2 dudt \rightarrow 0, \quad n \rightarrow \infty.$$

We show that the weak convergence of weighted $\mathcal{Z}_k$ process under H_1 also holds for the same optimal class of weight functions as in the case of the null hypothesis H_0 .

The sequence of direct products $F_{1n} \times \dots \times F_{nn}$, $n = 1, 2, \dots$ is contiguous to the sequence $F \times \dots \times F$ (cf. Section 2.3). For accommodating changepoint alternatives, we assume that there exists $\eta \in (0, 1)$ such that

$$(6.3.4) \quad g(t, u) = \mathbf{1}\{t \geq \eta\}g(u)$$

for some square integrable function g .

For the sake of stating and proving our results, we will assume without loss of generality that under H_0 the X_i are of the form $F^{-1}(U_i)$, where U_i are independent, identically distributed uniform $-(0, 1)$ random variables, since $X_i \stackrel{\mathcal{D}}{=} F^{-1}(U_i)$. Whenever convenient and not misleading, we will simply continue writing X_i .

Define the centered log-likelihood ratio process $L_n(\cdot)$ by

$$(6.3.5) \quad L_n(t) = \sum_{i=1}^{[nt]} \left\{ \log \frac{dF_{in}}{dF}(F^{-1}(U_i)) - E \log \frac{dF_{in}}{dF}(F^{-1}(U_i)) \right\}.$$

First we consider a symmetric kernel, i.e. $h(x, y) = h(y, x)$. As in Section 6.1, under the null assumption H_0 we assume (6.1.1), i.e. $Eh^2(X_1, X_2) < \infty$ and let $Eh(X_1, X_2) = \theta$, $\tilde{h}(t) = E\{h(X_1, t) - \theta\}$. Condition (6.1.1) implies that $E\tilde{h}^2(X_1) < \infty$ and we assume (6.1.2), i.e. $0 < \sigma^2 = E\tilde{h}^2(X_1)$.

Let (Γ, L) be the two-dimensional Gaussian process, where Γ is the process defined by (6.1.4), L is a Gaussian process with mean zero and covariance function $Q_L(t_1 \wedge t_2)$, where

$$(6.3.6) \quad Q_L(t) = \int_0^t \int_0^1 g^2(s, u) \, du \, ds,$$

and the covariance function of Γ and L is

$$(6.3.7) \quad E\Gamma(t_1)L(t_2) = C(t_1, t_2) = (1 - t_1)G(t_1 \wedge t_2) + t_1\{G(t_2) - G(t_1 \wedge t_2)\},$$

where

$$(6.3.8) \quad G(t) = \frac{1}{\sigma} \int_0^t \int_0^1 g(s, u) \tilde{h}(F^{-1}(u)) \, du \, ds.$$

Theorem 6.3.1. *Let $h(x, y) = h(y, x)$. We assume that under H_0 (6.1.1) and (6.1.2) hold, i.e.*

$$Eh^2(X_1, X_2) < \infty \quad \text{and} \quad 0 < \sigma^2 = E\tilde{h}^2(X_1).$$

(a) *Assuming that H_0 holds, we have, as $n \rightarrow \infty$,*

$$\left(\frac{n^{-3/2}}{\sigma} \mathcal{U}_{[(n+1)t]}, L_n(t) \right) \xrightarrow{\mathcal{D}} (\Gamma(t), L(t)) \quad \text{in } D^2[0, 1].$$

(b) *If H_1 holds, then, as $n \rightarrow \infty$,*

$$\frac{n^{-3/2}}{\sigma} \mathcal{U}_{[(n+1)t]} \xrightarrow{\mathcal{D}} \Gamma(t) + D(t) \quad \text{in } D[0, 1],$$

where

$$(6.3.9) \quad D(t) = \frac{(1-t)}{\sigma} \int_0^t \int_0^1 g(s, u) \dot{h}(F^{-1}(u)) du ds \\ + \frac{t}{\sigma} \left(\int_0^1 \int_0^1 g(s, u) \dot{h}(F^{-1}(u)) du ds - \int_0^t \int_0^1 g(s, u) \dot{h}(F^{-1}(u)) du ds \right).$$

Proof. (a) As noted at the beginning of the proof of Theorem 6.1.1, we have

$$\sup_{0 < t < 1} \left| \frac{n^{-3/2}}{\sigma} \mathcal{U}_{[(n+1)t]} - \left\{ \left(1 - \frac{[(n+1)t]}{n} \right) \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^{[(n+1)t]} \dot{h}(X_i) \right. \right. \\ \left. \left. + \frac{[(n+1)t]}{n} \left(\frac{1}{\sigma\sqrt{n}} \sum_{i=1}^n \dot{h}(X_i) - \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^{[(n+1)t]} \dot{h}(X_i) \right) \right\} \right| = o_P(1),$$

so it is sufficient to prove the desired joint limit theorem for $(\sigma^{-1} \bar{U}_n, L_n)$, where

$$\bar{U}_n(t) = n^{-1/2} \left\{ \left(1 - \frac{[(n+1)t]}{n} \right) \sum_{i=1}^{[(n+1)t]} \dot{h}(X_i) + \frac{[(n+1)t]}{n} \left(\sum_{i=1}^n \dot{h}(X_i) \right. \right. \\ \left. \left. - \sum_{i=1}^{[(n+1)t]} \dot{h}(X_i) \right) \right\}.$$

Hence the result follows from Lemma 2.4.1.

(b) By part (a) and Le Cam's third lemma (cf Lemma 2.1.F), all finite dimensional distributions of $\frac{n^{-3/2}}{\sigma} \mathcal{U}_{[(n+1)t]}$ converge to those of $\Gamma + D$, where $D(t) = \text{Cov}(\Gamma(t), L(1))$. Since contiguity preserves tightness, using Theorem 6.1.A, part b) of theorem is also proved.

Corollary 6.3.1. Let $h(x, y) = h(y, x)$. Assume that under H_0 (6.1.1) and (6.1.2) hold. Then, under the changepoint alternatives (6.3.4), we have, as $n \rightarrow \infty$

$$\frac{n^{-3/2}}{\sigma} \mathcal{U}_{[(n+1)t]} \xrightarrow{D} \Gamma(t) + D^n(t) \quad \text{in } D[0, 1],$$

where

$$(6.3.10) \quad D^q(t) = [(1-2t)(t-\eta)\mathbf{1}\{t \geq \eta\} + t(1-\eta)] \frac{1}{\sigma} \int_0^1 g(u) \dot{h}(F^{-1}(u)) du.$$

Proof. The result is immediate on account of

$$\int_0^t \int_0^1 \mathbf{1}\{s \geq \eta\} g(u) \dot{h}(F^{-1}(u)) duds = (t-\eta) \int_0^1 g(u) \dot{h}(F^{-1}(u)) du.$$

Let $\mathcal{Q}^*$ and $I^*(q, c)$ be defined as in Section 6.1. Next we show that weighted versions of our results in this section so far will hold under the same, two moments only, conditions.

Theorem 6.3.2. *Let $h(x, y) = h(y, x)$. Assume that under H_0 (6.1.1) and (6.1.2) are satisfied.*

(a) *Let $q \in \mathcal{Q}^*$ and $I^*(q, c) < \infty$ for all $c > 0$. Assume that H_0 holds. Then as $n \rightarrow \infty$*

$$\left(\frac{n^{-3/2} \mathcal{U}_{[(n+1)t]}}{\sigma q(t)}, L_n(t) \right) \xrightarrow{\mathcal{D}} \left(\frac{\Gamma(t)}{q(t)}, L(t) \right) \quad \text{in } D^2[0, 1].$$

(b) *Let $q \in \mathcal{Q}^*$ and $I^*(q, c) < \infty$ for all $c > 0$. Assume that H_1 holds. Then as $n \rightarrow \infty$*

$$\frac{n^{-3/2} \mathcal{U}_{[(n+1)t]}}{\sigma q(t)} \xrightarrow{\mathcal{D}} \frac{\Gamma(t) + D(t)}{q(t)} \quad \text{in } D[0, 1],$$

where $D(t)$ is as in (6.3.9).

(c) *Let $q : (0, 1) \rightarrow (0, \infty)$ be any function such that $\inf_{\delta \leq t \leq 1-\delta} q(t) > 0$ for all $\delta \in (0, 1/2)$. Assume that H_1 holds. Then as $n \rightarrow \infty$*

$$\frac{n^{-3/2} \mathcal{U}_{[(n+1)t]}}{\sigma q(t)} \xrightarrow{\mathcal{D}} \frac{\Gamma(t) + D(t)}{q(t)} \quad \text{in } D[\delta, 1-\delta],$$

for any $0 < \delta < \frac{1}{2}$.

Proof. (a) The weak convergence of $n^{-3/2}\mathcal{U}_{[(n+1)t]}/\sigma q(t)$ to $\Gamma(t)/q(t)$ follows from Theorem 6.1.1. As in the proof of Theorem 6.1.1, we have $\sup_{0 < t < 1} |n^{-3/2}\mathcal{U}_{[(n+1)t]} - \bar{\mathcal{U}}_n(t)|/q(t) = o_p(1)$, where

$$\begin{aligned}\bar{\mathcal{U}}_n(t) = n^{-1/2} & \left\{ \left(1 - \frac{[(n+1)t]}{n}\right) \sum_{i=1}^{[(n+1)t]} \dot{h}(X_i) \right. \\ & \left. + \frac{[(n+1)t]}{n} \left(\sum_{i=1}^n \dot{h}(X_i) - \sum_{i=1}^{[(n+1)t]} \dot{h}(X_i) \right) \right\}.\end{aligned}$$

We note that

$$E \frac{\sigma^{-1} \bar{\mathcal{U}}_n(t_1)}{q(t_1)} L_n(t_2) = \frac{1}{q(t_1)} E \sigma^{-1} \bar{\mathcal{U}}_n(t_1) L_n(t_2) \longrightarrow \frac{1}{q(t_1)} C(t_1, t_2),$$

where $C(t_1, t_2)$ is defined as in (6.3.7). The covariance structure is seen to be finite by similar arguments as used in the proof of Theorem 5.4.1. Namely, since $g \in L^2[0, 1]^2$ and $0 < \int_0^1 \dot{h}(F^{-1}(u))^2 du < \infty$, $D(t)$ is finite, where $D(t) = C(t, 1)$ is as in (6.3.9). We have also

$$\begin{aligned}|D(t)| &= \left| \frac{(1-t)}{\sigma} \int_0^t \int_0^1 g(s, u) \dot{h}(F^{-1}(u)) du ds + \frac{t}{\sigma} \int_t^1 \int_0^1 g(s, u) \dot{h}(F^{-1}(u)) du ds \right| \\ &\leq (1-t)t^{1/2} \left(\int_0^t \int_0^1 g^2(s, u) du ds \right)^{1/2} + t(1-t)^{1/2} \left(\int_t^1 \int_0^1 g^2(s, u) du ds \right)^{1/2},\end{aligned}$$

which implies

$$\sup_{0 < t < 1} |D(t)|/q(t) < \infty$$

due to the fact that for any $q \in \mathcal{Q}^*$ and such that $I^*(q, c) < \infty$ for all $c > 0$, we have (6.1.6). The rest of the proof is similar to that of Theorem 6.3.1.

(b) As in the proof of Theorem 6.3.1 (b), Le Cam's third lemma implies the result.

(c) By Theorem 6.3.1 (b), when H_1 holds, we have as $n \rightarrow \infty$

$$\frac{n^{-3/2}}{\sigma} \mathcal{U}_{[(n+1)t]} \xrightarrow{\mathcal{D}} \Gamma(t) + D(t) \quad \text{in } D[0, 1].$$

By the Skorohod-Dudley-Wichura theorem there exists a probability space with $\mathcal{Y}_n$, $n \geq 1$ and $\mathcal{Y}$ on it such that

$$(6.3.11) \quad \{\mathcal{Y}_n(t), 0 \leq t < 1\} \stackrel{D}{=} \left\{ \frac{n^{-3/2}}{\sigma} \mathcal{U}_{[(n+1)t]}, 0 \leq t \leq 1 \right\}, \quad n \geq 1,$$

$$\{\mathcal{Y}(t), 0 \leq t \leq 1\} \stackrel{D}{=} \{\Gamma(t) + D(t), 0 \leq t \leq 1\}$$

and

$$\sup_{0 \leq t < 1} |\mathcal{Y}_n(t) - \mathcal{Y}(t)| = o_P(1).$$

Since for any q such that $\inf_{\delta \leq t \leq 1-\delta} q(t) > 0$ for all $\delta \in (0, 1/2)$ we have

$$\sup_{\delta \leq t \leq 1-\delta} \left| \frac{\mathcal{Y}_n(t) - \mathcal{Y}(t)}{q(t)} \right| = o_P(1),$$

we get also

$$\frac{\mathcal{Y}_n}{q} \xrightarrow{D} \frac{\mathcal{Y}}{q} \quad \text{in } D[\delta, 1 - \delta].$$

Hence, on account of (6.3.11), we conclude

$$\frac{n^{-3/2} \mathcal{U}_{[(n+1)t]}}{\sigma q(t)} \xrightarrow{D} \frac{\Gamma(t) + D(t)}{q(t)}$$

in $D[\delta, 1 - \delta]$, for all $0 < \delta < 1/2$.

Corollary 6.3.2. Assume that $h(x, y) = h(y, x)$ and under H_0 (6.1.1), (6.1.2) are satisfied. Then under the changepoint alternatives (6.3.4) the following hold.

(a) Let $q \in \mathcal{Q}^*$ and $I^*(q, c) < \infty$ for all $c > 0$. Then as $n \rightarrow \infty$

$$(6.3.12) \quad \frac{n^{-3/2} \mathcal{U}_{[(n+1)t]}}{\sigma q(t)} \xrightarrow{D} \frac{\Gamma(t) + D^\eta(t)}{q(t)} \quad \text{in } D[0, 1],$$

where D^η is as in (6.3.10).

(b) Let $q : (0, 1) \rightarrow (0, \infty)$ be any function such that $\inf_{\delta \leq t \leq 1-\delta} q(t) > 0$ for all $\delta \in (0, 1/2)$. Then (6.3.12) holds in $D[\delta, 1-\delta]$, with $0 < \delta < \frac{1}{2}$.

When considering antisymmetric kernels, i.e., $h(x, y) = -h(y, x)$, similar results are obtained but, just like in the case of null hypothesis, the limiting processes are different. We recall that in this case $Eh(X_1, X_2) = 0$ and, similarly to the symmetric case, we let $\tilde{h}(t) = Eh(t, X_1)$. As in Section 6.2, we assume (6.2.1), i.e.,

$$Eh^2(X_1, X_2) < \infty \quad \text{and} \quad 0 < \sigma^2 = E\tilde{h}^2(X_1).$$

Define the two-dimensional Gaussian process (B, L) , where B is a Brownian bridge process, L is a Gaussian process with mean zero and covariance function $Q_L(t_1 \wedge t_2)$, where $Q_L(t)$ is defined by (6.3.6), and

$$(6.3.13) \quad EB(t_1)L(t_2) = S(t_1, t_2) = G(t_1 \wedge t_2) - t_1 G(t_2),$$

where

$$G(t) = \frac{1}{\sigma} \int_0^t \int_0^1 g(s, u) \tilde{h}(F^{-1}(u)) du ds.$$

Theorem 6.3.3. Assume that $h(x, y) = -h(y, x)$ and under H_0 (6.2.1) holds, i.e.

$$Eh^2(X_1, X_2) < \infty \quad \text{and} \quad 0 < \sigma^2 = E\tilde{h}^2(X_1).$$

(a) If H_0 holds, then we have, as $n \rightarrow \infty$,

$$\left(\frac{n^{-3/2}}{\sigma} \mathcal{U}_{[(n+1)t]}, L_n(t) \right) \xrightarrow{\mathcal{D}} (B(t), L(t)) \quad \text{in } D^2[0, 1].$$

(b) If H_1 holds, then as $n \rightarrow \infty$,

$$\frac{n^{-3/2}}{\sigma} \mathcal{U}_{[(n+1)t]} \xrightarrow{\mathcal{D}} B(t) + A(t) \quad \text{in } D[0, 1].$$

where

(6.3.14)

$$A(t) = \frac{1}{\sigma} \int_0^t \int_0^1 g(s, u) \tilde{h}(F^{-1}(u)) du ds - \frac{t}{\sigma} \int_0^1 \int_0^1 g(s, u) \tilde{h}(F^{-1}(u)) du ds.$$

Proof: (a) As noted at the beginning of the proof of Theorem 6.2.1, we have

$$\sup_{0 \leq t < 1} \left| \frac{n^{-3/2} \mathcal{U}_{[(n+1)t]}}{\sigma} - \left\{ \frac{1}{\sigma \sqrt{n}} \sum_{i=1}^{[(n+1)t]} \tilde{h}(X_i) - \frac{[(n+1)t]}{n} \frac{1}{\sigma \sqrt{n}} \sum_{i=1}^n \tilde{h}(X_i) \right\} \right| = o_P(1).$$

Hence it is sufficient to prove the desired joint limit theorem for $(\sigma^{-1} \bar{\bar{\mathcal{U}}}_n, L_n)$,

where

$$\bar{\bar{\mathcal{U}}}_n(t) = n^{-1/2} \left(\sum_{i=1}^{[(n+1)t]} \tilde{h}(X_i) - \frac{[(n+1)t]}{n} \sum_{i=1}^n \tilde{h}(X_i) \right).$$

Since $\tilde{h}(X_i)$, $i \geq 1$, are i.i.d.r.v.'s with two moments being finite, the result follows from Lemma 2.4.1.

(b) By part (a) and Le Cam's third lemma (cf. Lemma 2.1.F), all finite dimensional distributions of $\frac{n^{-3/2}}{\sigma} \mathcal{U}_{[(n+1)t]}$ converge to those of $B(t) + A(t)$, where $A(t) = \text{Cov}(B(t), L(1)) = G(t) - tG(1)$. Since contiguity preserves tightness, we use Theorem 6.2.A to conclude the result.

Corollary 6.3.3. Let $h(x, y) = -h(y, x)$. Assume that under H_0 (6.2.1) holds. Then, under the changepoint alternatives (6.3.4), we have

$$\frac{n^{-3/2}}{\sigma} \mathcal{U}_{[(n+1)t]} \xrightarrow{P} B(t) + A^\eta(t), \quad \text{in } D[0, 1],$$

where

$$(6.3.15) \quad A^\eta(t) = \{(t - \eta) \mathbf{1}\{t \geq \eta\} - t(1 - \eta)\} \frac{1}{\sigma} \int_0^1 g(u) \tilde{h}(F^{-1}(u)) du.$$

The following theorem is an analog of Theorem 6.3.2.

Theorem 6.3.4. Assume that $h(x, y) = -h(y, x)$ and under H_0 (6.2.1) is satisfied

(a) Let $q \in \mathcal{Q}^*$ and $I^*(q, c) < \infty$ for all $c > 0$. Assume that H_0 holds. Then, as $n \rightarrow \infty$,

$$\left(\frac{n^{-3/2} \mathcal{U}_{[(n+1)t]}}{\sigma q(t)}, L_n(t) \right) \xrightarrow{D} \left(\frac{B(t)}{q(t)}, L(t) \right) \quad \text{in } D^2[0, 1].$$

(b) Let $q \in \mathcal{Q}^*$ and $I^*(q, c) < \infty$ for all $c > 0$. Assume that H_1 holds. Then, as $n \rightarrow \infty$,

$$\frac{n^{-3/2} \mathcal{U}_{[(n+1)t]}}{\sigma q(t)} \xrightarrow{D} \frac{B(t) + A(t)}{q(t)} \quad \text{in } D[0, 1],$$

where $A(t)$ is as in (6.3.14).

(c) Let $q : (0, 1) \rightarrow (0, \infty)$ be any function such that $\inf_{\delta \leq t \leq 1-\delta} q(t) > 0$ for all $\delta \in (0, 1/2)$. Assume that H_1 holds. Then, as $n \rightarrow \infty$,

$$\frac{n^{-3/2} \mathcal{U}_{[(n+1)t]}}{\sigma q(t)} \xrightarrow{D} \frac{B(t) + A(t)}{q(t)} \quad \text{in } D[\delta, 1 - \delta],$$

for any $0 < \delta < \frac{1}{2}$.

Proof: (a) When H_0 holds, the weak convergence of $n^{-3/2} \mathcal{U}_{[(n+1)t]}/\sigma q(t)$ to $B(t)/q(t)$ follows from Theorem 6.2.1. As in the proof of Theorem 6.2.1 we have

$$\sup_{0 \leq t < 1} \left| n^{-3/2} \mathcal{U}_{[(n+1)t]} - \bar{\bar{U}}_n(t) \right| / q(t) = o_P(1),$$

where

$$\bar{\bar{U}}_n(t) = n^{-1/2} \left(\sum_{i=1}^{[(n+1)t]} \tilde{h}(X_i) - \frac{[(n+1)t]}{n} \sum_{i=1}^n \tilde{h}(X_i) \right).$$

Hence it is enough to prove the desired joint limit theorem for $(\bar{\bar{U}}_n(t)/\sigma q(t), L_n(t))$.

We note that

$$E \frac{\sigma^{-1} \bar{\bar{U}}_n(t_1)}{q(t_1)} L_n(t_2) = \frac{1}{q(t_1)} E \sigma^{-1} \bar{\bar{U}}_n(t_1) L_n(t_2) \rightarrow \frac{1}{q(t_1)} S(t_1, t_2),$$

where $S(t_1, t_2)$ is as in (6.3.13). The covariance function is seen to be finite by using similar arguments as in the proof of Theorem 6.3.2 (a).

(b) As in the proof of Theorem 6.3.1(b), Le Cam's third lemma implies the result.

(c) By Theorem 6.3.3(b), when H_1 holds, we have as $n \rightarrow \infty$

$$\frac{n^{-3/2}}{\sigma} \mathcal{U}_{[(n+1)t]} \xrightarrow{\mathcal{D}} B(t) + A(t) \quad \text{in } D[0, 1].$$

Using Skorohod-Dudley-Wichura theorem, the proof goes along the lines of that of Theorem 6.3.2(c).

Corollary 6.3.4. *Assume that $h(x, y) = -h(y, x)$ and under H_0 (6.2.1) is satisfied. Then under the changepoint alternatives (6.3.4) we have:*

(a) *If $q \in \mathcal{Q}^*$ and $I^*(q, c) < \infty$ for all $c > 0$, then, as $n \rightarrow \infty$,*

$$(6.3.16) \quad \frac{n^{-3/2} \mathcal{U}_{[(n+1)t]}}{\sigma q(t)} \xrightarrow{\mathcal{D}} \frac{B(t) + A^n(t)}{q(t)} \quad \text{in } D[0, 1],$$

where A^n is as in (6.3.15).

(b) *If q is as in c) of Theorem 6.3.4, then (6.3.16) holds in $D[\delta, 1 - \delta]$, for any $0 < \delta < \frac{1}{2}$.*

CHAPTER 7

Pontograms (Poisson bridges) with changepoints in their intensity function

7.0. Introduction

Let $\{N(x), x \geq 0\}$ be a nonhomogeneous, also called nonstationary, Poisson process with intensity function $\lambda(x)$, $x \geq 0$. We consider the problem of testing the null hypothesis of $\lambda(x)$ having a constant value against the alternative that $\lambda(x)$ is a function of x .

When looking for a possible changepoint in the Land's End data, Kendall and Kendall (1980) considered test-statistics based on the stochastic process

$$X_n(t) = \left(\frac{N(nt)}{nt} - \frac{N(n) - N(nt)}{n - nt} \right) \left(\frac{nt(n - nt)}{N(n)} \right)^{\frac{1}{2}}, \quad 0 < t < 1, \\ X_n(0) = X_n(1) = 0,$$

called by them a pontogram. We let

$$(7.0.1) \quad K_n(t) = n^{-\frac{1}{2}}(N(nt) - tN(n)), \quad 0 \leq t \leq 1$$

and note that

$$\left\{ \left(\frac{n}{N(n)} \right)^{\frac{1}{2}} \frac{K_n(t)}{(t(1-t))^{\frac{1}{2}}}, \quad 0 < t < 1, \quad n \geq 1 \right\} = \left\{ X_n(t), \quad 0 < t < 1, \quad n \geq 1 \right\}.$$

Assuming that $\lambda(x)$ has a constant, but unknown value $\lambda > 0$, Kendall and Kendall (1980) proved that, for all $0 < \epsilon_1 < 1 - \epsilon_2 < 1$, we have

$$(7.0.2) \quad \sup_{\epsilon_1 \leq t \leq 1 - \epsilon_2} \left(\frac{n}{N(n)} \right)^{\frac{1}{2}} \frac{|K_n(t)|}{(t(1-t))^{\frac{1}{2}}} \xrightarrow{D} \sup_{0 \leq s \leq \frac{1}{2} \log[(1 - \epsilon_2)/\epsilon_1][(1 - \epsilon_1)/\epsilon_2]} |V(s)|,$$

where $\{V(s), -\infty < s < \infty\}$ is an Ornstein-Uhlenbeck process with covariance function $\exp(-|t - s|)$. The problem of replacing ϵ_1 and ϵ_2 by zero, or by sequences

3 of/de 3

PM-1 3 1/2" x 4" PHOTOGRAPHIC MICROCOPY TARGET
NBS 1010a ANSI/ISO #2 EQUIVALENT

PRECISION™ RESOLUTION TARGETS

PRECISION OPTICALS, INC. FORT WORTH, TEXAS 76101

PRECISION

PRECISION OPTICALS, INC. FORT WORTH, TEXAS 76101
TEL. 817-735-1100 FAX 817-735-1101

$\epsilon_i(n) \rightarrow 0$ as $n \rightarrow \infty$, $i = 1, 2$, in (7.0.2) was solved by Csörgő and Horváth (1987c).

Let

$$a(x) = (2 \log x)^{\frac{1}{2}}$$

$$b(x) = 2 \log x + \frac{1}{2} \log \log x - \frac{1}{2} \log \pi,$$

$$c(x) = \log((1-x)/x),$$

and Y be a random variable with distribution function $\exp(-2 \exp(-x))$, $-\infty < x < \infty$. Assuming that $\lambda(x) \equiv \lambda > 0$ they prove the following result.

Theorem 7.0.A. *If $k_n \rightarrow \infty$ and $k_n/n \rightarrow 0$ as $n \rightarrow \infty$, then*

$$\begin{aligned} a(\log n) \lambda^{-\frac{1}{2}} \sup_{0 < t < 1} |K_n(t)| / (t(1-t))^{\frac{1}{2}} - b(\log n) &\xrightarrow{D} Y, \\ a\left(\frac{1}{2} \log k_n\right) \lambda^{-\frac{1}{2}} \sup_{0 < t \leq k_n/n} |K_n(t)| / (t(1-t))^{\frac{1}{2}} - b\left(\frac{1}{2} \log k_n\right) &\xrightarrow{D} Y, \\ a\left(\frac{1}{2} \log k_n\right) \lambda^{-\frac{1}{2}} \sup_{1-k_n/n \leq t < 1} |K_n(t)| / (t(1-t))^{\frac{1}{2}} - b\left(\frac{1}{2} \log k_n\right) &\xrightarrow{D} Y, \end{aligned}$$

and

$$a(c(k_n/n)) \lambda^{-\frac{1}{2}} \sup_{k_n/n \leq t \leq 1-k_n/n} |K_n(t)| / (t(1-t))^{\frac{1}{2}} - b(c(k_n/n)) \xrightarrow{D} Y,$$

as $n \rightarrow \infty$.

By the law of large numbers, λ can be replaced by $N(n)/n$ in Theorem 7.0.A. The statistics thus obtained were of interest to, and inspired by Kendall and Kendall (1980) (cf. e.g., (7.0.2)). We note that the normalizing factors $a(x)$, $b(x)$ and $c(x)$ are essential, since

$$\sup_{0 < t < 1} |K_n(t)| / (t(1-t))^{\frac{1}{2}} \xrightarrow{P} \infty$$

already under the null assumption. Such a test statistic will, however, converge in distribution to a nondegenerate random variable if we introduce weight functions which are less severe on the tails than $(t(1-t))^{\frac{1}{2}}$ is.

Here we consider the asymptotic behaviour of the process

$$\left\{ \left(\frac{n}{N(n)} \right)^{\frac{1}{2}} K_n(t)/q(t), \quad 0 < t < 1 \right\}$$

for the largest possible class of weight functions q (cf. also Szyszkowicz 1992d,e). Statistics based on such weighted processes should be very sensitive on the tails while remaining sensitive also in the middle. It will be seen that these weight functions q are such that

$$(7.0.3) \quad \lim_{t \downarrow 0} q(t)/t^{\frac{1}{2}} = \lim_{t \uparrow 1} q(t)/(1-t)^{\frac{1}{2}} = \infty.$$

Hence the process

$$(7.0.4) \quad \left(\frac{n}{N(n)} \right)^{\frac{1}{2}} K_n(t)/q(t), \quad 0 < t < 1,$$

can be also written as

$$\left(\frac{n}{N(n)} \right)^{\frac{1}{2}} \frac{K_n(t)}{(t(1-t))^{\frac{1}{2}}} / g(t), \quad 0 < t < 1,$$

where $\lim_{t \downarrow 0} g(t) = \lim_{t \uparrow 1} g(t) = \infty$. Consequently (7.0.4) can be viewed as weighted versions of pontograms considered by Kendall and Kendall (1980).

For example, assuming that $\lambda(x) \equiv \lambda > 0$, we have, as $n \rightarrow \infty$

$$(7.0.5) \quad \sup_{0 < t < 1} \left(\frac{n}{N(n)} \right)^{\frac{1}{2}} |K_n(t)|/q(t) \xrightarrow{D} \sup_{0 < t < 1} |B(t)|/q(t),$$

where $\{B(t), 0 \leq t \leq 1\}$ is a Brownian bridge, with

$$q(t) = \left(t(1-t) \log \log \frac{1}{t(1-t)} \right)^{\frac{1}{2}}, \quad 0 < t < 1.$$

We note that, modulo constants, this weight function appears to be optimal for having the weak convergence of the sup-functional as in (7.0.5) at all, while it

is too severe to allow the weak convergence of the whole process in (7.0.4) (cf. Proposition 7.1.2).

Studying the asymptotic behaviour of the stochastic process $K_n(t)$, in Section 7.1 we assume that $\lambda(x) \equiv \lambda > 0$, and describe necessary and sufficient conditions for weight functions q to have the weak convergence of $K_n(t)$ process and then that of its sup-functionals for a larger class of functions q . The asymptotics of weighted L_p norms, $0 < p < \infty$, of $K_n(t)$ are studied in Section 7.2.

Having the description of the behaviour of $K_n(t)/q(t)$ under the null hypothesis

$$H: \lambda(x) = \lambda > 0, \quad x \geq 0,$$

in Section 7.3 we study its asymptotic behaviour under the sequence of contiguous alternatives

$$A: \lambda(x) = \lambda + \xi\left(\frac{[x]}{n}\right)/n^{\frac{1}{2}}, \quad 0 \leq x \leq n, \quad n \geq 1,$$

where $\xi^2(t)$ is a Riemann integrable function.

In particular, A accommodates the so-called changepoint alternative

A^η : there exists $\eta \in (0, 1)$ such that

$$\lambda_1(x) = \lambda \quad \text{for } 0 \leq x < n\eta,$$

$$\lambda(x) = \lambda + \rho/n^{\frac{1}{2}} \quad \text{for } n\eta \leq x \leq n, \quad n \geq 1,$$

by taking $\xi(t) = \rho \mathbf{1}\{t \geq \eta\}$, $\rho \neq 0$. The value of ρ is not known a priori. One may want to pick an appropriate value for it, for a particular problem at hand. We note only that $\rho > 0$ ($\rho < 0$) corresponds to a higher (lower) value of the intensity function $\lambda(x)$ of the Poisson process after the changepoint η of this parameterization.

7.1. Weighted pontograms in sup-metrics

Let $\{N(x), x \geq 0\}$ be a nonhomogeneous Poisson process with intensity function $\lambda(x)$, $x \geq 0$, and $\{K_n(t), 0 < t < 1\}$ be a nonhomogeneous Poisson bridge as defined in (7.0.1). In this section we consider the asymptotic behaviour of the process $\{K_n(t)/q(t), 0 < t < 1\}$ for optimal classes of the weight functions q under the null hypothesis

$$H: \lambda(x) \equiv \lambda > 0, \quad x \geq 0.$$

As a consequence, we obtain also the limiting distribution of the weighted pontogram $\left\{ \left(\frac{n}{N(n)} \right)^{\frac{1}{2}} K_n(t)/q(t), 0 < t < 1 \right\}$.

Let Q^* be the class of functions q defined on $(0, 1)$ which are nondecreasing in a neighbourhood of zero and nonincreasing in a neighbourhood of one, and are such that

$$(7.1.1) \quad \inf_{\delta \leq t \leq 1-\delta} q(t) > 0 \quad \text{for all } \delta \in (0, 1/2).$$

We introduce the following integrals

$$E^*(q, c) = \int_0^1 (t(1-t))^{-\frac{3}{2}} q(t) \exp(-c(t(1-t))^{-1} q^2(t)) dt$$

and

$$I^*(q, c) = \int_0^1 (t(1-t))^{-1} \exp(-c(t(1-t))^{-1} q^2(t)) dt$$

for some constant $0 < c < \infty$.

The following two propositions summarize the asymptotic behaviour of the process K_n with weight functions $q \in Q^*$ under the null hypothesis H .

Proposition 7.1.1. *We assume that H holds. We can define a sequence of Brownian bridges $\{B_n(t), 0 \leq t \leq 1\}$ such that, as $n \rightarrow \infty$, we have:*

(a) Let $q \in \mathcal{Q}^*$. Then

$$(7.1.2) \quad \sup_{0 < t < 1} |\lambda^{-\frac{1}{2}} K_n(t) - B_n(t)|/q(t) = o_P(1)$$

if and only if $I^*(q, c) < \infty$ for all $c > 0$.

(b) Let $q \in \mathcal{Q}^*$ and $I^*(q, c) < \infty$ for some $c > 0$. Then

$$(7.1.3) \quad \sup_{0 < t < 1} |\lambda^{-\frac{1}{2}} K_n(t) - B_n(t)|/q(t) = O_P(1).$$

Conversely, if $q \in \mathcal{Q}^*$ is such that (7.0.3) holds then (7.1.3) implies $I^*(q, c) < \infty$ for some $c > 0$.

Remark 7.1.1. According to Theorem 5.1.C, the statement (7.1.2) of Proposition 7.1.1 holds also if $E^*(q, c) < \infty$ for all $c > 0$ and (7.0.3) holds, and, conversely, if (7.1.2) is true for $q \in \mathcal{Q}^*$ and such that (7.0.3) holds, then $E^*(q, c) < \infty$ for all $c > 0$.

Similarly, according to Theorem 5.1.B, the statement (7.1.3) holds also if $E^*(q, c) < \infty$ for some $c > 0$ and (7.0.3) holds. Conversely, we have as well that (7.1.3) being true for $q \in \mathcal{Q}^*$ and such that (7.0.3) holds, implies $E^*(q, c) < \infty$ for some $c > 0$.

Proof. Let $q \in \mathcal{Q}^*$. By Y_1 we denote the time of the first jump of $N(x)$ and let Y_n be the time of the last jump of $N(x)$ before n , where $\{N(x), x \geq 0\}$ is a Poisson process with intensity parameter $\lambda > 0$. First we show that if $q \in \mathcal{Q}^*$ and (7.0.3) holds, i.e.,

$$\lim_{t \downarrow 0} q(t)/t^{1/2} = \lim_{t \uparrow 1} q(t)/(1-t)^{1/2} = \infty,$$

then there exists a sequence of Brownian bridges $\{B_n(t), 0 \leq t \leq 1\}$ such that, as $n \rightarrow \infty$

$$(7.1.4) \quad \sup_{Y_1/n \leq t < Y_n/n} |\lambda^{-\frac{1}{2}} K_n(t) - B_n(t)|/q(t) = o_P(1).$$

It was already proved by Csörgő and Horváth (1987c) (cf. Theorem 3.4.A) that there exists a sequence of Brownian bridges $\{B_n(t), 0 \leq t \leq 1\}$ such that

$$(7.1.5) \quad \sup_{\mu/n \leq t \leq 1-\mu/n} n^\nu |\lambda^{-\frac{1}{2}} K_n(t) - B_n(t)| / (t(1-t))^{\frac{1}{2}-\nu} = O_P(1)$$

for all $\mu > 0$ and $0 \leq \nu < \frac{1}{2}$.

In order to show that (7.1.4) holds, we note that Y_1 and $n - Y_n$ are exponential, with mean $1/\lambda$, random variables and hence, for each $\epsilon > 0$, there is a $\mu > 0$ such that $P\{Y_1 < \mu\} \leq \epsilon$ and $P\{n - Y_n < \mu\} \leq \epsilon$. Using now (7.1.5) with $\nu = 0$, we conclude that, as $n \rightarrow \infty$,

$$(7.1.6) \quad \sup_{Y_1/n \leq t < Y_n/n} |\lambda^{-\frac{1}{2}} K_n(t) - B_n(t)| / (t(1-t))^{\frac{1}{2}} = O_P(1).$$

Thus, with any $\delta \in (0, \frac{1}{2})$,

$$(7.1.7) \quad \sup_{Y_1/n \leq t < \delta} |\lambda^{-\frac{1}{2}} K_n(t) - B_n(t)| / q(t) \leq O_P(1) \sup_{0 < t < \delta} (t(1-t))^{\frac{1}{2}} / q(t),$$

and

$$(7.1.8) \quad \sup_{1-\delta < t < Y_n/n} |\lambda^{-\frac{1}{2}} K_n(t) - B_n(t)| / q(t) \leq O_P(1) \sup_{1-\delta < t < 1} (t(1-t))^{\frac{1}{2}} / q(t).$$

From the Proof of Theorem 1.1 of Csörgő and Horváth (1987c) we can conclude also that, as $n \rightarrow \infty$,

$$\sup_{0 < t < 1} |\lambda^{-\frac{1}{2}} K_n(t) - B_n(t)| = o_P(1).$$

Consequently, by (7.1.1), we have

$$(7.1.9) \quad \sup_{\delta \leq t \leq 1-\delta} |\lambda^{-\frac{1}{2}} K_n(t) - B_n(t)| / q(t) = o_P(1)$$

for any $\delta \in (0, \frac{1}{2})$. Since $\delta > 0$ can be taken arbitrarily small, (7.1.9), (7.1.7) and (7.1.8) together with (7.0.3) imply (7.1.4).

Next we note that with $q \in \mathcal{Q}^*$ and such that (7.0.3) holds, we have, as $n \rightarrow \infty$

$$(7.1.10) \quad \sup_{0 < t < Y_1/n} |\lambda^{-\frac{1}{2}} K_n(t)|/q(t) = \sup_{0 < t < Y_1/n} \lambda^{-\frac{1}{2}} t N(n)/n^{\frac{1}{2}} q(t) \\ \leq \lambda^{-\frac{1}{2}} Y_1^{\frac{1}{2}} N(n)/n \sup_{0 < t < Y_1/n} t^{\frac{1}{2}}/q(t) = o_P(1)$$

and

$$(7.1.11) \quad \sup_{Y_n/n \leq t < 1} |\lambda^{-\frac{1}{2}} K_n(t)|/q(t) = \sup_{Y_n/n \leq t < 1} \lambda^{-\frac{1}{2}} (1-t) N(n)/n^{\frac{1}{2}} q(t) \\ \leq \lambda^{-\frac{1}{2}} (n - Y_n)^{\frac{1}{2}} N(n)/n \sup_{Y_n/n \leq t < 1} (1-t)^{\frac{1}{2}}/q(t) = o_P(1),$$

on account of $Y_1/n \xrightarrow{a.s.} 0$, $Y_n/n \xrightarrow{a.s.} 1$ and $N(n)/n \xrightarrow{a.s.} \lambda$ as $n \rightarrow \infty$.

Let

$$(7.1.12) \quad \bar{B}_n(t) = \begin{cases} B_n(t) & \text{if } Y_1/n \leq t < Y_n/n \\ 0 & \text{otherwise,} \end{cases}$$

where, as before, Y_1 denotes the time of the first jump of $N(x)$ and Y_n is the last jump of $N(x)$ before n . So far we have shown (cf. (7.1.4), (7.1.10) and (7.1.11)) that there exists a sequence of Brownian bridges $\{B_n(t), 0 \leq t \leq 1\}$ such that with the sequence $\{\bar{B}_n(t), 0 \leq t \leq 1\}$ defined as in (7.1.12) and $q \in \mathcal{Q}^*$ such that (7.0.3) holds, we have, as $n \rightarrow \infty$

$$(7.1.13) \quad \sup_{0 < t < 1} |\lambda^{-\frac{1}{2}} K_n(t) - \bar{B}_n(t)|/q(t) = o_P(1).$$

In order to prove (a), we assume that $q \in \mathcal{Q}^*$ and $I^*(q, c) < \infty$ for all $c > 0$. Then we have (7.0.3) (cf. Theorem 5.1.A) and, consequently, also (7.1.13). Using now Theorem 5.1.C in combination with (7.1.13), we conclude (7.1.2).

Conversely, if $q \in \mathcal{Q}^*$ and (7.1.2) holds, then we have

$$\sup_{0 < t < 1} |\lambda^{-\frac{1}{2}} K_n(t)|/q(t) \xrightarrow{P} \sup_{0 < t < 1} |B(t)|/q(t),$$

where $\{B(t), 0 \leq t \leq 1\}$ is a Brownian bridge. Consequently

$$P \left\{ \sup_{0 < t < 1} |B(t)|/q(t) < \infty \right\} = 1.$$

Hence

$$P \left\{ \limsup_{t \downarrow 0} |B(t)|/q(t) < \infty \right\} = 1 = P \left\{ \limsup_{t \uparrow 1} |B(t)|/q(t) < \infty \right\}$$

and so, by Theorem 5.1.B, $q \in \mathcal{Q}^*$ must be such that $I^*(q, c) < \infty$ for some $c > 0$, which implies (7.0.3). Consequently we have (7.1.10) and (7.1.11) which together with (7.1.2) imply, as $n \rightarrow \infty$

$$\sup_{0 < t < Y_1/n} |B_n(t)|/q(t) \stackrel{P}{=} \sup_{0 < t < Y_1/n} |B(t)|/q(t) = o_P(1)$$

and

$$\sup_{Y_n/n \leq t < 1} |B_n(t)|/q(t) \stackrel{P}{=} \sup_{Y_n/n \leq t < 1} |B(t)|/q(t) = o_P(1).$$

Since $Y_1/n \xrightarrow{a.s.} 0$ and $Y_n/n \xrightarrow{a.s.} 1$, we now use Theorem 5.1.C to conclude that $q \in \mathcal{Q}^*$ has to be such that $I^*(q, c) < \infty$ for all $c > 0$.

In order to prove (b), we assume that $q \in \mathcal{Q}^*$ and $I^*(q, c) < \infty$ for some $c > 0$, which again implies (7.0.3). Hence (7.1.13) holds and we use Theorem 5.1.B to conclude (7.1.3).

Conversely, if we assume that (7.1.3) holds with $q \in \mathcal{Q}^*$ and such that (7.0.3) holds, then we have also (7.1.10) and (7.1.11). Hence we conclude that, as $n \rightarrow \infty$

$$\sup_{0 < t < Y_1/n} |B_n(t)|/q(t) \stackrel{P}{=} \sup_{0 < t < Y_1/n} |B(t)|/q(t) = O_P(1)$$

and

$$\sup_{Y_n/n \leq t < 1} |B_n(t)|/q(t) \stackrel{P}{=} \sup_{Y_n/n \leq t < 1} |B(t)|/q(t) = O_P(1).$$

Using now Theorem 5.1.B we have $I^*(q, c) < \infty$ for some $c > 0$.

Proposition 7.1.2. *We assume that H holds and let $\{B(t), 0 \leq t \leq 1\}$ be a Brownian bridge.*

(a) *Let $q \in \mathcal{Q}^*$. Then, as $n \rightarrow \infty$,*

$$(7.1.14) \quad \lambda^{-\frac{1}{2}} K_n(t)/q(t) \xrightarrow{\mathcal{D}} B(t)/q(t) \quad \text{in } D[0, 1]$$

if and only if $I^(q, c) < \infty$ for all $c > 0$.*

(b) *Let $q \in \mathcal{Q}^*$. Then, as $n \rightarrow \infty$,*

$$(7.1.15) \quad \sup_{0 < t < 1} \lambda^{-\frac{1}{2}} |K_n(t)|/q(t) \xrightarrow{\mathcal{D}} \sup_{0 < t < 1} |B(t)|/q(t)$$

if and only if $I^(q, c) < \infty$ for some $c > 0$.*

(c) *Let $q \in (0, 1) \rightarrow (0, \infty)$ be such that (7.1.1) holds. Then, as $n \rightarrow \infty$,*

$$\lambda^{-\frac{1}{2}} K_n(t)/q(t) \xrightarrow{\mathcal{D}} B(t)/q(t) \quad \text{in } D[\delta, 1 - \delta]$$

for any $\delta \in (0, \frac{1}{2})$.

A remark, similar to Remark 7.1.1, is also valid here.

Proof. (a) Obviously, with $q \in \mathcal{Q}^*$ and such that $I^*(q, c) < \infty$ for all $c > 0$, by Proposition 7.1.1 (a) we have that (7.1.14) holds. Conversely, assuming now (7.1.14) with $q \in \mathcal{Q}^*$, by the Skorohod-Dudley-Wichura theorem (cf. Theorem A.2.1) there exists $\tilde{K}_n$, $n \geq 1$, and $\tilde{B}$ such that

$$\{\tilde{K}_n, 0 \leq t \leq 1\} \xrightarrow{\mathcal{D}} \{\lambda^{-\frac{1}{2}} K_n(t), 0 \leq t \leq 1\}, \quad n \geq 1,$$

$$\{\tilde{B}(t), 0 \leq t \leq 1\} \xrightarrow{\mathcal{D}} \{B(t), 0 \leq t \leq 1\}$$

and

$$\sup_{0 < t < 1} |\tilde{K}_n(t) - \tilde{B}(t)|/q(t) = o_P(1).$$

Hence, using the technique of Lemma 3.1.2 in Csörgő (1983) (cf. Section A.1), there exists a sequence of Brownian bridges $\{B_n(t), 0 \leq t \leq 1\}$ such that, as $n \rightarrow \infty$,

$$\sup_{0 < t < 1} |\lambda^{-\frac{1}{2}} K_n(t) - B_n(t)|/q(t) = o_P(1),$$

which, by Proposition 7.1.1 (a) implies $I^*(q, c) < \infty$ for all $c > 0$.

(b) When proving Proposition 7.1.1, we have shown (cf. (7.1.13)) that with the sequence $\bar{B}_n$ defined as in (7.1.12) we have, as $n \rightarrow \infty$

$$\sup_{0 < t < 1} |\lambda^{-\frac{1}{2}} K_n(t) - \bar{B}_n(t)|/q(t) = o_P(1)$$

for any $q \in \mathcal{Q}^*$ and such that (7.0.3) holds. Consequently $\sup_{0 < t < 1} |\lambda^{-\frac{1}{2}} K_n(t)|/q(t)$ converges in distribution to wherever $\sup_{Y_1/n \leq t < Y_n/n} |B_n(t)|/q(t)$ does, as $n \rightarrow \infty$. Hence, for the "if" part of (b) it suffices to show that, when $q \in \mathcal{Q}^*$ and $I^*(q, c) < \infty$ for some $c > 0$, then, as $n \rightarrow \infty$

$$\sup_{Y_1/n \leq t < Y_n/n} |B_n(t)|/q(t) \xrightarrow{D} \sup_{0 < t < 1} |B(t)|/q(t),$$

where $\{B(t), 0 \leq t \leq 1\}$ is a Brownian bridge. Since $Y_1/n \xrightarrow{a.s.} 0$, $Y_n/n \xrightarrow{a.s.} 1$ and $\{B_n(t), 0 \leq t \leq 1\}$ is a Brownian bridge for each $n \geq 1$, this can be done along the lines of the proof of Lemma 4.2.2 of Cs-Cs-H-M (1986).

Conversely, with $q \in \mathcal{Q}^*$ we assume (7.1.15). We take Skorohod-Dudley-Wichura versions $\hat{K}_n(t)$ and $\hat{B}_n(t)$ of $\lambda^{-\frac{1}{2}} K_n(t)$ and $B(t)$ respectively (cf. Theorem A.2.1). Then, as $n \rightarrow \infty$

$$\left| \sup_{0 < t < 1} |\hat{K}_n(t)|/q(t) - \sup_{0 < t < 1} |\hat{B}_n(t)|/q(t) \right| \xrightarrow{a.s.} o(1),$$

and consequently

$$P \left\{ \sup_{0 < t < 1} |\hat{B}_n(t)|/q(t) < \infty \right\} = 1.$$

Hence

$$P \left\{ \limsup_{t \downarrow 0} |\dot{B}(t)|/q(t) < \infty \right\} = 1 = P \left\{ \limsup_{t \uparrow 1} |\dot{B}(t)|/q(t) < \infty \right\},$$

and so, by Theorem 5.1.B, $q \in \mathcal{Q}^*$ must be such that $I^*(q, c) < \infty$ for some $c > 0$.

(c) By Proposition 7.1.1(a) with $q \equiv 1$, there exists a sequence of Brownian bridges $\{B_n(t), 0 \leq t \leq 1\}$ such that, as $n \rightarrow \infty$,

$$\sup_{0 \leq t \leq 1} |\lambda^{-\frac{1}{2}} K_n(t) - B_n(t)| = o_P(1).$$

Hence, with any $q: (0, 1) \rightarrow (0, \infty)$ such that (7.2.1) holds we have, as $n \rightarrow \infty$,

$$\sup_{\delta \leq t \leq 1-\delta} |\lambda^{-\frac{1}{2}} K_n(t) - B_n(t)|/q(t) = o_P(1)$$

for any $\delta \in (0, \frac{1}{2})$, which implies (c).

Remark 7.1.2. We should note that part (a) of Proposition 7.1.1 gives the limiting distribution of any continuous in sup-norm functional of $\lambda^{-\frac{1}{2}} K_n(t)/q(t)$. However, for the sup-functional, part (b) of Proposition 7.1.2 gives the desired convergence in distribution result for a bigger class of weight functions. For example part (b) of Proposition 7.1.2 holds true with

$$q(t) = \left(t(1-t) \log \log \frac{1}{t(1-t)} \right)^{\frac{1}{2}}, \quad 0 < t < 1.$$

The “if” part of Proposition 7.1.1(a) and that of Proposition 7.1.2(b) is due to, and proved also by Csörgő and Horváth (1992).

Corollary 7.1.1. *We assume that H holds and let $\{B(t), 0 \leq t \leq 1\}$ be a Brownian bridge. Then Proposition 7.1.2 holds with $\lambda^{-\frac{1}{2}}$ replaced by $\left(\frac{n}{N(n)} \right)^{\frac{1}{2}}$.*

As we noted already, $I^*(q, c) < \infty$ for some $c > 0$ implies (7.0.3). Hence, when $q \in \mathcal{Q}^*$ and $I^*(q, c) < \infty$ for some $c > 0$ then $\left(\frac{n}{N(n)}\right)^{\frac{1}{2}} K_n(t)/q(t)$ can be also written as

$$\left(\frac{n}{N(n)}\right)^{\frac{1}{2}} \frac{K_n(t)}{(t(1-t))^{\frac{1}{2}}} / g(t),$$

where $\lim_{t \downarrow 0} g(t) = \lim_{t \uparrow 1} g(t) = \infty$, the weighted versions of the Kendall-Kendall (1980) pontograms, as mentioned already in Section 7.0.

7.2. L_p -norms of weighted pontograms

We noted already that the weak convergence of $K_n(t)$ in weighted sup-norm metrics does not give the convergence in distribution of the supremum of the weighted $K_n(t)$ process for all possible weight functions of interest. Naturally then, L_p -functionals of such processes should be studied also directly, on their own. We present a necessary and sufficient condition for the weak convergence of $K_n(t)$ process to a Brownian bridge in weighted L_p -norms. In this case, unlike in supremum norm, weighted L_p -functionals of the $K_n(t)$ process converge in distribution to the corresponding functionals of a Brownian bridge for the same class of weight functions.

Proposition 7.2.1. *We assume that H holds and $0 < p < \infty$. Let q be positive, i.e. such that $\inf_{\delta \leq t \leq 1-\delta} q(t) > 0$ for all $\delta \in (0, 1/2)$.*

(a) *There is a sequence of Brownian bridges $\{B_n(t), 0 \leq t \leq 1\}$ such that, as $n \rightarrow \infty$*

$$(7.2.1) \quad \int_0^1 |\lambda^{-\frac{1}{2}} K_n(t) - B_n(t)|^p / q(t) dt = o_P(1)$$

if and only if

$$(7.2.2) \quad \int_0^1 (t(1-t))^{\frac{p}{2}} / q(t) dt < \infty.$$

(b) Let $\{B(t), 0 \leq t \leq 1\}$ be a Brownian bridge. Then, as $n \rightarrow \infty$

$$(7.2.3) \quad \int_0^1 |\lambda^{-\frac{1}{2}} K_n(t)|^p / q(t) dt \xrightarrow{P} \int_0^1 |B(t)|^p / q(t) dt$$

if and only if (7.2.2) holds.

This result is an analog of a weighted L_p approximation of the empirical process due to Csörgő, Horváth and Shao (1991). Our proof follows the lines of that of their Theorem 1.1 and is given here for the convenience of the reader. In the proof we will use the following result of Csörgő, Horváth and Shao (1991).

Theorem 7.2.A. Let $\{B(t), 0 \leq t \leq 1\}$ be a Brownian bridge, $0 < p < \infty$ and $-\infty < \beta < \infty$. Then

$$\int_0^1 |B(t) + \beta t(1-t)|^p / q(t) dt < \infty \quad a.s.$$

if and only if (7.2.2) holds.

Proof of Proposition 7.2.1. (a) First we assume that (7.2.2) holds. By (7.1.2) (with $q \equiv 1$ there) we have, as $n \rightarrow \infty$

$$(7.2.4) \quad \int_{\delta}^{1-\delta} |\lambda^{-\frac{1}{2}} K_n(t) - B_n(t)|^p / q(t) dt = o_P(1)$$

for all $\delta \in (0, \frac{1}{2})$. Using (7.1.5) with $\nu = 0$ gives

$$\begin{aligned} & \int_{\mu/n}^{\delta} |\lambda^{-\frac{1}{2}} K_n(t) - B_n(t)|^p / q(t) dt \\ & \leq \left\{ \sup_{\mu/n \leq t \leq \delta} |\lambda^{-\frac{1}{2}} K_n(t) - B_n(t)| / (t(1-t))^{\frac{1}{2}} \right\}^p \int_{\mu/n}^{\delta} (t(1-t))^{\frac{p}{2}} / q(t) dt \\ & \leq \left\{ \sup_{\mu/n \leq t \leq 1-\mu/n} |\lambda^{-\frac{1}{2}} K_n(t) - B_n(t)| / (t(1-t))^{\frac{1}{2}} \right\}^p \int_0^{\delta} (t(1-t))^{\frac{p}{2}} / q(t) dt \\ (7.2.5) \quad & = O_P(1) \int_0^{\delta} (t(1-t))^{\frac{p}{2}} / q(t) dt \end{aligned}$$

and similarly

$$\begin{aligned}
 & \int_{1-\delta}^{1-\mu/n} |\lambda^{-\frac{1}{2}} K_n(t) - B_n(t)|^p / q(t) \\
 (7.2.6) \quad & \leq O_P(1) \int_{1-\delta}^1 (t(1-t))^{\frac{p}{2}} / q(t) dt
 \end{aligned}$$

for all $0 < \mu < \infty$. By combining (7.2.4), (7.2.5), (7.2.6) and taking $\delta > 0$ arbitrarily small, we have

$$(7.2.7) \quad \int_{\mu/n}^{1-\mu/n} |\lambda^{-\frac{1}{2}} K_n(t) - B_n(t)|^p / q(t) dt = o_P(1)$$

as $n \rightarrow \infty$ for all $0 < \mu < \infty$. By Theorem 7.2.A we have, as $n \rightarrow \infty$

$$(7.2.8) \quad \int_0^{\mu/n} |B_n(t)|^p / q(t) dt = o_P(1)$$

and also

$$(7.2.9) \quad \int_{1-\mu/n}^1 |B_n(t)|^p / q(t) dt = o_P(1).$$

Again, let Y_1 denote the time of the first jump of $N(x)$ and Y_n be the time of the last jump of $N(x)$ before n . Then

$$\begin{aligned}
 & \int_0^{Y_1/n} |\lambda^{-\frac{1}{2}} K_n(t)|^p / q(t) dt \\
 & = \int_0^{Y_1/n} |(\lambda n)^{-\frac{1}{2}} t N(n)|^p / q(t) dt \\
 (7.2.10) \quad & \leq \lambda^{-\frac{p}{2}} (Y_1)^{\frac{p}{2}} (N(n)/n)^p \int_0^{Y_1/n} t^{\frac{p}{2}} / q(t) dt \\
 & = o_P(1),
 \end{aligned}$$

and also

$$\begin{aligned}
 & \int_{Y_n/n}^1 |\lambda^{-\frac{1}{2}} K_n(t)|^p / q(t) dt \\
 (7.2.11) \quad & = \int_{Y_n/n}^1 |(\lambda n)^{-\frac{1}{2}} (1-t) N(n)|^p / q(t) dt \\
 & \leq \lambda^{-\frac{p}{2}} (n - Y_n)^{\frac{p}{2}} (N(n)/n)^p \int_{Y_n/n}^1 (1-t)^{\frac{p}{2}} / q(t) dt \\
 & = o_P(1),
 \end{aligned}$$

due to the fact that Y_1 and $n - Y_n$ are exponential random variables, $N(n)/n \xrightarrow{a.s.} \lambda$, $Y_1/n \xrightarrow{a.s.} 0$ and $Y_n/n \xrightarrow{a.s.} 1$. We note also that for each $\epsilon > 0$ there is a $\mu > 0$ such that $P\{Y_1 < \mu\} \leq \epsilon$ and $P\{n - Y_n < \mu\} \leq \epsilon$. Consequently (7.2.7)–(7.2.11) yield (7.2.1).

Next we assume that (7.2.1) holds true. Hence

$$\begin{aligned}
 (7.2.12) \quad & \int_0^{Y_1/n} |\lambda^{-\frac{1}{2}} K_n(t) - B_n(t)|^p / q(t) dt \\
 &= \int_0^{Y_1/n} |B_n(t) + n^{\frac{1}{2}} t \lambda^{-\frac{1}{2}} N(n)/n|^p / q(t) dt \\
 &= o_P(1)
 \end{aligned}$$

and also

$$\begin{aligned}
 (7.2.13) \quad & \int_{Y_n/n}^1 |\lambda^{-\frac{1}{2}} K_n(t) - B_n(t)|^p / q(t) dt \\
 &= \int_{Y_n/n}^1 |B_n(t) + n^{\frac{1}{2}} (1-t) \lambda^{-\frac{1}{2}} N(n)/n|^p / q(t) dt \\
 &= o_P(1),
 \end{aligned}$$

as $n \rightarrow \infty$.

Since $Y_1/n \xrightarrow{a.s.} 0$, $Y_n/n \xrightarrow{a.s.} 1$ and $N(n)/n \xrightarrow{a.s.} \lambda$, there are constants n_0, δ, β and ϵ such that

$$P \left\{ \int_0^\delta |B_{n_0}(t) + \beta t|^p / q(t) dt > \epsilon \right\} \leq \frac{1}{2}$$

and

$$P \left\{ \int_{1-\delta}^1 |B_{n_0}(t) + \beta(1-t)|^p / q(t) dt > \epsilon \right\} \leq \frac{1}{2}.$$

Consequently

$$P \left\{ \int_0^1 |B_{n_0}(t) + \beta t(1-t)|^p / q(t) dt > M \right\} \leq \frac{1}{2}$$

with some constant M . The distribution of B_{n_0} does not depend on n_0 and, therefore we have

$$(7.2.14) \quad P \left\{ \int_0^1 |B(t) + \beta t(1-t)|^p / q(t) dt > M \right\} \leq \frac{1}{2}$$

for any Brownian bridge $\{B(t), 0 \leq t \leq 1\}$. The 0-1 law for the Brownian bridge implies

$$P \left\{ \int_0^1 |B(t) + \beta t(1-t)|^p / q(t) dt < \infty \right\} = 0 \text{ or } 1,$$

hence, by Theorem 7.2.A, (7.2.14) yields (7.2.2).

(b) First we assume (7.2.2) to hold. Since q is such that (7.1.1) holds, by Proposition 7.1.1(a) we have

$$\sup_{\delta < t < 1-\delta} |\lambda^{-\frac{1}{2}} K_n(t) - B_n(t)| / q(t) = o_P(1)$$

for any $\delta \in (0, \frac{1}{2})$, and consequently also

$$(7.2.15) \quad \int_{\delta}^{1-\delta} |\lambda^{-\frac{1}{2}} K_n(t)|^p / q(t) dt \xrightarrow{P} \int_{\delta}^{1-\delta} |B(t)|^p / q(t) dt,$$

where $\{B(t), 0 \leq t \leq 1\}$ is a Brownian bridge. By Theorem 7.2.A we have

$$(7.2.16) \quad \lim_{\delta \rightarrow 0} \int_0^{\delta} |B(t)|^p / q(t) dt = 0 \quad a.s.$$

and

$$(7.2.17) \quad \lim_{\delta \rightarrow 0} \int_{1-\delta}^1 |B(t)|^p / q(t) dt = 0 \quad a.s.$$

Since

$$\int_0^{\delta} |\lambda^{-\frac{1}{2}} K_n(t)|^p / q(t) dt \leq 2^p \int_0^{\delta} |B_n(t)|^p / q(t) dt + 2^p \int_0^{\delta} |\lambda^{-\frac{1}{2}} K_n(t) - B_n(t)|^p / q(t) dt,$$

we use (7.2.1) and (7.2.16) to get

$$(7.2.18) \quad \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} P \left\{ \int_0^{\delta} |\lambda^{-\frac{1}{2}} K_n(t)|^p / q(t) dt > \epsilon \right\} = 0,$$

and, similarly, in the neighbourhood of 1 we have

$$(7.2.19) \quad \lim_{\delta \rightarrow 0} \limsup_{n \rightarrow \infty} P \left\{ \int_{1-\delta}^1 |\lambda^{-\frac{1}{2}} K_n(t)|^p / q(t) dt > \epsilon \right\} = 0,$$

for all $\epsilon > 0$. Combining (7.2.15), (7.2.18) and (7.2.19) yields (7.2.3).

Next we assume (7.2.3) to hold. Then the limit is finite almost surely and Theorem 7.2.A implies (7.2.2).

As in the case of the supremum norm, different types of normalization can be used to get a nondegenerate limiting distribution. Csörgő and Horváth (1987c) prove a number of results considering L_2 norms of pontograms, in particular the L_2 analog of Theorem 7.0.A. These can be generalized for L_p , $1 \leq p < \infty$, functionals, using the method of proof of Csörgő and Horváth (1988b, 1992), where they study weighted uniform and quantile processes.

Proposition 7.2.2. *Let $1 \leq p < \infty$ and $N(0,1)$ be a standard normal variable. We assume that H holds. Let*

$$(7.2.20) \quad M = M(p) = (2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} |x|^p \exp(-x^2/2) dx$$

and

$$(7.2.21) \quad D = D(p) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |xy|^p \times \left\{ \frac{1}{2\pi(1 - e^{-2|u|})^{\frac{1}{2}}} \exp\left(-\frac{x^2 + y^2 - 2e^{-|u|}xy}{2(1 - e^{-2|u|})^{\frac{1}{2}}}\right) - \phi(x)\phi(y) \right\} dx dy du.$$

(a) *We assume $0 < a_n < 1 - b_n < 1$ and*

$$\frac{(1 - b_n)(1 - a_n)}{a_n b_n} \rightarrow \infty \quad \text{as } n \rightarrow \infty.$$

Then we have, as $n \rightarrow \infty$,

(7.2.22)

$$(2D \log c_n^*)^{-\frac{1}{2}} \left(\int_{a_n}^{1-b_n} |\lambda^{-\frac{1}{2}} K_n(t)|^p / (t(1-t))^{1+\frac{p}{2}} dt - m \log c_n^* \right) \xrightarrow{\mathcal{D}} N(0, 1),$$

where

$$c_n^* = \frac{(1 - \max(b_n, 1/n))(1 - \max(a_n, 1/n))}{\max(a_n, 1/n) \max(b_n, 1/n)}.$$

(b) We assume $a_n \rightarrow 0$ and $na_n \rightarrow \infty$ as $n \rightarrow \infty$. Then we have, as $n \rightarrow \infty$

$$(4D \log \frac{1}{a_n})^{-\frac{1}{2}} \left(\int_{a_n}^{1-a_n} |\lambda^{-\frac{1}{2}} K_n(t)|^p / (t(1-t))^{1+\frac{p}{2}} dt - 2M \log \frac{1}{a_n} \right) \xrightarrow{\mathcal{D}} N(0, 1),$$

$$(2D \log(na_n))^{-\frac{1}{2}} \left(\int_0^{a_n} |\lambda^{-\frac{1}{2}} K_n(t)|^p / (t(1-t))^{1+\frac{p}{2}} dt - M \log(na_n) \right) \xrightarrow{\mathcal{D}} N(0, 1),$$

and

$$(2D \log(na_n))^{-\frac{1}{2}} \left(\int_{1-a_n}^1 |\lambda^{-\frac{1}{2}} K_n(t)|^p / (t(1-t))^{1+\frac{p}{2}} dt - M \log(na_n) \right) \xrightarrow{\mathcal{D}} N(0, 1).$$

(c) We have, as $n \rightarrow \infty$

$$(4D \log n)^{-\frac{1}{2}} \left(\int_0^1 |\lambda^{-\frac{1}{2}} K_n(t)|^p / (t(1-t))^{1+\frac{p}{2}} dt - 2M \log n \right) \xrightarrow{\mathcal{D}} N(0, 1).$$

The proof is based on the following two results (cf. Csörgő and Horváth, 1992).

Let $\{V(t), -\infty < t < \infty\}$ be an Ornstein-Uhlenbeck process, i.e. a Gaussian process with $EV(t) = 0$ and $EV(t)V(s) = \exp(-|t-s|)$.

Lemma 7.2.A. We assume that $0 < p < \infty$. We have, as $T \rightarrow \infty$,

$$\left(\int_0^T |V(t)|^p dt - mT \right) / (DT)^{\frac{1}{2}} \xrightarrow{D} N(0, 1),$$

where $N(0, 1)$ is a standard normal random variable and M, D are as in Proposition 7.2.2.

Lemma 7.2.B We assume that $0 < p < \infty$. Let $0 < a_n < 1 - b_n < 1$ and

$$c_n = \frac{(1 - b_n)(1 - a_n)}{a_n b_n} \rightarrow \infty \quad \text{as } n \rightarrow \infty.$$

Then we have, as $n \rightarrow \infty$

$$\left(\int_{a_n}^{1-b_n} |B(t)|^p / (t(1-t))^{1+\frac{p}{2}} dt - M \log c_n \right) / (2D \log c_n)^{\frac{1}{2}} \xrightarrow{D} N(0, 1),$$

where $\{B(t), 0 \leq t \leq 1\}$ is a Brownian bridge, $N(0, 1)$ is a standard normal random variable and M, D are as in Proposition 7.2.2.

Proof of Proposition 7.2.2. By (7.1.5) we can define a sequence of Brownian bridges $\{B_n(t), 0 \leq t \leq 1\}$ such that with $0 < \nu \leq \frac{1}{2}$ we have as $n \rightarrow \infty$

$$\begin{aligned} & \int_{\mu/n}^{1-\mu/n} |\lambda^{-\frac{1}{2}} K_n(t) - B_n(t)|^p / (t(1-t))^{1+\frac{p}{2}} dt \\ & \leq \left(\sup_{\mu/n \leq t \leq 1-\mu/n} |\lambda^{-\frac{1}{2}} K_n(t) - B_n(t)| / (t(1-t))^\nu \right)^p \\ & \quad \times \int_{\mu/n}^{1-\mu/n} (t(1-t))^{\nu p - 1 - \frac{p}{2}} dt \\ & = O_P(1) n^{(\nu - \frac{1}{2})p} \int_{\mu/n}^{1-\mu/n} (t(1-t))^{-1 + (\nu - \frac{1}{2})p} dt \\ (7.2.27) \quad & = O_P(1), \end{aligned}$$

for all $0 < \mu < \infty$. Using (7.2.27) and the method of proof of Theorem 2.3 of Csörgő and Horváth (1988b), we have

$$(7.2.28) \quad \int_{\mu/n}^{1-\mu/n} \left| |\lambda^{-\frac{1}{2}} K_n(t)|^p - |B_n(t)|^p \right| / (t(1-t))^{1+\frac{p}{2}} dt = O_P(1)$$

for all $0 < \mu < \infty$.

Namely, we note that if $p \geq 1$, then we have

$$\left| |a|^p - |b|^p \right| \leq p2^{p-1}|a-b|^p + p2^{p-1}|b|^{p-1}|a-b|.$$

Hence

$$\begin{aligned} & \int_{\mu/n}^{1-\mu/n} \left| |\lambda^{-\frac{1}{2}} K_n(t)|^p - |B_n(t)|^p \right| / (t(1-t))^{1+\frac{p}{2}} dt \\ & \leq p2^p \int_{\mu/n}^{1-\mu/n} |\lambda^{-\frac{1}{2}} K_n(t) - B_n(t)|^p / (t(1-t))^{1+\frac{p}{2}} dt \\ & + p2^{p-1} \int_{\mu/n}^{1-\mu/n} |\lambda^{-\frac{1}{2}} K_n(t) - B_n(t)| |B_n(t)|^{p-1} / (t(1-t))^{1+\frac{p}{2}} dt \\ & = O_P(1), \end{aligned}$$

due to (7.2.27).

Let Y_1 be the first jump point of $N(t)$. We have

$$\begin{aligned} & \int_0^{Y_1/n} |\lambda^{-\frac{1}{2}} K_n(t)|^p / (t(1-t))^{1+\frac{p}{2}} dt \\ & = \int_0^{Y_1/n} \left(\lambda^{-\frac{1}{2}} n^{-\frac{1}{2}} t N(n) \right)^p / (t(1-t))^{1+\frac{p}{2}} dt \\ & = \lambda^{-\frac{p}{2}} n^{-\frac{p}{2}} (N(n))^p \int_0^{Y_1/n} (t(1-t))^{\frac{p}{2}-1} dt \\ & \leq \lambda^{-\frac{p}{2}} (N(n)/n)^p n^{\frac{p}{2}} (Y_1/n)^{\frac{p}{2}} \\ (7.2.29) \quad & = O_P(1), \end{aligned}$$

on account of $N(n)/n \xrightarrow{d} \lambda$ and Y_1 being exponential random variable. Similarly to (7.2.29) we obtain

$$(7.2.30) \quad \int_{Y_n/n}^1 |\lambda^{-\frac{1}{2}} K_n(t)|^p / (t(1-t))^{1+\frac{p}{2}} dt = O_P(1),$$

where Y_n is the time of the last jump of $N(t)$ before n .

(a) By (7.2.28), (7.2.29) and (7.2.30) there exist μ and C for all $\epsilon > 0$ such that

$$\limsup_{n \rightarrow \infty} P \left\{ \left| \int_{a_n}^{1-b_n} |\lambda^{-\frac{1}{2}} K_n(t)|^p / (t(1-t))^{1+\frac{p}{2}} dt - \int_{\max(a_n, \mu/n)}^{1-\max(b_n, \mu/n)} |B_n(t)|^p / (t(1-t))^{1+\frac{p}{2}} dt \right| > C \right\} \leq \epsilon.$$

Now we apply Lemma 7.2.B to conclude the result.

(b) Using the assumption $na_n \rightarrow \infty$ as $n \rightarrow \infty$, (7.2.23) follows from (7.2.22) by taking $b_n = a_n$. In order to prove (7.2.24) we use Lemma 7.2.B. Namely, we have for all $0 < \mu < \infty$

$$(7.2.31) \quad \left(\int_{\mu/n}^{a_n} |B_n(t)|^p / (t(1-t))^{1+\frac{p}{2}} dt - M \log(na_n) \right) / \left(2D \log(na_n) \right)^{\frac{1}{2}} \xrightarrow{D} N(0, 1),$$

where M and D are defined by (7.2.20) and (7.2.21), respectively. Combining (7.2.28), (7.2.29) and (7.2.31) gives (7.2.24). We obtain (7.2.25) similarly.

(c) By Lemma 7.2.B we have for all $0 < \mu < \infty$

$$\left(\int_{\mu/n}^{1-\mu/n} |B_n(t)|^p / (t(1-t))^{1+\frac{p}{2}} dt - 2M \log n \right) / \left(4D \log n \right)^{\frac{1}{2}} \xrightarrow{D} N(0, 1),$$

which, together with (7.2.28), (7.2.29) and (7.2.30), yields (7.2.26).

For L_p , $1 \leq p < \infty$, norms of generalized pontograms we refer to Eastwood (1990).

7.3. Contiguous alternatives and a changepoint problem

In this section we consider the situation when the intensity function $\lambda(x)$, $x \geq 0$, of a nonhomogeneous Poisson process $\{N(x), x \geq 0\}$ satisfies the condition

$$A: \lambda(x) = \lambda + \xi([x]/n)/n^{\frac{1}{2}}, \quad 0 \leq x \leq n$$

for each $n \geq 1$, where $\lambda > 0$ and $\xi^2(t)$ is a Riemann integrable function. We study the asymptotic behaviour of the process $\{K_n(t), 0 < t < 1\}$ as defined in (7.0.1), and its weighted versions with weight functions q as in Section 7.1. Under the same assumption about $\lambda(x)$ we obtain also the limiting distribution of the weighted pontogram $\left\{ \left(\frac{n}{N(n)} \right)^{\frac{1}{2}} K_n(t)/q(t), 0 < t < 1 \right\}$.

Under the sequence of contiguous alternatives of A we have the following result for the Poisson bridge.

We recall that $\mathcal{Q}^*$ is the class of functions $q(t) : (0, 1) \rightarrow (0, \infty)$ which are nondecreasing near zero and nonincreasing near one and such that $\inf_{\delta \leq t \leq 1-\delta} q(t) > 0$ for all $\delta \in (0, 1/2)$. As before, let

$$E^*(q, c) = \int_0^1 (t(1-t))^{-\frac{1}{2}} q(t) \exp(-c(t(1-t))^{-1} q^2(t)) dt$$

and

$$I^*(q, c) = \int_0^1 (t(1-t))^{-1} \exp(-c(t(1-t))^{-1} q^2(t)), \quad 0 < c < \infty.$$

Theorem 7.3.1. *Let $\{B(t), 0 \leq t \leq 1\}$ be a Brownian bridge, and we assume that A holds.*

(a) *Let $q \in \mathcal{Q}^*$ and $I^*(q, c) < \infty$ for all $c > 0$. Then, as $n \rightarrow \infty$*

$$(7.3.1) \quad \lambda^{-\frac{1}{2}} K_n(t)/q(t) \xrightarrow{\mathcal{D}} \left(B(t) + \lambda^{-\frac{1}{2}} \left(\int_0^t \xi(s) ds - t \int_0^1 \xi(s) ds \right) \right) / q(t)$$

in $D[0, 1]$.

(b) *Let $q : (0, 1) \rightarrow (0, \infty)$ be any function such that (7.1.1) holds. Then, as $n \rightarrow \infty$ (7.3.1) holds in $D[\delta, 1-\delta]$ for any $\delta \in (0, \frac{1}{2})$.*

Remark 7.3.1. According to Theorem 3.1.C, part (a) of Theorem 7.3.1 can be stated equivalently in terms of the integral $E^*(q, c)$ (cf. Remark 7.1.1).

The proof of Theorem 7.3.1 will be given in the end of this section.

Remark 7.3.2. We note that by taking $q \equiv 1$ in Theorem 7.3.1(a) we have also the limiting distribution of the process $K_n(t)$ under the sequence of contiguous alternatives of A .

Considering now the pontogram, the following result holds.

Corollary 7.3.1. *Theorem 7.3.1 remains true when replacing*

$$\{\lambda^{-\frac{1}{2}} K_n(t)/q(t), 0 < t < 1\}$$

by the weighted pontogram process

$$\left\{ \left(\frac{n}{N(n)} \right)^{\frac{1}{2}} K_n(t)/q(t), 0 < t < 1 \right\}.$$

As a particular case, we consider the so-called changepoint alternative

A^η : there exists $\eta \in (0, 1)$ such that

$$\lambda(x) = \lambda \quad \text{for } 0 \leq x < n\eta,$$

$$\lambda(x) = \lambda + \rho/n^{\frac{1}{2}} \quad \text{for } n\eta \leq x \leq n,$$

where $\lambda > 0$ and $\rho \neq 0$.

Corollary 7.3.2. *We assume that A^η holds.*

(a) *Let $q \in Q^*$ and $I^*(q, c) < \infty$ for all $c > 0$. Then as $n \rightarrow \infty$*

$$(7.3.2) \quad \left(\frac{n}{N(n)} \right)^{\frac{1}{2}} K_n(t)/q(t) \\ \xrightarrow{D} \left(B(t) + \lambda^{-\frac{1}{2}} [-t\rho(1-\eta)1\{t < \eta\} - (1-t)\rho\eta 1\{t \geq \eta\}] \right) / q(t).$$

(b) *Let $q: (0, 1) \rightarrow (0, \infty)$ be any function such that (7.1.1) holds. Then, as $n \rightarrow \infty$, (7.3.2) holds in $D[\delta, 1-\delta]$ for any $\delta \in (0, \frac{1}{2})$.*

Corollaries 7.3.1 and 7.3.2 give the limiting distribution for any continuous in sup-norm functional of the weighted pontogram under A and A'' respectively. For example, by Corollary 7.3.2 we have that under A'' ,

$$\sup_{0 < t < 1} \left(\frac{n}{N(n)} \right)^{\frac{1}{2}} K_n(t)/q(t) \xrightarrow{D} \sup_{0 < t < 1} \left(B(t) + \lambda^{-\frac{1}{2}} [-t\rho(1-\eta)1\{t < \eta\} - (1-t)\rho\eta 1\{t \geq \eta\}] \right) / q(t),$$

where $\{B(t), 0 \leq t \leq 1\}$ is a Brownian bridge.

Proof of Theorem 7.3.1. Let $\{N(x), x \geq 0\}$ be a nonhomogeneous Poisson process with intensity function $\lambda(x)$, $x \geq 0$. We note that when assuming H

$$(7.3.3) \quad Z_i = N(i) - N(i-1), \quad i \geq 1$$

are independent, identically distributed Poisson random variables with mean λ and, under the conditions of the alternative A , Z_i , $1 \leq i \leq n$, are independent, Poisson random variables with respective means equal to $\lambda + \xi(i/n)/n^{\frac{1}{2}}$.

Let

$$K_n(t) = \begin{cases} n^{-\frac{1}{2}} \left(\sum_{i=1}^{[(n+1)t]} Z_i - \frac{[(n+1)t]}{n} \sum_{i=1}^n Z_i \right), & 0 \leq t < 1 \\ 0, & t = 1. \end{cases}$$

Let $q \in Q^*$ and $I^*(q, c) < \infty$ for all $c > 0$. It follows from Corollary 5.2.1 that assuming H

$$(7.3.4) \quad K_n(t)/\lambda^{\frac{1}{2}}q(t) \xrightarrow{D} B(t)/q(t) \quad \text{in } D[0, 1],$$

where $\{B(t), 0 \leq t \leq 1\}$ is a Brownian bridge.

Now we derive the limiting distribution of K_n under A .

Let

$$f(x, \theta) = e^{-\theta} \frac{\theta^x}{x!}, \quad x = 0, 1, \dots$$

and for each Z_i , $1 \leq i \leq n$, we consider the likelihood ratio

$$\Lambda_n(Z_i) := \frac{f(Z_i, \lambda + \xi(i/n)/n^{\frac{1}{2}})}{f(Z_i, \lambda)} = e^{-\xi(i/n)/n^{\frac{1}{2}}} (1 + \xi(i/n)/\lambda n^{\frac{1}{2}})^{Z_i}.$$

We define the centered log-likelihood ratio process $L_n(\cdot)$ by

$$L_n(t) = \sum_{i=1}^{nt} \{\log \Lambda_n(Z_i) - E \log \Lambda_n(Z_i)\}, \quad 0 \leq t \leq 1.$$

Let (B, L) be the two-dimensional Gaussian process, where B is a Brownian bridge, L is a Gaussian process with mean zero and covariance function

$$(7.3.5) \quad \text{Cov } L(t_1)L(t_2) = EL(t_1)L(t_2) = \lambda^{-1} \int_0^{t_1 \wedge t_2} \xi^2(s) ds,$$

and the covariance function of B and L is

$$(7.3.6) \quad \text{Cov } B(t_1)L(t_2) = EB(t_1)L(t_2) = \lambda^{-\frac{1}{2}} \left(\int_0^{t_1 \wedge t_2} \xi(s) ds - t_1 \int_0^{t_2} \xi(s) ds \right).$$

We note that assuming H we have, as $n \rightarrow \infty$

$$\sum_{i=1}^n \log \Lambda_n(Z_i) \rightarrow N \left(-\frac{1}{2} \sigma^2, \sigma^2 \right)$$

with

$$\sigma^2 = \frac{1}{\lambda} \int_0^1 \xi^2(s) ds,$$

which, by Le Cam's first lemma (cf. Corollary 2.1.A) implies contiguity of the sequence of probability measures of $(Z_1, \dots, Z_n)$ under H to those under A .

Clearly, the process $\lambda^{-\frac{1}{2}} K_n(t)$ converges to Brownian bridge and $L_n(t)$ being a process of partial sums of independent random variables with mean zero and finite variances for which the Lindeberg's condition holds, converges to L (cf.

(7.3.5)). Because of the structure of both processes $\lambda^{-\frac{1}{2}}\mathcal{K}_n(t)$ and $L_n(t)$ it is easy to compute their covariance function, which is seen to converge to $\text{Cov } B(t_1)L(t_2)$ (cf. (7.3.6)). Using the Cramér-Wold theorem (cf. Billingsley (1968)) we conclude that assuming H

$$(7.3.7) \quad (\lambda^{-\frac{1}{2}}\mathcal{K}_n(t), L_n(t)) \xrightarrow{\mathcal{D}} (B(t), L(t)) \text{ in } D[0, 1].$$

We note that since $\xi^2(t)$ is a Riemann integrable function, $\text{Cov} B(t_1)L(t_2) < \infty$. In fact, we have

$$\begin{aligned} \left| \int_0^t \xi(s) ds \right| &\leq \left(\int_0^t \xi^2(s) ds \right)^{1/2} \left(\int_0^t ds \right)^{1/2} \\ &= \left(\int_0^t \xi^2(s) ds \right)^{1/2} t^{1/2}. \end{aligned}$$

Similarly

$$\left| \int_t^1 \xi(s) ds \right| \leq \left(\int_t^1 \xi^2(s) ds \right)^{1/2} (1-t)^{1/2}.$$

Hence

$$\begin{aligned} \int_0^t \xi(s) ds - t \int_0^1 \xi(s) ds &= (1-t) \int_0^t \xi(s) ds - t \int_t^1 \xi(s) ds \\ &\leq (1-t)t^{1/2} \left(\int_0^t \xi^2(s) ds \right)^{1/2} - t(1-t)^{1/2} \left(\int_t^1 \xi^2(s) ds \right)^{1/2}. \end{aligned}$$

Let $q : (0, 1) \rightarrow (0, \infty)$ be such that (7.0.3) and (7.1.1) hold. Then

$$\begin{aligned} &\sup_{0 < t < 1} \left| \int_0^t \xi(s) ds - t \int_0^1 \xi(s) ds \right| / q(t) \\ &\leq \sup_{0 < t < 1} (1-t)^{1/2} \left(\int_0^t \xi^2(s) ds \right)^{1/2} - t^{1/2} \left(\int_0^1 \xi^2(s) ds \right)^{1/2} \sup_{0 < t < 1} \frac{(t(1-t))^{1/2}}{q(t)} < \infty. \end{aligned}$$

Let $q \in \mathcal{Q}^*$. Since $I^*(q, c) < \infty$ for some $c > 0$ implies (7.0.3) and

$$E \frac{\mathcal{K}_n(t_1)}{\lambda^{\frac{1}{2}} q(t_1)} L_n(t_2) = \frac{1}{q(t_1)} E \lambda^{-\frac{1}{2}} \mathcal{K}_n(t_1) L_n(t_2) \rightarrow \frac{1}{q(t_1)} \text{Cov } B(t_1)L(t_2),$$

we use (7.3.4) and (7.3.7) to conclude that, under H

$$(7.3.8) \quad \left(\lambda^{-\frac{1}{2}} \mathcal{K}_n(t)/q(t), L_n(t) \right) \xrightarrow{\mathcal{D}} (B(t)/q(t), L(t)) \quad \text{in } D[0, 1].$$

Now we assume that A holds. By Le Cam's third lemma (cf. Lemma 2.1.F), all finite dimensional distributions of $\lambda^{-\frac{1}{2}} \mathcal{K}_n(t)/q(t)$ converge to those of $(B(t) + b(t))/q(t)$, where $b(t) = \text{Cov } B(t)L(1)$. Since contiguity preserves tightness, it follows that, assuming A

$$(7.3.9) \quad \lambda^{-\frac{1}{2}} \mathcal{K}_n(t)/q(t) \xrightarrow{\mathcal{D}} \left(B(t) + \lambda^{-\frac{1}{2}} \left(\int_0^t \xi(s) ds - t \int_0^1 \xi(s) ds \right) \right) / q(t) \quad \text{in } D[0, 1]$$

We note here that (7.3.9) can be also deduced from Theorem 5.4.1, or more general result, Theorem 6.3.2 which concerns the asymptotic behaviour of U -statistic type processes under the sequence of contiguous alternatives parameterized in a more general way.

We note that for the Poisson process $\{N(x), x > 0\}$ with intensity parameter $\lambda > 0$

$$(7.3.10) \quad \sup_{k \leq x < k+1} |N(x) - N(k)| = N(k+1) - N(k) \stackrel{a.s.}{=} O(\log k),$$

as $k \rightarrow \infty$. Hence, when assuming H

$$\sup_{0 < t < 1} |\mathcal{K}_n(t) - \mathcal{K}_n(t)| / (t(1-t))^{\frac{1}{2}} g(t) = o_P(1)$$

with any function $g(t)$ such that $\lim_{t \downarrow 0} g(t) = \lim_{t \uparrow 1} g(t) = \infty$. Let $q \in \mathcal{Q}^*$. Since $I^*(q, c) < \infty$ for all $c > 0$ implies (7.0.3), we have, under H

$$(7.3.11) \quad \sup_{0 < t < 1} |\mathcal{K}_n(t) - \mathcal{K}_n(t)| / q(t) = o_P(1)$$

for $q \in \mathcal{Q}^*$ and such that $I^*(q, c) < \infty$ for all $c > 0$. Hence, by contiguity, we have (7.3.11) under A as well. This implies part (a) of Theorem 7.3.1 on account of (7.3.9).

(b) We assume A to hold. Using part (a) with $q \equiv 1$, we have, as $n \rightarrow \infty$

$$\lambda^{-\frac{1}{2}} K_n(t) \xrightarrow{D} B(t) + b(t), \quad \text{in } D[0, 1]$$

where

$$b(t) = \lambda^{-\frac{1}{2}} \left(\int_0^t \xi(s) ds - t \int_0^1 \xi(s) ds \right), \quad 0 \leq t \leq 1.$$

By the Skorohod-Dudley-Wichura theorem, there exist versions $\mathcal{Y}_n$, $n \geq 1$, and $\mathcal{Y}$ such that

$$(7.3.12) \quad \mathcal{Y}_n \xrightarrow{D} \lambda^{-\frac{1}{2}} K_n \quad \text{and} \quad \mathcal{Y} \xrightarrow{D} B + b$$

and

$$\sup_{0 < t < 1} |\mathcal{Y}_n(t) - \mathcal{Y}(t)| = o_P(1).$$

Hence for any q such that $\inf_{\delta \leq t \leq 1-\delta} q(t) > 0$ for all $\delta \in (0, \frac{1}{2})$ we have

$$\sup_{\delta \leq t \leq 1-\delta} |\mathcal{Y}_n(t) - \mathcal{Y}(t)|/q(t) = o_P(1),$$

which gives

$$\mathcal{Y}_n/q \xrightarrow{D} \mathcal{Y}/q \quad \text{in } D[\delta, 1-\delta].$$

Hence, on account of (7.3.12), we conclude

$$\lambda^{-\frac{1}{2}} K_n(t)/q(t) \xrightarrow{D} (B(t) + b(t))/q(t) \quad \text{in } D[\delta, 1-\delta]$$

for all $\delta \in (0, \frac{1}{2})$.

Proof of Corollary 7.3.1. By Theorem 7.3.1, we have that under A

$$K_n(t)/q(t) \xrightarrow{D} \left(\lambda^{\frac{1}{2}} B(t) + \int_0^t \xi(s) ds - t \int_0^1 \xi(s) ds \right) / q(t).$$

Given H , $N(n)/n \rightarrow \lambda$ in probability. Hence, contiguity implies that $N(n)/n \rightarrow \lambda$ in probability under A as well. In this example it can be easily seen directly by noting that, under A , $N(n)$ is Poisson random variable for each $n \geq 1$, with mean

$$n\lambda + \sum_{i=1}^n \xi\left(\frac{i}{n}\right) / n^{\frac{1}{2}},$$

where $\xi^2(s)$ is a Riemann integrable function, and so we have as $n \rightarrow \infty$

$$E \frac{N(n)}{n} = \lambda + n^{-\frac{1}{2}} \frac{1}{n} \sum_{i=1}^n \xi\left(\frac{i}{n}\right) \rightarrow \lambda.$$

Consequently, $N(n)/n \xrightarrow{P} \lambda$. Hence

$$\left(\frac{n}{N(n)} \right)^{\frac{1}{2}} K_n(t)/q(t) \xrightarrow{D} \left(B(t) + \lambda^{-\frac{1}{2}} \left(\int_0^t \xi(s) ds - t \int_0^1 \xi(s) ds \right) \right) / q(t).$$

CHAPTER 8

Weighted Two-Time Parameter Empirical Processes Under Contiguous Measures

8.0. Introduction

Let $R_{1n}, \dots, R_{nn}$ denote the normalized ranks

$$R_{kn} = n^{-1} \sum_{i=1}^n 1\{X_i \leq X_k\}, \quad k = 1, \dots, n,$$

and let $\xi_1, \dots, \xi_n$ be the normalized sequential ranks

$$\xi_k = k^{-1} \sum_{i=1}^k 1\{X_i < X_k\}, \quad k = 1, \dots, n,$$

of the first n_k of the random variables $X_1, X_2, \dots$.

We consider the two-time parameter empirical processes

$$(8.0.1) \quad \alpha_n(s, t) = n^{-1/2} \sum_{i=1}^{[nt]} (1\{F(X_i) \leq s\} - s),$$

$$(8.0.2) \quad \beta_n(s, t) = n^{-1/2} \sum_{i=1}^{[nt]} (1\{R_{in} \leq s\} - \frac{[ns]}{n}),$$

$$(8.0.3) \quad \gamma_n(s, t) = n^{-1/2} \sum_{i=1}^{[nt]} (1\{\xi_i \leq s\} - \frac{[is]}{i}),$$

where $0 \leq s, t \leq 1$, and F is assumed to be continuous. Such empirical processes were studied by Pardzhanadze and Khmaladze (1986) under the class of contiguous alternatives H_1 .

They consider the null hypothesis

$$H_0: X_i, \quad 1 \leq i \leq n, \quad \text{have the same distribution } F,$$

versus the alternative hypothesis

$H_1: X_i, 1 \leq i \leq n$, have the respective distribution functions F_{in} ,

where we assume that all F_{in} are absolutely continuous with respect to the distribution function F and

$$(8.0.4) \quad \left[\frac{dF_{in}}{dF}(F^{-1}(u)) \right]^{1/2} = 1 + \frac{1}{2\sqrt{n}} g_n(t, u), \quad \frac{i-1}{n} < t \leq \frac{i}{n},$$

where $F^{-1}(u) = \inf\{x : (F(x) \geq u)\}$, $0 < u \leq 1$, $F^{-1}(0) = F^{-1}(0+)$ (cf. (2.3.1)).

We assume also that there exists a function $g \in L^2[0, 1]^2$, such that

$$(8.0.5) \quad \int_0^1 g(t, u) du = 0 \quad \text{for almost all } t \in [0, 1]$$

and

$$(8.0.6) \quad \int_0^1 \int_0^1 \left[g_n(t, u) - g\left(\frac{[nt]}{n}, u\right) \right]^2 du dt \rightarrow 0, \quad n \rightarrow \infty.$$

The sequence of direct products $F_{1n} \times \dots \times F_{nn}$, $n = 1, 2, \dots$ is contiguous to the sequence $F \times \dots \times F$ (cf. Section 2.3). In particular, for the changepoint alternatives we assume that there exists $\eta \in (0, 1)$ such that

$$(8.0.7) \quad g(t, u) = 1\{t \geq \eta\}g(u)$$

for some square integrable function g .

The process $\alpha_n(s, t)$ is a two-time parameter version of the usual empirical process. Csörgő and Horváth (1987a) propose nonparametric procedures for the changepoint problem based on the $\beta_n(s, t)$ process. Using a different parameterization of contiguity from that of (8.0.4)–(8.0.6) and a method of proof that differs from ours, Lejpus (1988, 1989) studies the processes $\alpha_n(s, t)$ and $\beta_n(s, t)$ as well as some functionals of $\beta_n(s, t)$ in connection with changepoint problems. The process $\gamma_n(s, t)$ was used by Csörgő and Horváth (1987b) to construct sequential

procedures for detecting a possible changepoint in a random sequence. Here we introduce some weight functions $q(t)$ and prove weighted approximations of the empirical processes of (8.0.1), (8.0.2) and (8.0.3) under the null hypothesis H_0 , as well as under the class of contiguous alternatives H_1 . These results with $q \equiv 1$ constitute an alternative proof of Theorem 1 of Pardzhanadze and Khmaladze (1986).

It is known (cf. Picard, 1985; Deshayes and Picard, 1986) that when testing for a changepoint, possibly occurring in a sequence of independent, identically distributed observations, it may be of interest to make use of tests based on the “bridge” like (tied down at $t = 0$ and $t = 1$) process

$$\alpha_n(s, t) - t\alpha_n(s, 1).$$

Inspired by discussions in Picard (1985) and Deshayes and Picard (1986) concerning the problem of weight functions and tail sensitivity in connection with changepoint problems, we study the asymptotic behaviour of such a process for a wide class of weight functions $q(t)$. Statistics based on such processes should be sensitive to detect a possible changepoint not only in the middle but also on both tails.

8.1. Uniform empirical processes

Let $X_1, X_2, \dots$ be i.i.d.r.v.'s with distribution function $F(\cdot)$ and define the empirical distribution function of the sample $X_1, \dots, X_n$ by

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}\{X_i \leq x\}.$$

An equivalent definition of the empirical distribution function $F_n(\cdot)$ can be given in terms of the order statistics $X_{1,n} \leq X_{2,n} \leq \dots \leq X_{n,n}$ of the random sample $X_1, \dots, X_n$ as follows:

$$(8.1.1) \quad F_n(x) = \begin{cases} 0, & X_{1,n} > x \\ \frac{k}{n}, & X_{k,n} \leq x < X_{k+1,n}, \quad k = 1, 2, \dots, n-1, \\ 1, & X_{n,n} \leq x. \end{cases}$$

Clearly, for every fixed x , $F_n(x)$ is the relative frequency of successes in a Bernoulli sequence of trials with $EF_n(x) = F(x)$ and $\text{Var } F_n(x) = \frac{1}{n}F(x)(1 - F(x))$. Consequently, by the classical strong law of large numbers,

$$(8.1.2) \quad F_n(x) \xrightarrow{\text{a.s.}} F(x), \quad \text{for } x \text{ fixed.}$$

Hence, using the language of statistics, $F_n(x)$ is an unbiased and strongly consistent estimator of $F(x)$ for each fixed x . Viewing $\{F_n(x); -\infty < x < \infty\}$ as a stochastic process, its sample functions are distribution functions and it is of great importance that $F(x)$ can be uniformly estimated by this process with probability one, as proved by Cantelli (1933) and Glivenko (1933). Namely we have

$$(8.1.3) \quad \sup_{-\infty < x < \infty} |F_n(x) - F(x)| \xrightarrow{\text{a.s.}} 0,$$

which tells us that, sampling ad infinitum, $F(x)$ can be uniquely determined with probability one.

From a practical point of view it is also of interest to study the rate of convergence in (8.1.3). Towards this end we define *the empirical process*

$$(8.1.4) \quad \zeta_n(x) = n^{-1/2}(F_n(x) - F(x)), \quad -\infty < x < \infty$$

For each fixed x one has immediately the central limit theorem:

$$\zeta_n(x) \xrightarrow{D} N(0, F(x)(1 - F(x))).$$

As to the rate of convergence we have the following result of Kolmogorov (1933) and Smirnov (1939).

Theorem 8.1.A. *If $F(x)$ is a continuous distribution function, then*

$$P\left\{\sup_{-\infty < x < \infty} |\zeta_n(x)| \leq y\right\} \rightarrow K(y)$$

where

$$K(y) = \begin{cases} \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2 y^2}, & y > 0 \\ 0, & \text{otherwise} \end{cases}$$

and

$$P\left\{\sup_{-\infty < x < \infty} \zeta_n(x) \leq y\right\} \rightarrow S(y)$$

where

$$S(y) = \begin{cases} 1 - e^{-2y^2}, & y > 0 \\ 0, & \text{otherwise.} \end{cases}$$

Let $U_i = F(X_i)$, $i = 1, 2, \dots$. Then U_i are $U(0, 1)$ random variable provided $F(\cdot)$ is continuous. Let now $E_n(s)$ be the empirical distribution function of the sample $U_1, \dots, U_n$ and denote the resulting empirical process in this case by

$$\alpha_n(s) = \sqrt{n}(E_n(s) - s), \quad 0 \leq s \leq 1$$

or equivalently, by

$$\alpha_n(s) = n^{-1/2} \sum_{i=1}^n (1\{U_i \leq s\} - s), \quad 0 \leq s \leq 1.$$

Then $E\alpha_n(s) = 0$ and the covariance function of the process $\{\alpha_n(s); 0 \leq s \leq 1\}$ is

$$\rho(s_1, s_2) = E\alpha_n(s_1)\alpha_n(s_2) = s_1 \wedge s_2 - s_1 s_2,$$

which coincides with that of a Brownian bridge $\{B(s); 0 \leq s \leq 1\}$. The weak convergence of $\alpha_n(\cdot)$ to Brownian bridge was proved by Donsker (1952). The first strong approximation of the empirical process is due to Brillinger (1969). The following theorem which gives the optimal rate of approximation is due to Komlós, Major and Tusnády (1975) and to Bretagnolle and Massart (1989) for the evaluation of the constants.

Theorem 8.1.B. *Given independent $U(0,1)$ random variables $U_1, U_2, \dots$, there exists a sequence of Brownian bridges $\{B_n(s); 0 \leq s \leq 1\}$ such that for all n and x we have*

$$P\left\{n^{-1/2} \sup_{0 \leq s \leq 1} |\alpha_n(s) - B_n(s)| > x + 12 \log n\right\} \leq 2 \exp(-x/6).$$

Consequently

$$\sup_{0 \leq s \leq 1} |\alpha_n(s) - B_n(s)| \stackrel{a.s.}{=} O(n^{-1/2} \log n).$$

Starting with Rényi (1953), Chibisov (1964) and O'Reilly (1974), there has been considerable interest in the asymptotic behaviour of weighted uniform empirical and quantile processes. For an insightful treatise of this subject we refer to Csörgő, Csörgő, Horváth and Mason (1986), and to Shorack and Wellner (1986), as well as to the references in these works. There are now complete characterizations available for describing the asymptotic behaviour of the weighted uniform empirical and quantile processes in supremum norm. The proofs of the following theorem can be found in Csörgő, Csörgő, Horváth and Mason (1986).

As in Chapter 5, let $\mathcal{Q}^*$ be the class of positive functions on $(0,1)$, i.e. $\inf_{\delta \leq t \leq 1-\delta} q(t) > 0$ for all $\delta \in (0, 1/2)$, which are nondecreasing near zero and non-increasing near one. We recall also the integrals

$$E^*(q, c) = \int_0^1 (t(1-t))^{-3/2} q(t) \exp(-c(t(1-t))^{-1} q^2(t)) dt$$

and

$$I^*(q, c) = \int_0^1 (t(1-t))^{-1} \exp(-cq^2(t)/(t(1-t))) dt, \quad c > 0.$$

Theorem 8.1.C. *Let $U_1, U_2, \dots$ be independent $U(0, 1)$ random variables and $q \in \mathcal{Q}^*$. (a) We can define a sequence of Brownian bridges $\{B_n(s); 0 \leq s \leq 1\}$ such that*

$$\sup_{0 < s < 1} |\alpha_n(s) - B_n(s)|/q(s) = o_P(1)$$

if and only if $I^(q, c) < \infty$ for all $c > 0$.*

(b) *We have*

$$\sup_{0 < s < 1} |\alpha_n(s)|/q(s) \xrightarrow{D} \sup_{0 < s < 1} |B(s)|/q(s),$$

where $\{B(s); 0 \leq s \leq 1\}$ is a Brownian bridge, if and only if $I^(q, c) < \infty$ for some $c > 0$.*

Remark 8.1.A According to Theorems 5.1.C and 5.1.B, Theorem 8.1.C can be stated equivalently in terms of the integral $E^*(q, c)$.

Similar results hold true for uniform quantile processes.

In Section 8.9 we study $\alpha_n(s)/q(s)$ under contiguous alternatives H_1 .

8.2. Weighted Two-Time Parameter Empirical Processes Based on Observations.

Let $\{K(s, t); 0 \leq s \leq 1, 0 \leq t < \infty\}$ denote a Kiefer process, i.e. a two-time parameter Gaussian process with mean zero and covariance function

$$(8.2.1) \quad EK(s_1, t_1)K(s_2, t_2) = (s_1 \wedge s_2 - s_1 s_2)(t_1 \wedge t_2).$$

Kiefer was the first one to call attention to the fact that the empirical process should be viewed as a two-time parameter process and that it should be approximated by an appropriate two-dimensional Gaussian process. He also gave a solution to this problem and proved the first two-dimensional strong approximation theorem for $\alpha_n(s)$. (cf. Kiefer (1972)).

Theorem 8.2.A. *Given independent $U(0,1)$ random variables $U_1, U_2, \dots$, there exists a Kiefer process $\{K(s, t); 0 \leq s \leq 1, t \geq 0\}$ such that*

$$\sup_{0 \leq s \leq 1} |n^{1/2} \alpha_n(s) - K(s, n)| \stackrel{a.s.}{=} O(n^{1/3} (\log n)^{2/3}).$$

This result was extended to the multivariate space parameter by many authors. The situation there is much more complicated since the distribution of empirical process is not distribution free any more, and the problem is far from being completely solved (cf. Csörgő and Horváth, 1988d and the references given there). In the univariate case the best result is due to Komlós, Major and Tusnády (1975) and to Bonvalot and Castelle (1991) for the evaluation of the constants.

Theorem 8.2.B. *Given independent $U(0,1)$ random variables $U_1, U_2, \dots$, there exists a Kiefer process $\{K(s, t); 0 \leq s \leq 1, t \geq 0\}$ such that*

$$P \left\{ \sup_{1 \leq k \leq n} \sup_{0 \leq s \leq 1} |k^{1/2} \alpha_k(s) - K(s, k)| > (x + 76 \text{LOG} n) \text{LOG} n \right\} \leq 2.028 \exp(-x/41)$$

with $\text{LOG} n = \max(\log n, \log 4)$. Consequently

$$\sup_{0 \leq s \leq 1} |n^{1/2} \alpha_n(s) - K(s, n)| \stackrel{a.s.}{=} O(\log^2 n).$$

Let $X_1, X_2, \dots$ be independent random variables. Here we consider the two-time parameter empirical process

$$\alpha_n(s, t) = n^{-1/2} \sum_{i=1}^{\lfloor nt \rfloor} (1\{F(X_i) \leq s\} - s), \quad 0 \leq s, t \leq 1,$$

with weights in the time parameter $t \in (0, 1]$, where F is assumed to be continuous. We should note that, actually, the two-time parameter version of the empirical process $\zeta_n(s)$ as in (8.1.4) is

$$\zeta_n(x, t) = n^{-1/2} \sum_{i=1}^{[nt]} \left(\mathbf{1}\{X_i \leq x\} - F(x) \right), \quad -\infty < x < \infty, \quad 0 \leq t \leq 1.$$

Since for any random variable X having a continuous distribution function F , the random variable $F(X)$ is uniformly distributed on $[0, 1]$, for all continuous F we have

$$\left\{ \zeta_n(F^{-1}(s), t); \quad 0 \leq s, t \leq 1, \quad n \geq 1 \right\} = \left\{ \alpha_n(s, t), \quad 0 \leq s, t \leq 1, \quad n \geq 1 \right\},$$

i.e., the above two empirical processes are identical. Hence all results proved for $\alpha_n(s, t)$ will hold automatically also for $\zeta_n(x, t)$. Indeed, for every $\omega \in \Omega$, $\alpha_n(F(x), t) = \zeta_n(x, t)$, $x \in \mathbb{R}^1$, $n \geq 1$.

We will call q a positive function on $(0, 1]$ if $q : (0, 1] \rightarrow (0, \infty)$ is such that

$$(8.2.2) \quad \inf_{\delta \leq t \leq 1} q(t) > 0 \quad \text{for all } 0 < \delta < 1.$$

Theorem 8.2.1. *We assume that H_0 holds. Then there exists a Kiefer process $K(\cdot, \cdot)$ such that, with a weight function q which is positive on $(0, 1]$, we have, as $n \rightarrow \infty$*

(a) *if q is such that*

$$(8.2.3) \quad \lim_{t \downarrow 0} (t \log \log 1/t)^{1/2} / q(t) = 0,$$

then

$$\sup_{0 < t \leq 1} \sup_{0 \leq s \leq 1} |\alpha_n(s, t) - n^{-1/2} K(s, nt)| / q(t) = o_p(1),$$

(b) if q is such that

$$(8.2.4) \quad \limsup_{t \downarrow 0} (t \log \log 1/t)^{1/2} / q(t) < \infty,$$

then

$$\sup_{0 < t \leq 1} \sup_{0 \leq s \leq 1} |\alpha_n(s, t) - n^{-1/2} K(s, nt)| / q(t) = O_P(1),$$

(c) if q is such that (8.2.4) holds, then

$$\sup_{0 < t \leq 1} \sup_{0 \leq s \leq 1} |\alpha_n(s, t)| / q(t) \xrightarrow{P} \sup_{0 < t \leq 1} \sup_{0 \leq s \leq 1} |K(s, t)| / q(t).$$

Remark 8.2.1. We emphasize that (b) and (c) do not follow from (a), nor does (c) from (b).

In the proof of Theorem 8.2.1 we will use the following lemma.

Lemma 8.2.1 (a) *Let q be a positive function on $(0, 1]$ and such that (8.2.3) holds. Then, as $T \rightarrow \infty$, we have*

$$(8.2.5) \quad \sup_{0 < t \leq 1/T} \sup_{0 \leq s \leq 1} |K(s, t)| / q(t) \stackrel{a.s.}{=} o(1),$$

and

$$(8.2.6) \quad \sup_{0 \leq s \leq 1/T} \sup_{0 < t \leq 1} |K(s, t)| / q(t) \stackrel{a.s.}{=} o(1),$$

(b) *Let q be a positive function on $(0, 1]$ and such that (8.2.4) holds. Then, as $T \rightarrow \infty$, we have*

$$(8.2.7) \quad \sup_{0 < t \leq 1/T} \sup_{0 \leq s \leq 1} |K(s, t)| / q(t) \stackrel{a.s.}{=} O(1),$$

and

$$(8.2.8) \quad \sup_{0 \leq s \leq 1/T} \sup_{0 < t \leq 1} |K(s, t)| / q(t) \stackrel{a.s.}{=} O(1),$$

Proof of Lemma 8.2.1. By Corollary 1.15.1 in Csörgő and Révész (1981)

we have

$$\begin{aligned}
& \lim_{T \rightarrow \infty} \sup_{0 < t \leq 1/T} \sup_{0 \leq s \leq 1} |K(s, t)| / (t \log \log 1/t)^{1/2} \\
&= \lim_{T \rightarrow \infty} \sup_{1/t \leq 1/T} \sup_{0 \leq s \leq 1} |K(s, 1/t)| / (t^{-1} \log \log t)^{1/2} \\
&= \lim_{T \rightarrow \infty} \sup_{t \geq T} \sup_{0 \leq s \leq 1} t |K(s, 1/t)| / (t(t^{-1} \log \log t)^{1/2}) \\
&= \lim_{T \rightarrow \infty} \sup_{t \geq T} \sup_{0 \leq s \leq 1} |K^*(s, t)| / (t \log \log t)^{1/2} \\
&\stackrel{a.s.}{=} 2^{-1/2},
\end{aligned}$$

where $K^*(s, t)$ is again a Kiefer process. Hence

$$\begin{aligned}
& \sup_{0 < t \leq 1/T} \sup_{0 \leq s \leq 1} |K(s, t)| / q(t) \\
&\stackrel{a.s.}{=} O(1) \sup_{0 < t \leq 1/T} (t \log \log 1/t)^{1/2} / q(t)
\end{aligned}$$

as $T \rightarrow \infty$, which gives (8.2.5) and (8.2.7) by using (8.2.3) and (8.2.4) respectively.

As to the proof of (8.2.6) and (8.2.8) we have

$$\begin{aligned}
& \sup_{0 \leq s \leq 1/T} \sup_{0 < t \leq 1} |K(s, t)| / q(t) \\
&\leq \sup_{0 \leq s \leq 1/T} \sup_{0 < t < \delta} |K(s, t)| / q(t) + \sup_{0 \leq s \leq 1/T} \sup_{\delta \leq t \leq 1} |K(s, t)| / q(t) \\
&= I_1(T, t) + I_2(T, t).
\end{aligned}$$

Similarly as right above,

$$\begin{aligned}
I_1(T, t) &\leq \sup_{0 < t < \delta} \sup_{0 \leq s \leq 1/T} |K(s, t)| / (t \log \log 1/t)^{1/2} \sup_{0 < t < \delta} (t \log \log 1/t)^{1/2} / q(t) \\
&\stackrel{a.s.}{=} \begin{cases} \alpha(1) & \text{if } q \text{ satisfies (8.2.3)} \\ O(1) & \text{if } q \text{ satisfies (8.2.4)} \end{cases}
\end{aligned}$$

as $T \rightarrow \infty$ and taking $\delta > 0$ arbitrarily small. We have also

$$\begin{aligned}
I_2(T, t) &\leq \sup_{0 < s \leq 1/T} \sup_{\delta \leq t < 1} |K(s, t)| (1 / \inf_{\delta \leq t < 1} q(t)) \\
&\stackrel{a.s.}{=} \alpha(1)
\end{aligned}$$

as $T \rightarrow \infty$, for any positive function q on $(0, 1]$ on account of $K(s, t)$ being almost surely continuous.

Proof of Theorem 8.2.1. Let

$$S_{nt}(s) = \sum_{i=1}^{[nt]} (1\{F(X_i) \leq s\} - s).$$

By the Komlós, Major and Tusnády (1975) approximation there exists a Kiefer process $\tilde{K}(s, t)$ such that, as $nt \rightarrow \infty$

$$(8.2.9) \quad \sup_{0 \leq s \leq 1} |S_{nt}(s) - \tilde{K}(s, nt)| \stackrel{a.s.}{=} O(\log^2 nt).$$

Hence for any $\mu > 0$

$$(8.2.10) \quad \sup_{\mu/n \leq t \leq 1} \sup_{0 \leq s \leq 1} |S_{nt}(s) - \tilde{K}(s, nt)| / (nt)^{1/2} \stackrel{a.s.}{=} O(1).$$

Also, by (8.2.2) and (8.2.9) we get

$$(8.2.11) \quad \sup_{\delta \leq t \leq 1} \sup_{0 \leq s \leq 1} |S_{nt}(s) - \tilde{K}(s, nt)| / n^{1/2} q(t) \stackrel{a.s.}{=} O \left(\sup_{\delta \leq t \leq 1} (\log^2 nt) / n^{1/2} q(t) \right) \stackrel{a.s.}{=} o(1)$$

for all $\delta \in (0, 1)$. We note that (8.2.3) as well as (8.2.4) imply

$$(8.2.12) \quad \lim_{t \downarrow 0} t^{1/2} / q(t) = 0.$$

Hence, using also (8.2.10), we have

$$(8.2.13) \quad \sup_{1/n \leq t \leq \delta} \sup_{0 \leq s \leq 1} |S_{nt}(s) - \tilde{K}(s, nt)| / n^{1/2} q(t) \leq O(1) \sup_{0 \leq t \leq \delta} t^{1/2} / q(t) \stackrel{a.s.}{=} o(1)$$

as $n \rightarrow \infty$ and taking $\delta > 0$ arbitrarily small.

We have also, as $n \rightarrow \infty$

$$(8.2.14) \quad \sup_{0 < t < 1/n} \sup_{0 \leq s \leq 1} |S_{nt}(s)|/n^{1/2}q(t) \stackrel{a.s.}{=} o(1),$$

because of

$$\begin{aligned} \sup_{0 < t < 1/n} \sup_{0 \leq s \leq 1} |S_{nt}(s)|/n^{1/2}q(t) &\leq \sup_{0 < t < 1/n} \sup_{0 \leq s \leq 1} nt((1-s) \vee s)/n^{1/2}q(t) \\ &\leq \sup_{0 < t < 1/n} n^{1/2}t/q(t) \leq \sup_{0 < t < 1/n} t^{1/2}/q(t) = o(1) \end{aligned}$$

as $n \rightarrow \infty$.

So far we have shown that, with a positive function q on $(0, 1]$ and such that (8.2.12) holds, we have

$$(8.2.15) \quad \sup_{0 < t \leq 1} \sup_{0 \leq s \leq 1} |S_{nt}(s) - \bar{K}(s, nt)|/n^{1/2}q(t) \stackrel{a.s.}{=} o(1)$$

as $n \rightarrow \infty$, where

$$\bar{K}(s, nt) = \begin{cases} K(s, nt) & \text{for } t \in [1/n, 1] \\ 0 & \text{elsewhere} \end{cases}$$

and $K(\cdot, \cdot)$ is a Kiefer process as in (8.2.9).

Next, we have for each $n \geq 1$

$$(8.2.16) \quad \sup_{0 < t < 1/n} \sup_{0 \leq s \leq 1} |K(s, nt)|/n^{1/2}q(t) \stackrel{\mathcal{D}}{=} \sup_{0 < t < 1/n} \sup_{0 \leq s \leq 1} |K(s, t)|/q(t).$$

Using now (8.2.5) and (8.2.7) in Lemma 8.2.1, we get

$$(8.2.17) \quad \sup_{0 < t < 1/n} \sup_{0 \leq s \leq 1} |K(s, nt)|/n^{1/2}q(t) = \begin{cases} o_P(1) & \text{if (8.2.3) holds} \\ O_P(1) & \text{if (8.2.4) holds,} \end{cases}$$

which together with (8.2.15) gives (a) and (b) respectively.

As to the proof of (c) we note that by (8.2.15) we have

$$(8.2.18) \quad \lim_{n \rightarrow \infty} \left| P \left\{ \sup_{0 < t \leq 1} \sup_{0 \leq s \leq 1} |S_{nt}(s)|/n^{1/2}q(t) \leq x \right\} - P \left\{ \sup_{1/n \leq t \leq 1} \sup_{0 \leq s \leq 1} |K(s, t)|/q(t) \leq x \right\} \right| = 0$$

for any Kiefer process $K(\cdot, \cdot)$ and $-\infty < x < \infty$. We have also for any Kiefer process $K(\cdot, \cdot)$ and function q positive on $(0, 1]$ and such that (8.2.4) holds

$$R_n(x) := P \left\{ \sup_{1/n \leq t \leq 1} \sup_{0 \leq s \leq 1} |K(s, t)|/q(t) \leq x \right\} \rightarrow R(x) = P \left\{ \sup_{0 < t \leq 1} \sup_{0 \leq s \leq 1} |K(s, t)|/q(t) \leq x \right\}$$

as $n \rightarrow \infty$, which together with (8.2.18) gives (c).

Corollary 8.2.1. *We assume that H_0 holds with F continuous and let $\{K(s, t) : 0 \leq s \leq 1, 0 \leq t \leq 1\}$ be a Kiefer process. Then we have, as $n \rightarrow \infty$*

(a) *if q is a positive function on $(0, 1]$ and such that (8.2.3) holds, then*

$$(8.2.19) \quad \alpha_n(s, t)/q(t) \xrightarrow{D} K(s, t)/q(t)$$

in $D[0, 1]^2$;

(b) *if q is a positive function on $(0, 1]$, then (8.2.19) holds in $D([0, 1] \times [\delta, 1])$ for any $0 < \delta < 1$.*

Proof: (a) Follows from Theorem 8.2.1 (a).

(b) By Komlós, Major and Tusnády (1975) (or Theorem 8.2.1 (a) with $q \equiv 1$) and (8.2.2).

Picard (1985), Deshayes and Picard (1986) consider weak convergence of the sequence of weighted processes $\frac{\psi(t)}{t}\alpha_n(s, t)$, where ψ is a piecewise continuous nonnegative function on $(0, 1]$ satisfying the condition

$$\int_0^1 \left(\frac{\psi(t)}{t} \right)^2 dt < \infty.$$

For the proof they refer to Deshayes and Picard (1983). Consequently we may note that, on assuming (8.2.3), or square-integrability of the function $1/q$, we can conclude a convergence in distribution result for any continuous functional of $\alpha_n(s, t)/q(t)$. However, for the sup-functional, which is usually of the main interest in detecting changepoint problems, Theorem 8.2.1 (c) gives this desired convergence in distribution result assuming only (8.2.4). For example, part (c) of Theorem 8.2.1 holds true with

$$q(t) = (t \log \log 1/t)^{1/2}, \quad 0 < t \leq 1.$$

Using upper-lower class results (tests for upper and lower functions) for suprema of a Kiefer process we can also state Theorem 8.2.1 under different conditions. Namely, assuming that $q \in C(0, 1]$ and $q(t)/t^{1/2}$ is nonincreasing near zero, by Theorem 5.2 of Adler and Brown (1986) (cf. also Chung 1949 and Kiefer 1961) we can conclude

$$\sup_{0 < t < 1/n} \sup_{0 \leq s \leq 1} |K(s, nt)|/n^{1/2}q(t) = \begin{cases} o_P(1) & \text{if } \mathcal{J}(q, c) < \infty \text{ for all } c > 0 \\ O_P(1) & \text{if } \mathcal{J}(q, c) < \infty \text{ for some } c > 0, \end{cases}$$

where

$$\mathcal{J}(q, c) = \int_0^1 t^{-1} q^2(t) \exp(-ct^{-1} q^2(t)) dt, \quad c > 0.$$

This in turn will allow us to replace the condition (8.2.3) by

$$\mathcal{J}(q, c) < \infty \quad \text{for all } c > 0$$

and the condition (8.2.4) by

$$\mathcal{J}(q, c) < \infty \quad \text{for some } c > 0.$$

respectively (cf. Proof of Theorem 8.2.1).

The above remarks, as well as the one which follows, can be modified and used for any other two-time parameter empirical-type process considered in this Chapter.

We note that assuming (8.2.3) (or using any of the conditions discussed above) we obtain the weak convergence of any functional of $\alpha_n(s, t)$ continuous in $\|(\cdot, \cdot)/q(t)\|$ -metric. As was seen already, for the $\sup_{0 < t \leq 1} \sup_{0 \leq s \leq 1} |\alpha_n(s, t)|/q(t)$ functional the convergence in distribution holds for a bigger class of weight functions (cf. Theorem 8.2.1 (c)). Considering L_p , $0 < p < \infty$, functionals, one can show that if q is positive on $(0, 1]$ and

$$(8.2.19) \quad \int_0^1 t^{p/2}/q(t) dt < \infty$$

then

$$\int_0^1 \int_0^1 |\alpha_n(s, t)|^p / q(t) ds dt \xrightarrow{\mathcal{D}} \int_0^1 \int_0^1 |K(s, t)|^p / q(t) ds dt,$$

where $K(\cdot, \cdot)$ is a Kiefer process. In fact, assuming (8.2.19) with q positive on $(0, 1]$, we have, as $n \rightarrow \infty$,

$$\int_0^1 \int_0^1 |\alpha_n(s, t) - n^{-1/2} K(s, nt)|^p / q(t) ds dt = o_P(1),$$

where $0 < p < \infty$.

In order to see this, we note first that, by Theorem 8.2.1(a) with $q \equiv 1$, we obtain for any $\delta \in (0, 1)$

$$\begin{aligned} & \int_{\delta}^1 \int_0^1 |\alpha_n(s, t) - n^{-1/2} K(s, nt)|^p / q(t) ds dt \\ & \leq \sup_{\delta \leq t \leq 1} \sup_{0 \leq s \leq 1} \left(|\alpha_n(s, t) - n^{-1/2} K(s, nt)|^p \right) \int_{\delta}^1 1/q(t) dt \\ & = o_P(1) O(1) = o_P(1), \end{aligned}$$

where q is positive on $(0, 1]$, i.e. (8.2.2) holds. Using (8.2.10) we have also

$$\begin{aligned} & \int_{1/n}^{\delta} \int_0^1 |\alpha_n(s, t) - n^{-1/2} K(s, nt)|^p / q(t) ds dt \\ & \leq \sup_{1/n \leq t < \delta} \sup_{0 \leq s \leq 1} \left(|\alpha_n(s, t) - n^{-1/2} K(s, nt)| / t^{1/2} \right)^p \int_0^{\delta} t^{p/2} / q(t) dt \\ & = O_P(1) \int_{1/n}^{\delta} t^{p/2} / q(t) dt \end{aligned}$$

for any $\delta \in (0, 1]$. Taking $\delta > 0$ arbitrarily small and assuming (8.2.19), by the above two statements we obtain

$$(8.2.20) \quad \int_{1/n}^1 \int_0^1 |\alpha_n(s, t) - n^{-1/2} K(s, nt)|^p / q(t) ds dt = o_P(1).$$

Next we note that the sequence

$$V_n(\omega) = \int_{1/n}^1 \int_0^1 |K(s, t)|^p / q(t) ds dt$$

is nondecreasing for each $\omega \in \Omega$. Thus

$$\lim_{n \rightarrow \infty} V_n(\omega) = V(\omega) \leq \infty$$

exists and, by the monotone convergence theorem, we have

$$(8.2.21) \quad \lim_{n \rightarrow \infty} EV_n = EV.$$

Since $K(s, t)/(s(1-s)t)^{1/2}$ is a $N(0, 1)$ random variable, we have

$$(8.2.22) \quad \begin{aligned} EV_n &= E|N(0, 1)|^p \int_{1/n}^1 \int_0^1 (s(1-s)t)^{p/2}/q(t) ds dt \\ &\rightarrow E|N(0, 1)|^p \int_1^1 \int_1^1 (s(1-s)t)^{p/2}/q(t) ds dt. \end{aligned}$$

Therefore, by (8.2.19), (8.2.21) and (8.2.22) we get $EV < \infty$, which implies

$$\int_0^1 \int_0^1 |K(s, t)|^p/q(t) ds dt < \infty \quad \text{a.s. .}$$

Consequently,

$$\int_0^{1/n} \int_0^1 |n^{-1/2} K(s, nt)|^p/q(t) ds dt = o_p(1)$$

which, together with (8.2.20), implies

$$\int_0^1 \int_0^1 |\alpha_n(s, t) - n^{-1/2} K(s, nt)|^p/q(t) ds dt = o_p(1).$$

8.3. Two-time parameter empirical processes tied down at $t = 1$.

It is known that when testing for the existence of a changepoint in a random sequence, it may be of interest to make use of test statistics based on comparing the empirical distribution function of the first k observations to that of the last $n - k$ observations. For example, Picard (1985), and Deshayes and Picard (1986) consider

$$(8.3.1) \quad \begin{aligned} &\sup_{1 \leq k < n} \sup_{x \in \mathbb{R}} n^{1/2} \left| \frac{1}{k} \sum_{i=1}^k 1\{X_i \leq x\} - \frac{1}{n-k} \sum_{i=k+1}^n 1\{X_i \leq x\} \right| \\ &= \sup_{1 \leq k < n} \sup_{x \in \mathbb{R}} n^{1/2} \left| \sum_{i=1}^k 1\{X_i \leq x\} - \frac{k}{n} \sum_{i=1}^n 1\{X_i \leq x\} \right| / (k(n-k)/n). \end{aligned}$$

Since, given H_0 , as $n \rightarrow \infty$ we have

$$\sup_{1 \leq k < n} \sup_{x \in \mathbb{R}} \left| \sum_{i=1}^k 1\{X_i \leq x\} - \frac{k}{n} \sum_{i=1}^n 1\{X_i \leq x\} \right| / n^{1/2} \frac{k}{n} \left(1 - \frac{k}{n}\right) \xrightarrow{P} \infty,$$

as well as

$$\sup_{1 \leq k < n} \sup_{x \in \mathbb{R}} \left| \sum_{i=1}^k 1\{X_i \leq x\} - \frac{k}{n} \sum_{i=1}^n 1\{X_i \leq x\} \right| / n^{1/2} \left(\frac{k}{n} \left(1 - \frac{k}{n} \right) \right)^{1/2} \xrightarrow{P} \infty,$$

in order to have the above test statistics converge in distribution to a nondegenerate random variable, we have to renormalize them somehow. We may, for example, wish to introduce weight functions which are less severe on the tails than $((k/n)(1 - k/n))^{1/2}$ is.

Just like tests based on the classical Kolmogorov-Smirnov statistic, the ones based on

$$(8.3.2) \quad \sup_{1 \leq k < n} \sup_{x \in \mathbb{R}} \frac{k}{n} \left(1 - \frac{k}{n} \right) n^{1/2} \left| \frac{1}{k} \sum_{i=1}^k 1\{X_i \leq x\} - \frac{1}{n-k} \sum_{i=k+1}^n 1\{X_i \leq x\} \right|$$

$$= \sup_{1 \leq k < n} \sup_{x \in \mathbb{R}} n^{-1/2} \left| \sum_{i=1}^k 1\{X_i \leq x\} - \frac{k}{n} \sum_{i=1}^n 1\{X_i \leq x\} \right|$$

should be more powerful for detecting changes that occur in the middle, namely near $n/2$, where $\frac{k}{n} (1 - \frac{k}{n})$ has its maximum, than for noticing the ones occurring near the endpoints 0 and n . Thus, a weighted version of (8.3.2) should emphasize changes which may have occurred near the endpoints, while retaining sensitivity to possible changes in the middle as well. To see what kind of weight functions are possible for the statistic in (8.3.1), or in (8.3.2), we will study the asymptotic behavior of the process

$$(\alpha_n(s, t) - t\alpha_n(s, 1))/q(t)$$

for a wide class of functions q .

We will call q a positive function on $(0, 1)$ if

$$q : (0, 1) \rightarrow (0, \infty) \text{ is such that}$$

$$(8.3.3) \quad \inf_{\delta \leq t \leq 1-\delta} q(t) > 0 \quad \text{for all } 0 < \delta < 1/2.$$

Let $\{\Gamma(s, t); 0 \leq s \leq 1, 0 \leq t \leq 1\}$ denote a Gaussian process with mean zero and covariance function

$$(8.3.4) \quad E\Gamma(s_1, t_1)\Gamma(s_2, t_2) = (s_1 \wedge s_2 - s_1 s_2)(t_1 \wedge t_2 - t_1 t_2).$$

We introduce

$$\hat{\alpha}_n(s, t) = \begin{cases} n^{-1/2} \left(\sum_{i=1}^{[(n+1)t]} \mathbf{1}\{F(X_i) \leq s\} - \frac{[(n+1)t]}{n} \sum_{i=1}^n \mathbf{1}\{F(X_i) \leq s\} \right) & , 0 \leq t < 1, \\ & 0 \leq s \leq 1 \\ 0 & , t = 1. \end{cases}$$

Theorem 8.3.1. We assume that H_0 holds. Then, there exists a Kiefer process $K(\cdot, \cdot)$ such that with the sequence of stochastic processes $\Gamma_n(\cdot, \cdot)$

$$\{\Gamma_n(s, t), 0 \leq s \leq 1, 0 \leq t \leq 1\} = \{n^{-1/2}(K(s, nt) - tK(s, n)), 0 \leq s \leq 1, 0 \leq t \leq 1\} \\ \xrightarrow{D} \{\Gamma(s, t), 0 \leq s \leq 1, 0 \leq t \leq 1\} \text{ for each } n \geq 1,$$

and with a weight function q which is positive on $(0, 1)$ we have, as $n \rightarrow \infty$

(a) if

$$(8.3.5) \quad \begin{cases} \lim_{t \downarrow 0} (t(1-t) \log \log 1/(t(1-t)))^{1/2}/q(t) = 0 \\ \text{and} \\ \lim_{t \uparrow 1} (t(1-t) \log \log 1/(t(1-t)))^{1/2}/q(t) = 0, \end{cases}$$

then

$$\sup_{0 < t < 1} \sup_{0 \leq s \leq 1} |\hat{\alpha}_n(s, t) - \Gamma_n(s, t)|/q(t) = o_P(1),$$

(b) if

$$(8.3.6) \quad \begin{cases} \lim_{t \downarrow 0} (t(1-t) \log \log 1/(t(1-t)))^{1/2}/q(t) < \infty \\ \text{and} \\ \lim_{t \uparrow 1} (t(1-t) \log \log 1/(t(1-t)))^{1/2}/q(t) < \infty, \end{cases}$$

then

$$\sup_{0 < t < 1} \sup_{0 \leq s \leq 1} |\hat{\alpha}_n(s, t) - \Gamma_n(s, t)|/q(t) = O_P(1).$$

(c) if (8.3.6) holds then we have also

$$\sup_{0 < t < 1} \sup_{0 \leq s \leq 1} |\hat{\alpha}_n(s, t)|/q(t) \xrightarrow{P} \sup_{0 < t < 1} \sup_{0 \leq s \leq 1} |\Gamma(s, t)|/q(t).$$

Remark 8.3.1. We note here that with a function q which is positive on $(0, 1)$, we have, as $T \rightarrow \infty$

$$\sup_{0 < t < 1/T} \sup_{0 \leq s \leq 1} |K(s, t)|/q(1-t) \stackrel{a.s.}{=} \begin{cases} o(1) & \text{if (8.3.5) holds} \\ O(1) & \text{if (8.3.6) holds,} \end{cases}$$

where K is a Kiefer process (cf. Proof of Lemma 8.2.1).

Proof of Theorem 8.3.1. Let

$$S_{[(n+1)t]}(s) = \sum_{i=1}^{[(n+1)t]} (1\{F(X_i) \leq s\} - s)$$

and

$$S_{[(n+1)t]}^{(1)}(s) = \sum_{i=1}^{[(n+1)t]} (1\{F(X_i) \leq s\} - s), \quad 0 \leq t \leq 1/2, 0 \leq s \leq 1$$

$$S_{[(n+1)t]}^{(2)}(s) = \sum_{i=[(n+1)t]+1}^n (1\{F(X_i) \leq s\} - s), \quad 1/2 \leq t < 1, 0 \leq s \leq 1.$$

Then

$$S_{[(n+1)t]}(s) = \begin{cases} S_{[(n+1)t]}^{(1)}(s), & 0 \leq t \leq 1/2, 0 \leq s \leq 1 \\ S_{[\frac{n+1}{2}]}^{(1)}(s) + S_{[\frac{n+1}{2}]}^{(2)}(s) - S_{[(n+1)t]}^{(2)}(s), & 1/2 < t < 1, 0 \leq s \leq 1. \end{cases}$$

Let q be a positive function on $(0, 1)$. Using Theorem 8.2.1 we can construct two independent Kiefer processes $K^{(1)}(\cdot, \cdot)$ and $K^{(2)}(\cdot, \cdot)$ such that

$$(8.3.7) \quad \sup_{0 < t \leq 1/2} \sup_{0 \leq s \leq 1} n^{-1/2} \left| S_{[(n+1)t]}^{(1)}(s) - K^{(1)}(s, (n+1)t) \right| / q(t) \\ = \begin{cases} o_P(1) & \text{if (8.3.5) holds} \\ O_P(1) & \text{if (8.3.6) holds,} \end{cases}$$

and

$$\begin{aligned} & \sup_{0 < t < 1/2} \sup_{0 \leq s \leq 1} n^{-1/2} \left| S_{[(n+1)(1-t)]}^{(2)}(s) - K^{(2)}(s, (n+1)t) \right| / q(1-t) \\ &= \begin{cases} o_P(1) & \text{if (8.3.5) holds} \\ O_P(1) & \text{if (8.3.6) holds.} \end{cases} \end{aligned}$$

In order to see (8.3.8) we note that

$$S_{[(n+1)(1-t)]}^{(2)}(s) = \sum_{i=1}^{[(n+1)t]} \mathbf{1}\{F(X_{n-i+1}) \leq s\} - s, \quad 0 < t < 1/2, 0 \leq s \leq 1,$$

and also, as $n \rightarrow \infty$, we have for any Kiefer process $K(\cdot, \cdot)$

(8.3.9)

$$\sup_{0 < t < 1/(n+1)} \sup_{0 \leq s \leq 1} |n^{-1/2} K(s, (n+1)t| / q(1-t) = \begin{cases} o_P(1) & \text{if (8.3.5) holds} \\ O_P(1) & \text{if (8.3.6) holds,} \end{cases}$$

on account of (8.2.16) and Remark 8.3.1.

We define a Kiefer process $K(\cdot, \cdot)$ by

$$\{K(s, nt), 0 \leq t \leq 1, 0 \leq s \leq 1\}$$

(8.3.10)

$$= \begin{cases} K^{(1)}(s, nt), & 0 \leq t \leq 1/2, 0 \leq s \leq 1 \\ K^{(1)}(s, \frac{n}{2}) + K^{(2)}(s, \frac{n}{2}) - K^{(2)}(s, n - nt), & 1/2 < t \leq 1, 0 \leq s \leq 1. \end{cases}$$

Then we have

$$\begin{aligned} & \sup_{0 < t \leq 1/2} \sup_{0 \leq s \leq 1} |\hat{\alpha}_n(s, t) - n^{-1/2}(K(s, (n+1)t) - tK(s, n+1))| / q(t) \\ &= \sup_{0 < t \leq 1/2} \sup_{0 \leq s \leq 1} |n^{-1/2} \left(S_{[(n+1)t]}^{(2)}(s) - \frac{[(n+1)t]}{n} S_n(s) \right) \\ & \quad - n^{-1/2}(K(s, (n+1)t) - tK(s, n+1))| / q(t) \\ &\leq \sup_{0 < t \leq 1/2} \sup_{0 \leq s \leq 1} |n^{-1/2} S_{[(n+1)t]}^{(1)}(s) - n^{-1/2} K^{(1)}(s, (n+1)t)| / q(t) \\ & \quad + \sup_{0 < t \leq 1/2} \sup_{0 \leq s \leq 1} |n^{-3/2} [(n+1)t] S_n(s) - n^{-1/2} t K(s, n+1)| / q(t) \\ &= I_1^{(1)}(n) + I_2^{(1)}(n). \end{aligned}$$

By (8.3.7) we obtain

$$I_1^{(1)}(n) = \sup_{0 < t \leq 1/2} \sup_{0 \leq s \leq 1} |n^{-1/2} S_{[(n+1)t]}^{(1)}(s) - n^{-1/2} K^{(1)}(s, (n+1)t)|/q(t)$$

$$= \begin{cases} o_P(1) & \text{if (8.3.5) holds} \\ O_P(1) & \text{if (8.3.6) holds.} \end{cases}$$

When considering $I_2^{(1)}(n)$, we note that, by the definition of $S^{(1)}$ and $S^{(2)}$, we have for each $n \geq 1$

$$(8.3.11) \quad S_{[\frac{n+1}{2}]}^{(1)}(s) + S_{[\frac{n+1}{2}]}^{(2)}(s) = S_n(s), \quad 0 \leq s \leq 1$$

and by (8.3.10) with $t = 1$

$$(8.3.12) \quad K^{(1)}\left(s, \frac{n+1}{2}\right) + K^{(2)}\left(s, \frac{n+1}{2}\right) = K(s, n+1), \quad 0 \leq s \leq 1,$$

which implies

$$(8.3.13) \quad \begin{aligned} & \sup_{0 \leq s \leq 1} |n^{-1/2} S_n(s) - n^{-1/2} K(s, n+1)| \\ &= \sup_{0 \leq s \leq 1} |\alpha_n(s) - n^{-1/2} K(s, n+1)| \\ &= o_P(1). \end{aligned}$$

Also, for any q positive on $(0,1)$ and such that (8.3.5) or (8.3.6) hold, we have

$$\sup_{0 < t \leq 1/2} t/q(t) = \sup_{0 < t < \delta} t/q(t) \vee \sup_{\delta < t \leq 1/2} t/q(t) = O(1)$$

when taking $\delta > 0$ arbitrarily small. Hence, for q as above

$$\begin{aligned} I_2^{(1)}(n) &\leq \sup_{0 < t \leq 1/2} t/q(t) \sup_{0 \leq s \leq 1} |n^{-1/2} S_n(s) - n^{-1/2} K(s, n+1)| \\ &\quad + \sup_{0 < t \leq 1/2} \sup_{0 \leq s \leq 1} |([(n+1)t]/n - t)n^{-1/2} S_n(s)|/q(t) \\ &= o_P(1) + \epsilon^{(1)}(n), \end{aligned}$$

where

$$\begin{aligned} e^{(1)}(n) &= \sup_{0 \leq s \leq 1} |n^{-1/2} S_n(s)| \left(\sup_{0 < t < \delta} |[(n+1)t]/n - t|/q(t) \vee \sup_{\delta \leq t \leq 1/2} |[(n+1)t]/n - t|/q(t) \right) \\ &\leq O_P(1) \left((n+2)/n \sup_{0 < t < \delta} t/q(t) \vee o(1) \sup_{\delta \leq t \leq 1/2} 1/q(t) \right) \\ &= o_P(1), \end{aligned}$$

as $n \rightarrow \infty$ and taking $\delta > 0$ arbitrarily small, due to (8.3.3), and (8.3.5) or (8.3.6) (cf. (8.2.12)).

For $t \in (1/2, 1)$ we have

$$\begin{aligned} &\sup_{1/2 < t < 1} \sup_{0 \leq s \leq 1} |\hat{\alpha}_n(s, t) - n^{-1/2} (K(s, (n+1)t) - tK(s, n+1))|/q(t) \\ &= \sup_{1/2 < t < 1} \sup_{0 \leq s \leq 1} |n^{-1/2} \left(S_{[(n+1)t]}(s) - \frac{[(n+1)t]}{n} S_n(s) \right) \\ &\quad - n^{-1/2} (K(s, (n+1)t) - tK(s, n+1))|/q(t) \\ &\leq \sup_{1/2 < t < 1} \sup_{0 \leq s \leq 1} |n^{-1/2} S_{[(n+1)t]}^{(2)}(s) - n^{-1/2} K^{(2)}(s, (n+1)(1-t))|/q(t) \\ &\quad + \sup_{1/2 < t < 1} \sup_{0 \leq s \leq 1} |n^{-1/2} \left(S_{[\frac{n+1}{2}]}^{(1)}(s) + S_{[\frac{n+1}{2}]}^{(2)}(s) - \frac{[(n+1)t]}{n} S_n(s) \right) \\ &\quad - n^{-1/2} \left(K^{(1)}\left(s, \frac{n+1}{2}\right) + K^{(2)}\left(s, \frac{n+1}{2}\right) - tK(s, n+1) \right)|/q(t) \\ &= I_1^{(2)}(n) + I_2^{(2)}(n). \end{aligned}$$

By (8.3.8) we obtain

$$\begin{aligned} I_1^{(2)}(n) &= \sup_{1/2 < t < 1} \sup_{0 \leq s \leq 1} |n^{-1/2} S_{[(n+1)t]}^{(2)}(s) - n^{-1/2} K^{(2)}(s, (n+1)(1-t))|/q(t) \\ &= \sup_{0 < t < 1/2} \sup_{0 \leq s \leq 1} |n^{-1/2} S_{[(n+1)(1-t)]}^{(2)}(s) - n^{-1/2} K^{(2)}(s, (n+1)t)|/q(1-t) \\ &= \begin{cases} o_P(1) & \text{if (8.3.5) holds} \\ O_P(1) & \text{if (8.3.6) holds.} \end{cases} \end{aligned}$$

In order to show that $I_2^{(2)}(n) = o_P(1)$ when assuming (8.3.5) or (8.3.6), we note that for any function q positive on $(0,1)$ and such that (8.3.5) or (8.3.6) holds true, we have

$$\sup_{1/2 < t < 1} (1-t)/q(t) = \sup_{1/2 < t \leq 1-\delta} (1-t)/q(t) \vee \sup_{1-\delta < t < 1} (1-t)/q(t) = O(1)$$

when taking $\delta > 0$ arbitrarily small. Hence, and using also (8.3.13),

$$\begin{aligned} I_2^{(2)}(n) &\leq \sup_{1/2 < t < 1} \sup_{0 \leq s \leq 1} |n^{-1/2} S_n(s) - n^{-1/2} K(s, n+1)| (1-t)/q(t) \\ &\quad + \sup_{1/2 < t < 1} \sup_{0 \leq s \leq 1} \left| \left(t - \frac{[(n+1)t]}{n} \right) n^{-1/2} S_n(s) \right| / q(t) \\ &= o_P(1) + \epsilon^{(2)}(n), \end{aligned}$$

where

$$\begin{aligned} \epsilon^{(2)}(n) &\leq \sup_{0 \leq s \leq 1} |n^{-1/2} S_n(s)| \left(\sup_{1/2 < t \leq 1-\delta} \left| t - \frac{[(n+1)t]}{n} \right| / q(t) \right. \\ &\quad \vee \left. \left(\sup_{1-\delta < t < 1} (1-t)/q(t) + \sup_{1-\delta < t < 1} \left(1 - \frac{[(n+1)t]}{n} \right) / q(t) \right) \right) \\ &\leq O_P(1) \left(o(1) \sup_{1/2 < t \leq 1-\delta} 1/q(t) \vee \left(o(1) + \frac{n+1}{n} \sup_{1-\delta < t < 1} (1-t)/q(t) \right) \right) \\ &= o_P(1) \end{aligned}$$

as $n \rightarrow \infty$ and taking $\delta > 0$ arbitrarily small, due to (8.3.3), and (8.3.5) or (8.3.6). This concludes the proof of (a) and (b) of Theorem 8.3.1. The proof of (c) can be done along the lines of the proof of part (c) of Theorem 8.2.1.

Corollary 8.3.1. *We assume that H_0 holds.*

(a) *If q is a positive function on $(0,1)$ and such that (8.3.5) holds, then, as $n \rightarrow \infty$*

$$(8.3.10) \quad \hat{\alpha}_n(s, t)/q(t) \xrightarrow{P} \Gamma(s, t)/q(t)$$

in $D[0, 1]^2$.

(b) If q is a positive function on $(0, 1)$, then as $n \rightarrow \infty$, (8.3.10) holds in $D([0, 1] \times [\delta, 1 - \delta])$ for any $0 < \delta < 1/2$.

Remark 8.3.2. Obviously, we should note again that assuming (8.3.5), part (a) of Corollary 8.3.1 gives the limiting distribution of any continuous functional of $\hat{\alpha}_n(s, t)/q(t)$. In addition to that, Theorem 8.3.1 (c) gives convergence in distribution of sup-functional in $D[0, 1]^2$ of $\hat{\alpha}_n(s, t)/q(t)$ for example with

$$q(t) = \left(t(1-t) \log \log \frac{1}{t(1-t)} \right)^{1/2}, \quad 0 < t < 1.$$

Thus, one arrives at the following modification of (8.3.1) and (8.3.2): reject H_0 in favour of H_1 for large values of

$$\sup_{1 \leq k < n} \sup_{x \in \mathbb{R}} \left\{ \left| \sum_{i=1}^k \mathbf{1}\{X_i \leq x\} - \frac{k}{n} \sum_{i=1}^n \mathbf{1}\{X_i \leq x\} \right| n^{-1/2} \left(\frac{k}{n} \left(1 - \frac{k}{n} \right) \log \log \frac{1}{\frac{k}{n} \left(1 - \frac{k}{n} \right)} \right)^{-1/2} \right\}.$$

Deshayes and Picard (1986) conclude asymptotic behaviour of the test statistic

$$\sup_{1 \leq k < n} \sup_{x \in \mathbb{R}} n^{1/2} \psi \left(\frac{k}{n} \right) \left| \frac{1}{k} \sum_{i=1}^k \mathbf{1}\{X_i \leq x\} - \frac{1}{n-k} \sum_{i=k+1}^n \mathbf{1}\{X_i \leq x\} \right|$$

with ψ a non-negative, piecewise continuous function on the interval $(0, 1)$, under the condition $\int_0^1 \left(\frac{\psi(t)}{t(1-t)} \right)^2 dt < \infty$.

We note here that the limiting Gaussian process Γ appeared also in Hoeffding (1948) and Blum, Kiefer and Rosenblatt (1961) in the context of testing for independence.

8.4. Weighted Empirical Processes Based on Ranks.

Let $X_1, X_2, \dots$ be independent random variables with an unknown continuous distribution function F . For each $n \geq 1$, let $R_{1n}, \dots, R_{nn}$ denote the normalized ranks

$$R_{kn} = n^{-1} \sum_{i=1}^n \mathbf{1}\{X_i \leq X_k\}, \quad k = 1, \dots, n,$$

of the first n of the random variables $X_1, X_2, \dots$. In this section we consider asymptotic behaviour of the empirical process based on ranks with weight functions q , namely $\beta_n(s, t)/q(t)$, where

$$\beta_n(s, t) = n^{-1/2} \sum_{i=1}^{\lfloor nt \rfloor} \left(\mathbf{1}\{R_{in} \leq s\} - \frac{\lfloor ns \rfloor}{n} \right), \quad 0 \leq s, t \leq 1.$$

We note that $\beta_n(s, t)$ is mean zero process on account of $P\{nR_{in} = k\} = \frac{1}{n}$, $1 \leq i, k \leq n$.

Let $\{\Gamma(s, t), 0 \leq s \leq 1, 0 \leq t \leq 1\}$ denote a Gaussian process with mean zero and covariance function as in (8.3.4). We define

$$\hat{\beta}_n(s, t) = \begin{cases} n^{-1/2} \sum_{i=1}^{\lfloor (n+1)t \rfloor} \left(\mathbf{1}\{R_{in} \leq s\} - \frac{\lfloor ns \rfloor}{n} \right) & , 0 \leq s \leq 1, 0 \leq t < 1 \\ 0 & , 0 \leq s \leq 1, t = 1. \end{cases}$$

Theorem 8.4.1. *We assume that H_0 holds. Then there exists a Kiefer process $K(\cdot, \cdot)$ such that with the sequence of stochastic processes $\Gamma_n(\cdot, \cdot)$, $n \geq 1$*

$$\{\Gamma_n(s, t), 0 \leq s \leq 1, 0 \leq t \leq 1\} = \{n^{-1/2}(K(s, nt) - tK(s, n)), 0 \leq s \leq 1, 0 \leq t \leq 1\}$$

$$\stackrel{D}{=} \{\Gamma(s, t), 0 \leq s \leq 1, 0 \leq t \leq 1\},$$

and with a weight function q which is positive on $(0, 1)$ we have, as $n \rightarrow \infty$

(a) if (8.3.5) holds, then

$$\sup_{0 < t < 1} \sup_{0 \leq s \leq 1} |\hat{\beta}_n(s, t) - \Gamma_n(s, t)|/q(t) = o_P(1)$$

(b) if (8.3.6) holds, then

$$\sup_{0 < t < 1} \sup_{0 \leq s \leq 1} |\hat{\beta}_n(s, t) - \Gamma_n(s, t)|/q(t) = O_P(1)$$

(c) if (8.3.6) holds, then

$$\sup_{0 < t < 1} \sup_{0 \leq s \leq 1} |\hat{\beta}_n(s, t)|/q(t) \xrightarrow{P} \sup_{0 < t < 1} \sup_{0 \leq s \leq 1} |\Gamma(s, t)|/q(t).$$

In order to prove Theorem 8.4.1, we note that

$$\begin{aligned} \beta_n(s, t) &= n^{-1/2} \sum_{i=1}^{[nt]} (1\{R_{i,n} \leq s\} - \frac{[ns]}{n}) \\ &= n^{-1/2} \sum_{i=1}^{[nt]} (1\{F(X_i) \leq U_n(s)\} - \frac{[ns]}{n}), \end{aligned}$$

where $U_n(s)$ is defined by

$$(8.4.1) \quad U_n(s) = \begin{cases} 0 & \text{for } s \in [0, 1/n) \\ U_{k,n} = F(X_{k,n}) & \text{for } s \in [k/n, (k+1)/n), \quad k = 1, \dots, n-1 \\ U_{n,n} & \text{for } s = 1, \end{cases}$$

and $X_{1,n} \leq \dots \leq X_{n,n}$ denote the order statistics of $X_1, \dots, X_n$. We note also that the process $\beta_n(s, t)$ already has a "bridge" type structure as it is, namely we have

$$\beta_n(s, t) = \alpha_n(U_n(s), t) - \frac{[nt]}{n} \alpha_n(U_n(s), 1).$$

To prove Theorem 8.4.1 we will need the following result.

Lemma 8.4.1. *Let q be a positive function on $(0, 1]$. Then, as $n \rightarrow \infty$ we have*

$$\sup_{0 < t \leq 1} \sup_{0 \leq s \leq 1} |K(U_n(s), nt) - K(s, nt)|/(n^{1/2}q(t)) = \begin{cases} o_P(1) & \text{if (8.2.3) holds} \\ O_P(1) & \text{if (8.2.4) holds,} \end{cases}$$

where $K(\cdot, \cdot)$ is a Kiefer process and $U_n(s)$ is defined as in (8.4.1).

Proof of Lemma 8.4.1. By Lemma 4.5.1 in Csörgő and Révész (1981), we have

(8.4.2)

$$\sup_{1/n \leq s \leq 1} |K(U_n(s), nt) - K(s, nt)| \stackrel{a.s.}{=} O\left((nt)^{1/4}(\log \log nt)^{1/4}(\log nt)^{1/2}\right),$$

as $nt \rightarrow \infty$. Hence for any $\mu > 0$

$$\begin{aligned} & \sup_{\mu/n \leq t \leq 1} \sup_{1/n \leq s \leq 1} |K(U_n(s), nt) - K(s, nt)| / (n^{1/2} q(t)) \\ & \stackrel{a.s.}{=} O\left(\sup_{\mu/n \leq t \leq 1} ((nt)^{1/4}(\log \log nt)^{1/4}(\log nt)^{1/2} / (nt)^{1/2})(t^{1/2}/q(t))\right) \\ & \leq O\left(\sup_{\mu/n \leq t \leq \delta} (nt)^{-1/4}(\log \log nt)^{1/4}(\log nt)^{1/2} \sup_{\mu/n \leq t \leq \delta} t^{1/2}/q(t)\right. \\ & \quad \left.+ \sup_{\delta \leq t \leq 1} (nt)^{-1/4}(\log \log nt)^{1/4}(\log nt)^{1/2} \sup_{\delta \leq t \leq 1} t^{1/2}/q(t)\right) \\ & = O(O(1)\alpha(1) + \alpha(1)O(1)) \\ & \stackrel{a.s.}{=} \alpha(1), \end{aligned}$$

as $n \rightarrow \infty$ and taking $\delta > 0$ arbitrarily small, for any function q which is positive on $(0, 1]$ and such that $\lim q(t)/t^{1/2} = \infty$ (cf. (8.2.12)).

Next we have

$$\begin{aligned} & \sup_{0 < t < \mu/n} \sup_{1/n \leq s \leq 1} |K(U_n(s), nt) - K(s, nt)| / n^{1/2} q(t) \\ & \leq 2 \sup_{0 < t < \mu/n} \sup_{0 \leq s \leq 1} |K(s, nt)| / n^{1/2} q(t) \\ & = \begin{cases} o_p(1) & \text{if (8.2.3) holds} \\ O_p(1) & \text{if (8.2.4) holds} \end{cases} \end{aligned}$$

as $n \rightarrow \infty$, by (8.2.5) and (8.2.7) respectively. Hence, by the last two results, as $n \rightarrow \infty$

$$\sup_{0 < t < 1} \sup_{1/n \leq s \leq 1} |K(U_n(s), nt) - K(s, nt)| / n^{1/2} q(t) = \begin{cases} o_p(1) & \text{if (8.2.3) holds} \\ O_p(1) & \text{if (8.2.4) holds.} \end{cases}$$

Also, on account of $U_n(s) = 0$ for $s \in [0, 1/n)$,

$$\begin{aligned} & \sup_{0 < t \leq 1} \sup_{0 \leq s < 1/n} |K(U_n(s), nt) - K(s, nt)|/n^{1/2}q(t) \\ &= \sup_{0 < t \leq 1} \sup_{0 \leq s < 1/n} |K(s, nt)|/n^{1/2}q(t) \\ &\stackrel{D}{=} \sup_{0 < t \leq 1} \sup_{0 \leq s < 1/n} |K(s, t)|/q(t) \\ &\stackrel{a.s.}{=} \begin{cases} \alpha(1) & \text{if (8.2.3) holds} \\ O(1) & \text{if (8.2.4) holds,} \end{cases} \end{aligned}$$

where the latter is by (8.2.6) and (8.2.8) respectively. This concludes the proof of Lemma 8.4.1.

As to the proof of Theorem 8.4.1, we can say simply that the proof of Theorem 8.3.1 can be repeated here, since Theorem 8.2.1 and Lemma 8.4.1 imply that there exists a Kiefer process $K(\cdot, \cdot)$ such that, as $n \rightarrow \infty$

$$\sup_{0 < t \leq 1} \sup_{0 \leq s \leq 1} |\alpha_n(U_n(s), t) - n^{-1/2}K(s, nt)|/q(t) = \begin{cases} o_P(1) & \text{if (8.2.3) holds} \\ O_P(1) & \text{if (8.2.4) holds.} \end{cases}$$

However, for the convenience of the reader, we outline some of the details.

Proof of Theorem 8.4.1. We note that

$$\begin{aligned} & \hat{\beta}_n(s, t) \\ &= \begin{cases} n^{-1/2} \left(\sum_{i=1}^{[(n+1)t]} \mathbf{1}\{F(X_i) \leq U_n(s)\} - \frac{[(n+1)t]}{n} \sum_{i=1}^n \mathbf{1}\{F(X_i) \leq U_n(s)\} \right), \\ \quad 0 \leq t < 1, 0 \leq s \leq 1, \\ 0, \quad t = 1, 0 \leq s \leq 1. \end{cases} \end{aligned}$$

We define $S_{[(n+1)t]}^{(i)}$ and $S_{[(n+1)t]}^{(i)}(s)$, $i = 1, 2$, as in proof of Theorem 8.3.1. Then

$$S_{[(n+1)t]}(U_n(s)) = \sum_{i=1}^{[(n+1)t]} (\mathbf{1}\{F(X_i) \leq U_n(s)\} - U_n(s)); \quad 0 \leq t < 1, 0 \leq s \leq 1$$

and

$$S_{[(n+1)t]}^{(1)}(U_n(s)) = \sum_{i=1}^{[(n+1)t]} \mathbf{1}\{F(X_i) \leq U_n(s)\} - U_n(s), \quad 0 \leq t \leq 1/2, \quad 0 \leq s \leq 1$$

$$S_{[(n+1)t]}^{(2)}(U_n(s)) = \sum_{i=[(n+1)t]+1}^n (\mathbf{1}\{F(X_i) \leq U_n(s)\} - U_n(s)), \quad 1/2 < t < 1, \quad 0 \leq s \leq 1.$$

Also

$$S_{[(n+1)t]}(U_n(s)) = \begin{cases} S_{[(n+1)t]}^{(1)}(U_n(s)), & 0 < t \leq 1/2, \quad 0 \leq s \leq 1, \\ S_{[\frac{n+1}{2}]}^{(1)}(U_n(s)) + S_{[\frac{n+1}{2}]}^{(2)}(U_n(s)) - S_{[(n+1)t]}^{(2)}(U_n(s)), & 1/2 < t < 1, \quad 0 \leq s \leq 1. \end{cases}$$

Using Theorem 8.2.1 we can define two independent Kiefer processes $K^{(1)}(\cdot, \cdot)$ and $K^{(2)}(\cdot, \cdot)$ such that

$$(8.4.3) \quad \sup_{0 < t \leq 1/2} \sup_{0 \leq s \leq 1} n^{-1/2} \left| S_{[(n+1)t]}^{(1)}(s) - K^{(1)}(s, (n+1)t) \right| / q(t) \\ = \begin{cases} o_P(1) & \text{if (8.3.5) holds} \\ O_P(1) & \text{if (8.3.6) holds} \end{cases}$$

and

$$(8.4.4) \quad \sup_{0 < t \leq 1/2} \sup_{0 \leq s \leq 1} n^{-1/2} \left| S_{[(n+1)t]}^{(2)}(s) - K^{(2)}(s, (n+1)t) \right| / q(1-t) \\ = \begin{cases} o_P(1) & \text{if (8.3.5) holds} \\ O_P(1) & \text{if (8.3.6) holds} \end{cases}$$

(cf. proof of Theorem 8.3.1). Similarly as in (8.3.10) we define a Kiefer process by

$$\{K(s, nt)\}, \quad 0 \leq t \leq 1, \quad 0 \leq s \leq 1\} \\ = \begin{cases} K^{(1)}(s, nt), & 1 \leq t \leq 1/2, \quad 0 \leq s \leq 1 \\ K^{(1)}(s, \frac{n}{2}) + K^{(2)}(s, \frac{n}{2}) - K^{(2)}(s, n - nt), & 1/2 < t \leq 1, \quad 0 \leq s \leq 1. \end{cases}$$

We have

$$\begin{aligned}
& \sup_{0 < t \leq 1/2} \sup_{0 \leq s \leq 1} |\hat{J}_n(s, t) - n^{-1/2}(K(s, (n+1)t) - tK(s, n+1))|/q(t) \\
&= \sup_{0 < t \leq 1/2} \sup_{0 \leq s \leq 1} |n^{-1/2}(S_{[(n+1)t]}(U_n(s)) - \frac{[(n+1)t]}{n} S_n(U_n(s)) \\
&\quad - n^{-1/2}(K(s, (n+1)t) - tK(s, n+1))|/q(t) \\
&\leq \sup_{0 < t \leq 1/2} \sup_{0 \leq s \leq 1} |n^{-1/2} S_{[(n+1)t]}^{(1)}(U_n(s)) - n^{-1/2} K^{(1)}(s, (n+1)t)|/q(t) \\
&\quad + \sup_{0 < t \leq 1/2} \sup_{0 \leq s \leq 1} |n^{-3/2}[(n+1)t]S_n(U_n(s)) - n^{-1/2}tK(s, n+1)|/q(t) \\
&= I_1^{(1)}(n) + I_2^{(1)}(n).
\end{aligned}$$

Using (8.4.3) and Lemma 8.4.1 we have

$$\begin{aligned}
I_1^{(1)}(n) &= \sup_{0 < t \leq 1/2} \sup_{0 \leq s \leq 1} |n^{-1/2} S_{[(n+1)t]}^{(1)}(U_n(s)) - n^{-1/2} K(s, (n+1)t)|/q(t) \\
&\leq \sup_{0 < t \leq 1/2} \sup_{0 \leq s \leq 1} |n^{-1/2} S_{[(n+1)t]}^{(1)}(U_n(s)) - n^{-1/2} K^{(1)}(U_n(s), (n+1)t)|/q(t) \\
&\quad + \sup_{0 < t \leq 1/2} \sup_{0 \leq s \leq 1} |n^{-1/2} K^{(1)}(U_n(s), (n+1)t) - n^{-1/2} K^{(1)}(s, (n+1)t)|/q(t) \\
&\leq \sup_{0 < t \leq 1/2} \sup_{0 \leq s \leq 1} |n^{-1/2} S_{[(n+1)t]}^{(1)}(s) - n^{-1/2} K^{(1)}(s, (n+1)t)|/q(t) \\
&\quad + \sup_{0 < t \leq 1/2} \sup_{0 \leq s \leq 1} |n^{-1/2} K^{(1)}(U_n(s), (n+1)t) - n^{-1/2} K^{(1)}(s, (n+1)t)|/q(t) \\
&= \begin{cases} o_P(1) & \text{if (8.3.5) holds} \\ O_P(1) & \text{if (8.3.6) holds.} \end{cases}
\end{aligned}$$

When considering $I_2^{(1)}(n)$, we note that, by similar arguments to the ones above, we have, as $n \rightarrow \infty$

$$\sup_{0 \leq s \leq 1} |n^{-1/2} S_n(U_n(s)) - n^{-1/2} K(s, n+1)| = o_P(1)$$

and hence, arguing the same way as in the proof of Theorem 8.3.1 right after (8.3.13), we conclude

$$I_2^{(1)}(n) = o_P(1)$$

for any q positive on $(0,1)$ and such that (8.3.5) or (8.3.6) holds.

For $t \in (1/2, 1)$ we have

$$\begin{aligned} & \sup_{0 < t \leq 1/2} \sup_{0 \leq s \leq 1} |\hat{\beta}_n(s, t) - n^{-1/2} (K(s, (n+1)t) - tK(s, n+1))|/q(t) \\ &= \sup_{1/2 < t < 1} \sup_{0 \leq s \leq 1} |n^{-1/2} (S_{[(n+1)t]}(U_n(s)) - \frac{[(n+1)t]}{n} S_n(U_n(s))) \\ & \quad - n^{-1/2} (K(s, (n+1)t) - tK(s, n+1))|/q(t) \\ &\leq \sup_{1/2 < t < 1} \sup_{0 \leq s \leq 1} |n^{-1/2} S_{[(n+1)t]}^{(2)}(U_n(s)) - n^{-1/2} K^{(2)}(s, (n+1)(1-t))|/q(t) \\ &+ \sup_{1/2 < t \leq 1} \sup_{0 \leq s \leq 1} |n^{-1/2} \left(S_{[\frac{n+1}{2}]}^{(1)}(U_n(s)) + S_{[\frac{n+1}{2}]}^{(1)}(U_n(s)) - \frac{[(n+1)t]}{n} S_n(U_n(s)) \right) \\ & \quad - n^{-1/2} \left(K^{(1)}\left(s, \frac{n+1}{2}\right) + K^{(2)}\left(s, \frac{n+1}{2}\right) - tK(s, n+1) \right)|/q(t) \\ &= I_1^{(2)}(n) + I_2^{(2)}(n). \end{aligned}$$

Using (8.4.4) and Lemma 8.4.1 as in estimating $I_1^{(1)}(n)$, we obtain

$$\begin{aligned} I_1^{(2)}(n) &= \sup_{1/2 < t < 1} \sup_{0 \leq s \leq 1} |n^{-1/2} S_{[(n+1)t]}^{(2)}(U_n(s)) - n^{-1/2} K^{(2)}(s, (n+1)(1-t))|/q(t) \\ &= \sup_{0 < t < 1/2} \sup_{0 \leq s \leq 1} |n^{-1/2} S_{[(n+1)(1-t)]}^{(2)}(U_n(s)) - n^{-1/2} K^{(2)}(s, (n+1)t)|/q(1-t) \\ &= \begin{cases} o_P(1) & \text{if (8.3.5) holds} \\ O_P(1) & \text{if (8.3.6) holds.} \end{cases} \end{aligned}$$

Finally, we have also, for any q positive on $(0,1)$ and such that (8.3.5) or (8.3.6) holds,

$$I_2^{(2)}(n) = o_P(1)$$

by the very same argument as in the proof of Theorem 8.3.1, combined with Lemma 8.4.1. This concludes the proof of (a) and (b) of Theorem 8.4.1. The proof of (c) can be done along the lines of the proof of part (c) of Theorem 8.2.1.

Corollary 8.4.1. *We assume that H_0 holds.*

(a) *If q is a positive function on $(0, 1)$ and such that (8.3.5) holds, then as $n \rightarrow \infty$*

$$(8.4.5) \quad \hat{\beta}_n(s, t)/q(t) \xrightarrow{D} \Gamma(s, t)/q(t)$$

in $D[0, 1]^2$.

(b) *If q is a positive function on $(0, 1)$, then as $n \rightarrow \infty$, (8.4.5) holds in $D([0, 1] \times [\delta, 1 - \delta])$ for any $0 < \delta < 1/2$.*

As in Section 8.3 (cf. Remark 8.3.2), we note also here that, assuming (8.3.5), part (a) of Corollary 8.4.1 gives the limiting distribution of any continuous functional of $\hat{\beta}_n(s, t)/q(t)$. However, when considering sup-functionals, Theorem 8.4.1 (c) gives convergence in distribution results, assuming only (8.3.6). For example, the just mentioned one holds true with

$$q(t) = \left(t(1-t) \log \log \frac{1}{t(1-t)} \right)^{1/2}, \quad 0 < t < 1,$$

but not as a consequence of (8.4.5).

8.5. Weighted Empirical Processes Based on Sequential Ranks

Let $X_1, X_2, \dots$ be independent random variables with an unknown continuous distribution function F . For each $n \geq 1$, let $\xi_1, \dots, \xi_n$ denote the normalized sequential ranks

$$\xi_i = i^{-1} \sum_{j=1}^i 1\{X_j < X_i\}, \quad i = 1, \dots, n,$$

of the first n of the random variables $X_1, X_2, \dots$. Here, we consider the two-time parameter empirical process

$$\gamma_n(s, t) = n^{-1/2} \sum_{i=1}^{[nt]} \mathbf{1}(\{\xi_i \leq s\} - \frac{[is]}{i}), \quad 0 \leq s, t \leq 1,$$

of these sequential ranks ξ_i , $i = 1, \dots, n$.

We note that $\gamma_n(s, t)$ is mean zero process on account of $P\{i\xi_i = k\} = \frac{1}{i}$, $1 \leq i \leq n$, $0 \leq k < i$.

Theorem 8.5.1. *We assume that H_0 holds. Then there exists a sequence of Kiefer processes $K_n(\cdot, \cdot)$ such that with a weight function q which is positive on $(0, 1]$, we have, as $n \rightarrow \infty$*

(a) if

$$(8.5.1) \quad \lim_{t \downarrow 0} (t \log \log 1/t)^{1/2} / q(t) = 0$$

then

$$\sup_{0 < t \leq 1} \sup_{0 \leq s \leq 1} |\gamma_n(s, t) - K_n(s, t)| / q(t) = o_P(1),$$

(b) if

$$(8.5.2) \quad \limsup_{t \downarrow 0} (t \log \log 1/t)^{1/2} / q(t) < \infty$$

then

$$\sup_{0 < t \leq 1} \sup_{0 \leq s \leq 1} |\gamma_n(s, t) - K_n(s, t)| / q(t) = O_P(1),$$

(c) if (8.5.2) holds, then

$$\sup_{0 < t \leq 1} \sup_{0 \leq s \leq 1} |\gamma_n(s, t)| / q(t) \xrightarrow{D} \sup_{0 < t \leq 1} \sup_{0 \leq s \leq 1} |K(s, t)| / q(t)$$

where $K(\cdot, \cdot)$ is a Kiefer process.

We note that Remark 8.2.1 is also valid here.

In order to prove our result, we will represent the normalized sequential ranks ξ_i in terms of uniformly distributed random variables. Since $i\xi_i$ are independent, the i^{th} being uniformly distributed on $\{0, \dots, i-1\}$, we have

$$(8.5.3) \quad \xi_i \stackrel{D}{=} \frac{[iU_i]}{i}, \quad i = 1, 2, \dots,$$

with independent random variables U_i , uniformly distributed on $[0,1]$ (cf. Chapter 4).

Proof of Theorem 8.5.1. Let $U_i = F(X_i)$, and

$$\tilde{S}_{nt}(s) = \sum_{i=1}^{[nt]} (1\{\frac{[iU_i]}{i} \leq s\} - s)$$

and

$$\tilde{\gamma}_n(s, t) = n^{-1/2} \tilde{S}_{nt}(s).$$

By Csörgő and Horváth (1987b), one can construct a Kiefer process $K(\cdot, \cdot)$ such that

$$(8.5.4) \quad \sup_{0 \leq s \leq 1} |\tilde{S}_{nt}(s) - K(s, nt)| = O(\log^2 nt),$$

as $nt \rightarrow \infty$. Hence, imitating the proof of Theorem 8.2.1 we get

$$(8.5.5) \quad \sup_{0 < t \leq 1} \sup_{0 \leq s \leq 1} |\tilde{\gamma}_n(s, t) - n^{-1/2} K(s, nt)|/q(t) = o_P(1),$$

if q is a positive function on $(0,1]$, such that (8.5.1) holds. If q is a positive function on $(0,1]$ such that (8.5.2) holds, then we have

$$(8.5.6) \quad \sup_{0 < t \leq 1} \sup_{0 \leq s \leq 1} |\tilde{\gamma}_n(s, t) - n^{-1/2} K(s, nt)|/q(t) = O_P(1)$$

and

$$(8.5.7) \quad \sup_{0 < t \leq 1} \sup_{0 \leq s \leq 1} |\tilde{\gamma}_n(s, t)|/q(t) \xrightarrow{D} \sup_{0 < t \leq 1} \sup_{0 \leq s \leq 1} |K(s, t)|/q(t).$$

Now, (8.5.5) and (8.5.6) can be reformulated by saying that there exists a sequence of Kiefer processes $\tilde{K}_n(s, t)$, namely $\tilde{K}_n(s, t) = n^{-1/2} K(s, nt)$, such that if (8.5.1) holds, then

$$\sup_{0 < t \leq 1} \sup_{0 \leq s \leq 1} |\tilde{\gamma}_n(s, t) - \tilde{K}_n(s, t)|/q(t) = o_P(1)$$

and if (8.5.2) holds, then

$$\sup_{0 < t \leq 1} \sup_{0 \leq s \leq 1} |\tilde{\gamma}_n(s, t) - \tilde{K}_n(s, t)|/q(t) = O_P(1).$$

Since, by (8.5.3),

$$(8.5.8) \quad \{\tilde{\gamma}_n(s, t), 0 \leq s \leq 1, 0 \leq t \leq 1\} \stackrel{D}{=} \{\tilde{\gamma}_n(s, t), 0 \leq s \leq 1, 0 \leq t \leq 1\},$$

where

$$\tilde{\gamma}_n(s, t) = n^{-1/2} \sum_{i=1}^{[nt]} (1\{\xi_i \leq s\} - s), \quad 0 \leq s, t \leq 1,$$

we can use Lemma 3.1.2 in Csörgő (1983) (cf. Section A.1) to conclude parts (a) and (b) with $\tilde{\gamma}_n(s, t)$ instead of $\gamma_n(s, t)$. Since, as $n \rightarrow \infty$

$$(8.5.9) \quad \sup_{0 < t \leq 1} \sup_{0 \leq s \leq 1} |\tilde{\gamma}_n(s, t) - \gamma_n(s, t)|/q(t) = o(1)$$

for any q positive on $(0, 1]$ and such that (8.5.2) is satisfied, (a) and (b) hold true.

In order to see (8.5.9) we note that the difference between $\tilde{\gamma}_n(s, t) - \gamma_n(s, t)$ is equal to

$$n^{-1/2} \sum_{i=1}^{[nt]} \left| \frac{[is]}{i} - s \right| = n^{-1/2} \sum_{i=1}^{[nt]} \left| \frac{[is] - is}{i} \right|$$

and we have

$$\sup_{0 \leq t \leq 1} \sup_{0 \leq s \leq 1} n^{-1/2} \sum_{i=1}^{[nt]} \left| \frac{[is] - is}{i} \right| \leq n^{-1/2} \log n,$$

which goes to 0 as $n \rightarrow \infty$, and for $nt \rightarrow \infty$ we have

$$\sum_{i=1}^{[nt]} \left| \frac{[is] - is}{i} \right| = O(\log nt) \text{ for any } s \in [0, 1],$$

which gives

$$\sup_{1/n \leq t \leq 1} \sup_{0 \leq s \leq 1} n^{-1/2} \sum_{i=1}^{[nt]} \left| \frac{[is] - is}{i} \right| / t^{1/2} = O(1).$$

Hence for any $\delta \in (0, 1)$ we get

$$\sup_{\delta \leq t \leq 1} \sup_{0 \leq s \leq 1} n^{-1/2} \sum_{i=1}^{[nt]} \left| \frac{[is] - is}{i} \right| / q(t) = o(1)$$

for q positive on $(0, 1]$, and

$$\begin{aligned} & \sup_{1/n \leq t < \delta} \sup_{0 \leq s \leq 1} n^{-1/2} \sum_{i=1}^{[nt]} \left| \frac{[is] - is}{i} \right| / q(t) \\ & \leq O(1) \sup_{0 < t < \delta} t^{1/2} / q(t). \end{aligned}$$

Consequently, taking $\delta > 0$ arbitrarily small we obtain

$$\sup_{1/n \leq t \leq 1} \sup_{0 \leq s \leq 1} n^{-1/2} \sum_{i=1}^{[nt]} \left| \frac{[is] - is}{i} \right| / q(t) = o(1)$$

for any q positive on $(0, 1]$ and such that $\lim_{t \downarrow 0} q(t)/t^{1/2} = \infty$, in particular for q such that (8.5.1) or (8.5.2) holds.

Obviously (8.5.7) together with (8.5.8) and (8.5.9) implies part (c).

Corollary 8.5.1 *We assume that H_0 holds. Let $\{K(s, t), 0 \leq s \leq 1, 0 \leq t \leq 1\}$ be a Kiefer process. Then we have, as $n \rightarrow \infty$*

(a) if q is a positive function on $(0, 1]$ and such that (8.5.1) holds, then

$$(8.5.10) \quad \gamma_n(s, t)/q(t) \xrightarrow{\mathcal{D}} K(s, t)/q(t)$$

in $D[0, 1]^2$.

(b) if q is a positive function on $(0, 1]$, then (8.5.10) holds in $D([0, 1] \times [\delta, 1])$ for any $0 < \delta < 1$.

8.6. Two-time parameter empirical processes of sequential ranks tied down at $t = 1$.

For the sake of completeness and theoretical interest we consider also "bridge" type empirical processes of sequential ranks. In order to study the asymptotic behaviour of the process

$$(\gamma_n(s, t) - t\gamma_n(s, 1))/q(t),$$

we introduce

$$\hat{\gamma}_n(s, t) = \begin{cases} n^{-1/2} \left(\sum_{i=1}^{[(n+1)t]} 1\{\xi_i \leq \frac{[is]}{i}\} - \frac{[(n+1)t]}{n} \sum_{i=1}^n 1\{\xi_i \leq \frac{[is]}{i}\} \right), & 0 < t \leq 1, \\ & 0 \leq s \leq 1, \\ 0 & , t = 1. \end{cases}$$

Let $\{\Gamma(s, t), 0 \leq s \leq 1, 0 \leq t \leq 1\}$ denote a Gaussian process with mean zero and covariance function (8.3.4).

Theorem 8.6.1. We assume that H_0 holds. Then there exists a sequence of stochastic processes $\Gamma_n, n \geq 1$,

$$\{\Gamma_n(s, t), 0 \leq s \leq 1, 0 \leq t \leq 1\} \xrightarrow{\mathcal{D}} \{\Gamma(s, t), 0 \leq s \leq 1, 0 \leq t \leq 1\}$$

such that with a weight function q which is positive on $(0, 1)$, we have as $n \rightarrow \infty$

(a) if

$$(8.6.1) \quad \left\{ \begin{array}{l} \lim_{t \downarrow 0} (t(1-t) \log \log 1/(t(1-t)))^{1/2}/q(t) = 0 \\ \text{and} \\ \lim_{t \uparrow 1} (t(1-t) \log \log 1/(t(1-t)))^{1/2}/q(t) = 0 \end{array} \right.$$

then

$$\sup_{0 < t < 1} \sup_{0 \leq s \leq 1} |\hat{\gamma}_n(s, t) - \Gamma_n(s, t)|/q(t) = o_P(1).$$

(b) if

$$(8.6.2) \quad \left\{ \begin{array}{l} \lim_{t \downarrow 0} (t(1-t) \log \log 1/(t(1-t)))^{1/2}/q(t) < \infty \\ \text{and} \\ \lim_{t \uparrow 1} (t(1-t) \log \log 1/(t(1-t)))^{1/2}/q(t) < \infty \end{array} \right.$$

then

$$\sup_{0 < t < 1} \sup_{0 \leq s \leq 1} |\hat{\gamma}_n(s, t) - \Gamma_n(s, t)|/q(t) = O_P(1)$$

(c) If (8.6.2) holds, then

$$\sup_{0 < t < 1} \sup_{0 \leq s \leq 1} |\hat{\gamma}_n(s, t)|/q(t) \xrightarrow{P} \sup_{0 < t < 1} \sup_{0 \leq s \leq 1} |\Gamma(s, t)|/q(t).$$

Proof: As in the proof of Theorem 8.5.1, let

$$\tilde{S}_{[(n+1)t]}(s) = \sum_{i=1}^{[(n+1)t]} (1\{\frac{[iU_i]}{i} \leq s\} - s).$$

We introduce also

$$S_{[(n+1)t]}^{(1)}(s) = \sum_{i=1}^{[(n+1)t]} (1\{\frac{[iU_i]}{i} \leq s\} - s), \quad 0 \leq t \leq 1/2, \quad 0 \leq s \leq 1,$$

$$\tilde{S}_{[(n+1)t]}^{(2)}(s) = \sum_{i=[(n+1)t]+1}^n (1\{\frac{[iU_i]}{i} \leq s\} - s), \quad 1/2 \leq t < 1, \quad 0 \leq s \leq 1.$$

Then

$$\tilde{S}_{[(n+1)t]^{1/n}} = \begin{cases} \tilde{S}^{(1)}(s), & 0 \leq t \leq 1/2, \quad 0 \leq s \leq 1 \\ \tilde{S}_{[\frac{n+1}{2}]}^{(1)}(s) + \tilde{S}_{[\frac{n+1}{2}]}^{(2)}(s) - \tilde{S}_{[(n+1)t]}^{(2)}(s), & 1/2 \leq t < 1, \quad 0 \leq s \leq 1. \end{cases}$$

Using representation of sequential ranks $\xi_i \stackrel{D}{=} [i\mathcal{U}_i]/i$ (cf. (8.5.3)) and the approximation of $\tilde{S}_{nt}$ (cf. (8.5.4)) one can construct two independent Kiefer processes $K^{(1)}(\cdot, \cdot)$ and $K^{(2)}(\cdot, \cdot)$ such that

$$(8.6.3) \quad \sup_{0 < t < 1} \sup_{0 \leq s \leq 1} n^{-1/2} \left| \tilde{S}_{[(n+1)t]}^{(1)}(s) - K^{(1)}(s, (n+1)t) \right| / q(t) \\ = \begin{cases} o_P(1) & \text{if (8.6.1) holds} \\ O_P(1) & \text{if (8.6.2) holds,} \end{cases}$$

and

$$(8.6.4) \quad \sup_{0 < t < 1} \sup_{0 \leq s \leq 1} n^{-1/2} \left| \tilde{S}_{[(n+1)t]}^{(2)}(s) - K^{(2)}(s, (n+1)t) \right| / q(1-t) \\ = \begin{cases} o_P(1) & \text{if (8.6.1) holds} \\ O_P(1) & \text{if (8.6.2) holds} \end{cases}$$

In order to get (8.6.4) we note that, along the lines of the proof of Theorem 2.1 in Csörgő and Horváth (1987b), we get, as $nt \rightarrow \infty$

$$\sup_{0 \leq s \leq 1} \left| \sum_{i=1}^{[(n+1)t]} \mathbf{1} \left\{ \frac{[(n-i+1)\mathcal{U}_{n-i+1}]}{n-i+1} \leq s \right\} - \sum_{i=1}^{[(n+1)t]} \mathbf{1} \{ U_{n-i+1} \leq s \} \right| = O(\log nt).$$

By this and (8.5.4), we have, as $nt \rightarrow \infty$

$$\sup_{0 \leq s \leq 1} \left| \sum_{i=1}^{[(n+1)t]} \mathbf{1} \left\{ \frac{[(n-i+1)\mathcal{U}_{n-i+1}]}{n-i+1} \leq s \right\} - K^{(2)}(s, (n+1)t) \right| = O(\log^2 nt),$$

Hence, and using also (8.3.9) with function q positive on $(0, 1)$, we have (cf. proofs of Theorems 8.2.1 and 8.5.1),

$$\sup_{1/2 < t < 1} \sup_{0 \leq s \leq 1} \left| n^{-1/2} \tilde{S}_{[(n+1)t]}^{(2)}(s) - n^{-1/2} K^{(2)}(s, (n+1)(1-t)) \right| / q(t) \\ = \sup_{0 < t < 1/2} \sup_{0 \leq s \leq 1} \left| n^{-1/2} \tilde{S}_{[(n+1)t]}^{(2)}(s) - n^{-1/2} K^{(2)}(s, (n+1)t) \right| / q(1-t) \\ = \begin{cases} o_P(1) & \text{if (8.6.1) holds} \\ O_P(1) & \text{if (8.6.2) holds.} \end{cases}$$

Using (8.6.3) and (8.6.4), as well as the method of proof of Theorem 8.3.1, with the Kiefer process

$$\begin{aligned} & \{K(s, nt), 0 \leq t \leq 1, 0 \leq s \leq 1\} \\ &= \begin{cases} K^{(1)}(s, nt), & 0 \leq t \leq 1/2, 0 \leq s \leq 1 \\ K^{(1)}\left(s, \frac{n}{2}\right) + K^{(2)}\left(s, \frac{n}{2}\right) - K^{(2)}(s, n - nt), & 1/2 < t \leq 1, 0 \leq s \leq 1, \end{cases} \end{aligned}$$

we have

$$\begin{aligned} & \sup_{0 < t \leq 1/2} \sup_{0 \leq s \leq 1} |\hat{\gamma}_n(s, t) - n^{-1/2}(K(s, (n+1)t) - tK(s, n+1))|/q(t) \\ &= \sup_{0 < t \leq 1/2} \sup_{0 \leq s \leq 1} |n^{-1/2}(\tilde{S}_{[(n+1)t]}(s) - \frac{[(n+1)t]}{n} \tilde{S}_n(s)) \\ & \quad - n^{-1/2}(K(s, (n+1)t) - tK(s, n+1))|/q(t) \\ &= \begin{cases} o_P(1) & \text{if (8.6.1) holds} \\ O_P(1) & \text{if (8.6.2) holds,} \end{cases} \end{aligned}$$

and also

$$\begin{aligned} & \sup_{1/2 < t < 1} \sup_{0 \leq s \leq 1} \left| \hat{\gamma}_n(s, t) - n^{-1/2}(K(s, (n+1)t) - tK(s, n+1)) \right|/q(t) \\ &= \sup_{1/2 < t < 1} \sup_{0 \leq s \leq 1} \left| n^{-1/2} \left(\tilde{S}_{[(n+1)t]}(s) - \frac{[(n+1)t]}{n} \tilde{S}_n(s) \right) \right. \\ & \quad \left. - n^{-1/2}(K(s, (n+1)t) - tK(s, n+1)) \right|/q(t) \\ &= \begin{cases} o_P(1) & \text{if (8.6.1) holds} \\ O_P(1) & \text{if (8.6.2) holds.} \end{cases} \end{aligned}$$

Consequently, with the sequence of stochastic processes Γ_n , $n \geq 1$,

$$\{\Gamma_n(s, t), 0 \leq s \leq 1, 0 \leq t \leq 1\} = \{n^{-1/2}(K(s, nt) - tK(s, n)), 0 \leq s \leq 1, 0 \leq t \leq 1\}$$

we have, as $n \rightarrow \infty$

$$\sup_{0 < t < 1} \sup_{0 \leq s \leq 1} \left| \hat{\gamma}_n(s, t) - \Gamma_n(s, t) \right|/q(t) = \begin{cases} o_P(1) & \text{if (8.6.1) holds} \\ O_P(1) & \text{if (8.6.2) holds.} \end{cases}$$

We note again that for each $n \geq 1$

$$\{\Gamma_n(s, t), 0 \leq s \leq 1, 0 \leq t \leq 1\} \stackrel{\mathcal{D}}{=} \{\Gamma(s, t), 0 \leq s \leq 1, 0 \leq t \leq 1\},$$

where $\Gamma(s, t)$ is a Gaussian process with mean zero and covariance as in (8.3.4). Since by (8.5.3)

$$\{\hat{\gamma}_n(s, t), 0 \leq s \leq 1, 0 \leq t \leq 1\} \stackrel{\mathcal{D}}{=} \{\hat{\gamma}_n(s, t), 0 \leq s \leq 1, 0 \leq t \leq 1\},$$

we can use Lemma 3.1.2 in Csörgő (1983) (cf. Section A.1) to conclude that, with a possibly different sequence of stochastic processes $\bar{\Gamma}_n(\cdot, \cdot)$, but still such that

$$\{\bar{\Gamma}_n(s, t), 0 \leq s \leq 1, 0 \leq t \leq 1\} \stackrel{\mathcal{D}}{=} \{\Gamma(s, t), 0 \leq s \leq 1, 0 \leq t \leq 1\}, \quad n \geq 1,$$

we have, as $n \rightarrow \infty$

$$\sup_{0 < t < 1} \sup_{0 \leq s \leq 1} |\hat{\gamma}_n(s, t) - \bar{\Gamma}_n(s, t)| / q(t) = \begin{cases} o_P(1) & \text{if (8.6.1) holds} \\ O_P(1) & \text{if (8.6.2) holds.} \end{cases}$$

where

$$\hat{\gamma}_n(s, t) = \begin{cases} n^{-1/2} \left(\sum_{i=1}^{[(n+1)t]} \mathbf{1}\{\xi_i \leq s\} - \frac{[(n+1)t]}{n} \sum_{i=1}^n \mathbf{1}\{\xi_i \leq s\} \right), & 0 \leq t < 1, \\ 0 & 0 \leq s \leq 1, \\ & t = 1. \end{cases}$$

This gives parts (a) and (b) of theorem, since

$$(8.6.5) \quad \sup_{0 < t < 1} \sup_{0 \leq s \leq 1} |\hat{\gamma}_n(s, t) - \bar{\gamma}_n(s, t)| / q(t) = o(1)$$

for any q positive on $(0, 1)$ and such that (8.6.2) holds. The proof of (8.6.5) can be done similarly to that of (8.5.9). The proof of (c) can be done along the lines of the proof of part (c) of Theorem 8.2.1.

8.7. Contiguous Alternatives

In this Section we prove that weak convergence of the two-time parameter empirical processes $\alpha_n(s, t)$, $\hat{\beta}_n(s, t)$ and $\gamma_n(s, t)$, as well as of the "bridge" type processes $\hat{\alpha}_n(s, t)$ and $\hat{\gamma}_n(s, t)$ with weight functions in their time parameter, continues to hold under the sequence of contiguous alternatives of H_1 as parametrized by (8.0.1)–(8.0.3) for the same class of weight functions as in the case of the null hypothesis H_0 .

Let

$$(8.7.1) \quad c(s, t) = \int_0^t \int_0^s g(\tau, u) \, du d\tau$$

and

$$(8.7.2) \quad r(s, t) = - \int_0^t \frac{c(s, \tau)}{\tau} \, d\tau + c(s, t),$$

where $g(\cdot, \cdot)$ is the function as given in the parametrization of H_1 .

Theorem 8.7.1. *We assume that H_1 holds.*

(a) *Let q be a positive function on $(0, 1]$ such that*

$$(8.7.3) \quad \lim_{t \downarrow 0} (t \log \log 1/t)^{1/2} / q(t) = 0.$$

Then, as $n \rightarrow \infty$,

$$(8.7.4) \quad \alpha_n(s, t) / q(t) \xrightarrow{D} (K(s, t) + c(s, t)) / q(t)$$

and

$$(8.7.5) \quad \gamma_n(s, t) / q(t) \xrightarrow{D} (K(s, t) + r(s, t)) / q(t)$$

in $D[0, 1]^2$, where $\{K(s, t), 0 \leq s \leq 1, 0 \leq t \leq 1\}$ is a Kiefer process.

(b) Let q be a positive function on $(0, 1]$. Then, as $n \rightarrow \infty$, (8.7.4) and (8.7.5) hold in $D([0, 1] \times [\delta, 1])$ for any $0 < \delta < 1$.

Let

$$(8.7.6) \quad d(s, t) = c(s, t) - tc(s, 1),$$

where $c(s, t)$ is defined by (8.7.1), i.e.,

$$d(s, t) = \int_0^t \int_0^s g(\tau, u) du d\tau - t \int_0^1 \int_0^s g(\tau, u) du d\tau,$$

and let

$$(8.7.7) \quad e(s, t) = r(s, t) - tr(s, 1),$$

where $r(s, t)$ is defined by (8.7.2), i.e.

$$e(s, t) = - \int_0^t \frac{c(s, \tau)}{\tau} d\tau + t \int_0^1 \frac{c(s, \tau)}{\tau} d\tau + c(s, t) - tc(s, 1).$$

Theorem 8.7.2. We assume that H_1 holds.

(a) Let q be positive on $(0, 1)$ and such that

$$(8.7.8) \quad \begin{cases} \lim_{t \downarrow 0} (t(1-t) \log \log 1/(t(1-t)))^{1/2} / q(t) = 0 \\ \text{and} \\ \lim_{t \uparrow 1} (t(1-t) \log \log 1/(t(1-t)))^{1/2} / q(t) = 0. \end{cases}$$

Then we have, as $n \rightarrow \infty$

$$(8.7.9) \quad \hat{\alpha}_n(s, t) / q(t) \xrightarrow{D} (\Gamma(s, t) + d(s, t)) / q(t),$$

$$(8.7.10) \quad \hat{\beta}_n(s, t) / q(t) \xrightarrow{D} (\Gamma(s, t) + d(s, t)) / q(t),$$

$$(8.7.11) \quad \hat{\gamma}_n(s, t) / q(t) \xrightarrow{D} (\Gamma(s, t) + e(s, t)) / q(t),$$

in $D(0, 1)^2$, where $\{\Gamma(s, t), 0 \leq s \leq 1, 0 \leq t \leq 1\}$ is a Gaussian process with mean zero and covariance as in (8.3.4).

(b) Let q be positive on $(0, 1)$. Then (8.7.9), (8.7.10) and (8.7.11) hold true in $D([0, 1] \times [\delta, 1 - \delta])$ for any $0 < \delta < 1/2$.

In order to prove our results, we assume without loss of generality that under H_0 the X_i are of the form $F^{-1}(U_i)$ where the U_i are independent, identically distributed uniform $- (0, 1)$ random variables, and so $X_i \stackrel{D}{=} F^{-1}(U_i)$, $i = 1, 2, \dots$.

Define the centered log-likelihood ratio process $L_n(\cdot)$ by

$$(8.7.12) \quad L_n(t) = \sum_{i=1}^{[nt]} \left\{ \log \frac{dF_{in}}{dF}(F^{-1}(U_i)) - E \log \frac{dF_{in}}{dF}(F^{-1}(U_i)) \right\},$$

and let $L(t)$ be a Gaussian process with mean zero and covariance function $Q_L(t_1 \wedge t_2)$, where

$$(8.7.13) \quad Q_L(t) = \int_0^t \int_0^1 g^2(s, u) \, du \, ds.$$

We need to introduce the following Gaussian processes:

(i) $(K(s, t), L(t))$, where $K(s, t)$ is a Kiefer process, i.e. the process defined by (8.2.1), and the covariance function of the process (K, L) is

$$Q_{K,L}(s, t_1, t_2) = EK(s, t_1)L(t_2) = c(s, t_1 \wedge t_2),$$

where $c(s, t)$ is defined by (8.7.1),

(ii) $(K^*(s, t), L(t))$, where $K^*(s, t)$ is again a Kiefer process, but the covariance function of the process (K^*, L) is

$$Q_{K^*,L}(s, t_1, t_2) = EK^*(s, t_1)L(t_2) = - \int_0^{t_1} \frac{c(s, \tau \wedge t_2)}{\tau} d\tau + c(s, t_1 \wedge t_2).$$

Lemma 8.7.1. We assume that H_0 holds. Then, as $n \rightarrow \infty$

$$(a) \quad (\alpha_n, L_n) \xrightarrow{D} (K, L),$$

$$(b) \quad (\gamma_n, L_n) \xrightarrow{D} (K^*, L),$$

in $D[0, 1]^2 \times D[0, 1]$.

Proof of Lemma 8.7.1. (a) Let $a(u) \in L^2(0, 1)$, where $\int_0^1 a(u) du = 0$ and $\int_0^1 a^2(u) du = 1$. Then $a(F(X_i))$ are i.i.d.r.v.'s with mean 0 and variance 1, and it follows from Lemma 2.4.1. that

$$(n^{-1/2} \sum_{i=1}^{[nt]} a(F(X_i)), L_n(t)) \xrightarrow{D} (W(t), L(t))$$

in $D[0, 1] \times D[0, 1]$. Here (W, L) is a two-dimensional Gaussian process, where W is a standard Wiener process, L is as in (8.7.13) and the covariance function of the process (W, L) is

$$EW(t_1)L(t_2) = \int_0^{t_1 \wedge t_2} \int_0^1 g(\tau, u)a(u) du d\tau.$$

In particular, the function $a(u)$ can be taken as

$$a(u) = a_s(u) = (\mathbf{1}\{u \leq s\} - s)/(s(1-s))^{1/2},$$

for any fixed $0 \leq s \leq 1$, and we note that for any fixed $0 \leq s \leq 1$

$$W(t) = \frac{K(s, t)}{(s(1-s))^{1/2}}, \quad t \geq 0$$

is a Wiener process. Consequently, we have

$$\begin{aligned} EK(s, t_1)L(t_2) &= (s(1-s))^{1/2} E \frac{K(s, t_1)}{(s(1-s))^{1/2}} L(t_2) \\ &= (s(1-s))^{1/2} \int_0^{t_1 \wedge t_2} \int_0^1 g(\tau, u) \frac{(\mathbf{1}\{u \leq s\} - s)}{(s(1-s))^{1/2}} du d\tau \\ &= \int_0^{t_1 \wedge t_2} \int_0^s g(\tau, u) du d\tau - s \int_0^{t_1 \wedge t_2} \int_0^1 g(\tau, u) du d\tau \\ &= \int_0^{t_1 \wedge t_2} \int_0^s g(\tau, u) du d\tau, \end{aligned}$$

due to (8.0.5). Hence, we conclude that all finite-dimensional distributions of the process $(\alpha_n(s, t), L_n(t))$ converge to those of the process $(K(s, t), L(t))$ with covariance function

$$EK(s, t_1)L(t_2) = \int_0^{t_1 \wedge t_2} \int_0^s g(\tau, u) du d\tau.$$

Tightness of $(\alpha_n(s, t), L_n(t))$ follows from the weak convergence of $\alpha_n(s, t)$ (cf., e.g., Theorem 8.2.1 (a), or (8.2.19) in Corollary 8.2.1 with $q \equiv 1$) and that of $L_n(t)$ (cf. Lemma 2.3.1).

(b) Using Theorem 3 in Khmaladze and Parjanadze (1986) (cf. (4.2.4)–(4.2.7) here), again with the function $a(u) = a_s(u) = (1\{u \leq s\} - s)/(s(1 - s))^{1/2}$, and Theorem 8.5.1 (a), or (8.5.10) in Corollary 8.5.1 with $q \equiv 1$, the proof is similar to that of part (a).

Proof of Theorem 8.7.1. (a) Since (α_n, L_n) converges weakly to (K, L) when we assume H_0 to hold, Le Cam's third lemma implies that, under contiguous alternatives H_1 ,

$$(8.7.14) \quad \alpha_n(s, t) \xrightarrow{D} K(s, t) + c(s, t)$$

in $D[0, 1]^2$ (cf., e.g., proof of Theorem 2.4.1). Now using (8.2.19) in Corollary 8.2.1 (a) and the fact that

$$E \frac{\alpha_n(s, t_1)}{q(t_1)} L_n(t_2) = \frac{1}{q(t_1)} E \alpha_n(s, t_1) L_n(t_2) \rightarrow \frac{1}{q(t_1)} c(s, t_1 \wedge t_2),$$

we conclude (8.7.4), and note also that the covariance function of $\alpha_n(s, t)/q(t)$ and $L_n(t)$ is seen to be finite by similar arguments as in the proof of Theorem 2.5.1.

Namely, since $g \in L^2[0, 1]^2$, we have

$$\begin{aligned}
 |c(s, t)| &= \left| \int_0^t \int_0^s g(\tau, u) du d\tau \right| \\
 &\leq \int_0^t \left(\int_0^s g^2(\tau, u) du \right)^{1/2} \left(\int_0^s du \right)^{1/2} d\tau \\
 &= s^{1/2} \int_0^t \left(\int_0^s g^2(\tau, u) du \right)^{1/2} d\tau \\
 &\leq s^{1/2} \left(\int_0^t \int_0^s g^2(\tau, u) du d\tau \right)^{1/2} \left(\int_0^t d\tau \right)^{1/2} \\
 &= s^{1/2} t^{1/2} \left(\int_0^t \int_0^s g^2(\tau, u) du d\tau \right)^{1/2},
 \end{aligned}$$

which implies

$$\sup_{0 < t \leq 1} |c(s, t)|/q(t) \leq s^{1/2} \sup_{0 < t \leq 1} \left(\int_0^t \int_0^s g^2(\tau, u) du d\tau \right)^{1/2} \sup_{0 < t \leq 1} t^{1/2}/q(t) < \infty$$

due the fact that q is positive on $(0, 1]$ and $\lim_{t \downarrow 0} t^{1/2}/q(t) = 0$.

The proof of (8.7.5) is similar, where the finiteness of the covariance function is shown similarly as in the proof of Theorem 4.2.1.

(b) By (8.7.18) and the Skorohod-Dudley-Wichura theorem (cf. Theorem A.2.1), there exists $\alpha_n^*(s, t)$ and $K^*(s, t)$ such that

$$\alpha_n^* \stackrel{D}{=} \alpha_n, \quad n \geq 1 \quad \text{and} \quad K^* \stackrel{D}{=} K + c(s, t)$$

and, under the contiguous alternatives H_1 ,

$$\sup_{0 \leq s \leq 1} \sup_{0 \leq t \leq 1} |\alpha_n^*(s, t) - K^*(s, t)| = o_P(1).$$

Hence for any function q such that $\inf_{0 \leq t \leq 1} q(t) > 0$, under H_1 we have

$$\sup_{0 \leq s \leq 1} \sup_{0 \leq t \leq 1} |\alpha_n^*(s, t) - K^*(s, t)|/q(t) = o_P(1)$$

for any $0 < \delta < 1$. The proof for $\gamma_n(s, t)$ is similar.

In order to prove Theorem 8.7.2, we need to introduce the following Gaussian processes:

(i) $(\Gamma(s, t), L(t))$, where $\Gamma(s, t)$ is a Gaussian process with mean zero and covariance function as in (8.3.4) and the covariance function of the process (Γ, L) is

$$Q_{\Gamma, L}(s, t_1, t_2) = E\Gamma(s, t_1)L(t_2) = c(s, t_1 \wedge t_2) - t_1 c(s, t_2),$$

where $c(s, t)$ is defined by (8.7.1).

(ii) $(\Gamma^*(s, t), L(t))$, where $\Gamma^*(s, t)$ is again a Gaussian process with mean zero and covariance as in (8.3.4) and the covariance function of the process (Γ^*, L) is

$$\begin{aligned} Q_{\Gamma^*, L}(s, t_1, t_2) &= E\Gamma^*(s, t_1)L(t_2) \\ &= - \int_0^{t_1} \frac{c(s, \tau \wedge t_2)}{\tau} d\tau + c(s, t_1 \wedge t_2) + t_1 \int_0^1 \frac{c(s, \tau \wedge t_2)}{\tau} d\tau - t_1 c(s, t_2), \end{aligned}$$

where $c(s, t)$ is defined by (8.7.1).

Lemma 8.7.2. *We assume that H_0 holds. Then, as $n \rightarrow \infty$*

$$(8.7.15) \quad (\hat{\alpha}_n, L_n) \rightarrow (\Gamma, L)$$

$$(8.7.16) \quad (\hat{\beta}_n, L_n) \rightarrow (\Gamma, L)$$

$$(8.7.17) \quad (\hat{\gamma}_n, L_n) \rightarrow (\Gamma^*, L)$$

in $D[0, 1]^2 \times D[0, 1]$.

Proof: The weak convergence statements of (8.7.15) and (8.7.17) follow from Lemma 8.7.1. As to (8.7.16), we note that

$$\begin{aligned} & \sup_{0 \leq t \leq 1} \sup_{0 \leq s \leq 1} |\hat{\beta}_n(s, t) - \hat{\alpha}_n(s, t)|/q(t) \\ &= \sup_{0 \leq t \leq 1} \sup_{0 \leq s \leq 1} |\hat{\alpha}_n(U_n(s), t) - \hat{\alpha}_n(s, t)|/q(t) \\ &= o_p(1). \end{aligned}$$

Proof of Theorem 8.7.2. Similar to that of Theorem 8.7.1 when using Lemma 8.7.2.

8.8. Changepoint alternatives

Corollary 8.8.1. *Under the changepoint alternatives (8.0.7), we have the following weak convergence results.*

(a) *Let q be a positive function on $(0, 1]$ and such that (8.7.3) holds. Then, as $n \rightarrow \infty$*

$$(8.8.1) \quad \alpha_n(s, t)/q(t) \xrightarrow{\mathcal{D}} (K(s, t) + c_\eta(s, t))/q(t),$$

$$(8.8.2) \quad \gamma_n(s, t)/q(t) \xrightarrow{\mathcal{D}} (K(s, t) + r_\eta(s, t))/q(t)$$

in $D[0, 1]^2$, where

$$c_\eta(s, t) = (t - \eta)1\{t \geq \eta\} \int_0^s g(u) du$$

and

$$r_\eta(s, t) = \eta \ln(t/\eta)1\{t \geq \eta\} \int_0^s g(u) du$$

(b) *Let q be a positive function on $(0, 1]$. Then, as $n \rightarrow \infty$, we have (8.8.1) and (8.8.2) in $D([0, 1] \times [\delta, 1])$ for any $0 < \delta < 1$.*

Corollary 8.8.2. *We assume the changepoint alternative of (8.0.7) to hold.*

(a) *Let q be positive on $(0, 1)$ and such that (8.7.8) holds. Then, as $n \rightarrow \infty$*

$$(8.8.3) \quad \hat{\alpha}_n(s, t)/q(t) \xrightarrow{\mathcal{D}} (\Gamma(s, t) + d_\eta(s, t))/q(t)$$

$$(8.8.4) \quad \hat{\beta}_n(s, t)/q(t) \xrightarrow{\mathcal{D}} (\Gamma(s, t) + d_\eta(s, t))/q(t)$$

$$(8.8.5) \quad \hat{\gamma}_n(s, t)/q(t) \xrightarrow{\mathcal{D}} (\Gamma(s, t) + e_\eta(s, t))/q(t)$$

in $D[0, 1]^2$, where

$$d_\eta(s, t) = -t(1 - \eta)\mathbf{1}\{t < \eta\} \int_0^s g(u)du - (1 - t)\eta\mathbf{1}\{t \geq \eta\} \int_0^s g(u)du$$

and

$$\epsilon_\eta(s, t) = (t\eta \log \eta \mathbf{1}\{t < \eta\} + \eta \log t/\eta + t\eta \log \eta) \int_0^s g(u)du$$

(b) Let q be positive on $(0, 1)$. Then (8.8.3), (8.8.4) and (8.8.5) hold true in $D([0, 1] \times [\delta, 1 - \delta])$ for any $0 < \delta < 1/2$.

8.9. Weighted uniform empirical processes under contiguous measures

We consider again the uniform empirical process $\alpha_n(s)$ as defined in Section 8.1, namely

$$\alpha_n(s) = n^{-1/2} \sum_{i=1}^n \left(\mathbf{1}\{F(X_i) \leq s\} - s \right), \quad 0 \leq s \leq 1,$$

where F is assumed to be continuous. Theorem 8.1.C describes the asymptotic behaviour of $\alpha_n(s)$ in weighted supremum metrics for the optimal class of weight functions under the null hypothesis that the X_i are i.i.d.r.v.'s with distribution F . Here we prove that on assuming the alternative hypothesis H_1 as parametrized by (8.0.4)–(8.0.6), the weak convergence of $\alpha_n(s)/q(s)$ continues to hold for the same optimal class of weight functions as in the case of H_0 .

As in Section 8.1., let $\mathcal{Q}^*$ be the class of positive functions on $(0, 1)$, i.e. $\inf_{\delta \leq t \leq 1-\delta} q(t) > 0$ for all $\delta \in (0, 1/2)$, which are nondecreasing near zero and non-increasing near one. We recall the integrals

$$E^*(q, c) = \int_0^1 (t(1 - t))^{-3/2} q(t) \exp(-c(t(1 - t))^{-1} q^2(t)) dt$$

and

$$I^*(q, c) = \int_0^1 (t(1-t))^{-1} \exp(-cq^2(t)/(t(1-t))) dt, \quad c > 0.$$

Theorem 8.9.1. *Let $q \in Q^*$ and $I^*(q, c) < \infty$ for all $c > 0$. Then under the sequence of alternatives of H_1 we have*

$$\alpha_n(s)/q(s) \xrightarrow{D} (B(s) + c(s))/q(s) \quad \text{in } D[0, 1],$$

where

$$c(s) = \int_0^1 \int_0^s g(\tau, u) du d\tau.$$

and $\{B(s); 0 \leq s \leq 1\}$ is a Brownian bridge.

Remark 8.9.1. According to Theorem 5.1.C, in Theorem 8.9.1 we can equivalently assume that $E^*(q, c) < \infty$ for all $c > 0$ and $\lim_{t \downarrow 0} q(t)/t^{1/2} = \lim_{t \uparrow 1} q(t)/(1-t)^{1/2} = \infty$.

Proof. Since $\alpha_n(s) = \alpha_n(s, 1)$, where $\alpha_n(s, t)$ is defined by (8.0.1) and $K(s, 1) \stackrel{D}{=} B(s)$, where $K(s, t)$ is a Kiefer process defined by (8.2.1) and $B(s)$ is a Brownian bridge, it follows from Lemma 8.7.1 (a) that

$$(\alpha_n(s), L_n(t)) \xrightarrow{D} (B(s), L(t)) \quad \text{in } D^2[0, 1],$$

where B is a Brownian bridge, L is a Gaussian process as in (8.1.5) and the covariance function of B and L is

$$Q_{B,L}(s, t) = EB(s)L(t) = c(s, t) = \int_0^t \int_0^s g(\tau, u) du d\tau.$$

Using Theorem 8.1.B and the fact that

$$E \frac{\alpha_n(s)}{q(s)} L_n(t) = \frac{1}{q(s)} E \alpha_n(s) L_n(t) \rightarrow \frac{1}{q(s)} c(s, t),$$

we obtain

$$(\alpha_n(s)/q(s), L_n(t)) \xrightarrow{D} (B(s)/q(s), L(t)),$$

and note also that the covariance function of $B(s)/q(s)$ and $L_n(t)$ is finite. Namely, we have

$$\begin{aligned} & \int_0^t \int_0^1 g(\tau, u) (\mathbf{1}\{u \leq s\} - s) du d\tau \\ &= \int_0^t \int_0^1 g(\tau, u) du d\tau - s \int_0^t \int_0^s g(\tau, u) du d\tau - s \int_0^t \int_s^1 g(\tau, u) du d\tau \\ &= (1-s) \int_0^t \int_0^1 g(\tau, u) du d\tau - s \int_0^t \int_s^1 g(\tau, u) du d\tau. \end{aligned}$$

Since

$$\begin{aligned} \left| \int_0^t \int_0^s g(\tau, u) du d\tau \right| &\leq \int_0^t \left(\int_0^s g^2(\tau, u) du \right)^{1/2} \left(\int_0^s du \right)^{1/2} d\tau \\ &\leq s^{1/2} \left(\int_0^t \int_0^s g^2(\tau, u) du d\tau \right)^{1/2} \end{aligned}$$

and

$$\left| \int_0^t \int_s^1 g(\tau, u) du d\tau \right| \leq (1-s)^{1/2} \left(\int_0^t \int_s^1 g^2(\tau, u) du d\tau \right)^{1/2},$$

where $g \in L^2[0, 1]^2$, we have

$$\begin{aligned} \sup_{0 < s < 1} |c(s, t)|/q(s) &\leq \sup_{0 < s < 1} \left((1-s)s^{1/2} \left(\int_0^t \int_0^s g^2(\tau, u) du d\tau \right)^{1/2} \right. \\ &\quad \left. - s(1-s)^{1/2} \left(\int_0^t \int_s^1 g^2(\tau, u) du d\tau \right)^{1/2} \right) / q(s) \\ &\leq \sup_{0 < s < 1} \left((1-s)^{1/2} \left(\int_0^t \int_0^s g^2(\tau, u) du d\tau \right)^{1/2} - s^{1/2} \left(\int_0^t \int_s^1 g^2(\tau, u) du d\tau \right)^{1/2} \right) \\ &\quad \times \sup_{0 < s < 1} (s(1-s))^{1/2} / q(t) < \infty, \end{aligned}$$

due to the fact that $q(s)$ is positive, i.e. $\inf_{\delta \leq s \leq 1-\delta} q(s) > 0$ for all $\delta \in (0, 1)$ and $I^*(q, c) < \infty$ for some $c > 0$ implies $\lim_{s \downarrow 0} q(s)/s^{1/2} = 0 = \lim_{s \uparrow 1} q(s)/(1-s)^{1/2}$.

APPENDIX

A.1. Some Measure Theoretical Results

In the proofs of approximations of stochastic processes we usually consider a representation of the original process and we get approximations for this version. The following two measure theoretical results make it possible to get approximations for the original process. The first one is Lemma 4.4.4 of Csörgő and Révész (1981) as stated in Csörgő (1983, p.21).

Theorem A.1.1 *Let μ be a probability measure defined on the Borel sets of the Skorohod space $D[0, 1] \times D[0, 1]$ and let X (resp. Y) be a $D[0, 1]$ -valued r.v. defined on a probability space $(\Omega_1, \mathcal{A}_1, P_1)$ (resp. on $(\Omega_2, \mathcal{A}_2, P_2)$) with*

$$P_1\{X \in A\} = \mu(A \times D[0, 1]), \quad P_2\{Y \in A\} = \mu(D[0, 1] \times A)$$

for any Borel sets A of $D[0, 1]$. Then there exists a probability measure P defined on $(\Omega_1 \times \Omega_2, \mathcal{A}_1 \times \mathcal{A}_2)$ such that for any Borel set B of $D[0, 1] \times D[0, 1]$ we have

$$P\{(\omega_1, \omega_2) \in (\Omega_1 \times \Omega_2) : (X(\omega_1), Y(\omega_2)) \in B\} = \mu(B).$$

The second result was obtained by Berkes and Philipp (1979). (cf. Csörgő, 1983, p.21).

Theorem A.1.2 *Let $(\Omega_i, \mathcal{A}_i)$ ($i = 1, 2, 3$) be measurable spaces with Ω_i ($i = 1, 2, 3$) being Polish spaces (i.e., complete separable metric spaces). Let X_i be random elements on Ω_i ($i = 1, 2, 3$). Let F be the distribution of (X_1, X_2) on $\Omega_1 \times \Omega_2$ and let G be the distribution of (X_2, X_3) on $\Omega_2 \times \Omega_3$ such that the second marginal of F equals the first marginal of G . Then there exists a probability measure H on $(\Omega_1 \times \Omega_2 \times \Omega_3, \mathcal{A}_1 \times \mathcal{A}_2 \times \mathcal{A}_3)$ such that the H -distribution of (X_1, X_2) is F and that of (X_2, X_3) is G .*

A.2. Skorohod-Dudley-Wichura theorem

Here we quote the well known Skorohod-Dudley-Wichura theorem following Shorack and Wellner (1986), p.47, together with their Remark 2, for easy reference.

Let (M, d) denote an arbitrary metric space. Let Y_n , $n \geq 0$, denote random elements from some probability space $(\Omega, \mathcal{F}, P)$ to the measurable space $(M, \mathcal{M}_d^B)$ ($\mathcal{M}_d^B$ denotes the σ -field generated by the collection of all open balls, where a ball is a subset of M of the form $\{y : d(y, x) < r\}$, for some $x \in M$ and some $r > 0$), and let P_n denote the probability measure induced on $(M, \mathcal{M}_d^B)$ by Y_n .

Theorem A.2.1. *Suppose*

$$Y_n \xrightarrow{D} Y_0 \quad \text{on } (M, \mathcal{M}_d^B, d) \quad \text{as } n \rightarrow \infty$$

and

$$P_0(M_s) = 1$$

for some $\mathcal{M}_d^B$ -measurable subset M_s of M that is d -separable (i.e., M_s has a countable subset that is d -dense in M_s). Then there exists a probability space $(\Omega^*, \mathcal{F}^*, P^*)$ and $\mathcal{M}_d^B - \mathcal{F}^*$ -measurable transformations Y_n^* , $n \geq 0$, from Ω^* to M inducing the probability measures P_n on $(M, \mathcal{M}_d^B)$ (i.e., Y_n and Y_n^* are equivalent processes) such that $d(Y_n^*, Y_0^*)$ is a.s. equal to a random variable and satisfies

$$d(Y_n^*, Y_0^*) \rightarrow 0 \quad \text{a.s.} \quad \text{as } n \rightarrow \infty.$$

Let $\| \cdot \|$ denote the supremum metric. Thus on $(D, \| \cdot \|)$ we have the following result. Let each Y_n , $n \geq 0$, be a process on $(D, \mathcal{D})$ with $P(X_0 \in C) = 1$. Suppose $Y_n \xrightarrow{D} Y$ on $(D, \mathcal{D}, \| \cdot \|)$ as $n \rightarrow \infty$. Then each Y_n may be replaced by an equivalent process Y_n^* (i.e., $Y_n^* \stackrel{D}{=} Y_n$) for which $\|Y_n^* - Y_n\| \rightarrow 0$ a.s. as $n \rightarrow \infty$. If $\|Y_n - Y_0\| \xrightarrow{P} 0$ as $n \rightarrow \infty$ is given, then $Y_n \xrightarrow{D} Y_0$ on $(D, \mathcal{D}, \| \cdot \|)$ as $n \rightarrow \infty$

can be concluded; in this latter case we can write either $Y_n \xrightarrow{D} Y_0$ on $(D, \mathcal{D}, \|\cdot\|)$ or $\|Y_n^* - Y_0^*\| \xrightarrow{P} 0$ as $n \rightarrow \infty$.

A.3. The Zero-One Law for Brownian Motion

If the sets of a σ -algebra of measurable sets of a measure space are null sets, or the complements of null sets, the σ -algebra is called *trivial*.

The following theorem and proof we quote from Doob (1984), p. 583.

Theorem A.3.1. *If $W(\cdot)$ is a Brownian motion starting from a point on the real line, the σ -algebras $\mathcal{F}_1$ and $\mathcal{F}_2$ generated, respectively, by the null sets and*

$$\bigcap_{t>0} \mathcal{F}\{W(s), s \leq t\}, \quad \bigcup_{t>0} \mathcal{F}\{W(s), s \geq t\}$$

are trivial.

Proof. It can be assumed (without loss of generality) that the initial point of the Brownian motion is the origin. If $A \in \mathcal{F}_1$ and if $t > 0$, then the set A is independent of the class of differences $\{W(s_2) - W(s_1), t \leq s_1 < s_2\}$ and therefore is independent of the σ -algebra $\mathcal{F}\{W(s_2) - W(s_1), 0 < s_1 < s_2\}$ which, since $W(0^+) = 0$ almost surely, together with the class of null sets generates a σ -algebra which includes $\mathcal{F}_1$. Thus A is independent of itself and so has probability either 0 or 1 ($P(A \cap A) = P(A)P(A) = P^2(A)$). Hence $\mathcal{F}_1$ is trivial. Triviality of $\mathcal{F}_2$ follows from that of $\mathcal{F}_1$ by applying this result to the Brownian motion $\{tW(1/t), 0 < t < \infty\}$ which for $t = 0$ is defined to be equal to zero.

Let q be a positive function on $(0, 1]$ which is nondecreasing in a neighbourhood of zero. Such a function is called a local function of a standard Wiener process $W(\cdot)$ if

$$\limsup_{t \downarrow 0} |W(t)|/q(t) < \infty \quad a.s.$$

Proposition A.3.1. *With q as above we have*

$$(A.3.1) \quad \limsup_{t \downarrow 0} |W(t)|/q(t) < \infty \quad \text{a.s.}$$

if and only if there exists a constant $0 \leq \beta < \infty$ such that

$$(A.3.2) \quad \limsup_{t \downarrow 0} |W(t)|/q(t) = \beta \quad \text{a.s.}$$

Proof. Obviously (A.3.2) implies (A.3.1). Assume that (A.3.1) is true. Then there exist constants $0 \leq \beta_1 \leq \beta_2 < \infty$ such that each of the events

$$A = \left\{ \beta_1 \leq \limsup_{t \downarrow 0} |W(t)|/q(t) \leq (\beta_1 + \beta_2)/2 \right\}$$

and

$$B = \left\{ (\beta_1 + \beta_2)/2 \leq \limsup_{t \downarrow 0} |W(t)|/q(t) \leq \beta_2 \right\}$$

has positive probability. By the 0 – 1 law (cf. Theorem A.3.1) we must have $P(A) = 0$ or 1 and $P(B) = 0$ or 1. Thus, since $P(A) > 0$ and $P(B) > 0$, we conclude $P(A) = P(B) = 1$. The latter is impossible, unless $\beta_1 = \beta_2$.

REFERENCES

- ADLER, R.J. and BROWN, L.D. (1986). Tail behaviour for suprema of empirical processes. *Ann. Probab.* **14** 1-30.
- ANDERSON, T.W. and DARLING, D.A. (1952). Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *Ann. Math. Statist.* **23** 193-212.
- BARNDORFF-NIELSEN, O. (1963). On the limit behaviour of extreme order statistics. *Ann. Math. Statist.* **34** 992-1002.
- BASSEVILLE, M. and BENVENISTE, A. eds. (1986). Detection of Abrupt Changes in Signals and Dynamical Systems. In *Lecture Notes in Control and Information Sciences* (eds. M. Thoma and A. Wyner) **77**. Springer-Verlag, Berlin.
- BERKES, I. and PHILIPP, W. (1979). Approximation theorems for independent and weakly dependent random vectors. *Ann. Prob.* **7**, 29-54.
- BHATTACHARYA, P.K. and FRIERSON, D. JR. (1981). A nonparametric control chart for detecting small disorders, *Ann. Statist.* **9** 544-554.
- BILLINGSLEY, P. (1968). *Convergence of Probability Measures*, J. Wiley, New York.
- BLUM, J.R., KIEFER, J. and ROSENBLATT, M. (1961). Distribution free tests of independence based on the sample distribution function. *Ann. Math. Statist.* **32** 485-498.
- BONVALOT, F. and CASTELLE, N. (1991). Strong approximation of uniform empirical process by Kiefer process. Prépublications 91-41, Université de Paris-Sud Mathématiques, Bât 425, Orsay, France.
- BRETAGNOLLE, J. and MASSART, P. (1989). Hungarian constructions from the nonasymptotic viewpoint. *Ann. Prob.* **17** 239-256.
- BRILLINGER, D.L. (1969). An asymptotic representation of the sample distribution function. *Bull. Amer. Math. Soc.* **75** 545-547.
- CANTELLI, F.P. (1933). Sulla determinazione empirica delle leggi di probabilità. *Giorn. Ist. Ital. Attuari* **4** 421-424.
- CHIBISOV, D. (1964). Some theorems on the limiting behaviour of empirical distribution functions, *Selected Transl. Math. Statist. Prob.* **6** 147-156.
- CHUNG, K.L. (1949). An estimate concerning the Kolmogoroff limit distribution. *Trans. Amer. Math. Soc.* **67** 36-50.
- CHUNG, K.L. and WILLIAMS, R.J. (1983). *Introduction to Stochastic Integration*. Birkhäuser, Boston.

- CSÖRGÖ, M. (1983). *Quantile processes with statistical applications*. SIAM 42. Philadelphia, Pennsylvania.
- CSÖRGÖ, M., CSÖRGÖ, S. and HORVÁTH, L. (1986). *An Asymptotic Theory for Empirical Reliability and Concentration Processes*. Lecture Notes in Statistics 33 (Springer-Verlag, Berlin).
- CSÖRGÖ, M., CSÖRGÖ, S., HORVÁTH, L. and MASON, D. (1986). Weighted empirical and quantile processes. *Ann. Probab.* **14** 31-85.
- CSÖRGÖ, M., DEHEUVELS, P. and HORVÁTH, L. (1987). An approximation of stopped sums with applications in queueing theory. *Adv. Appl. Prob.* **19** 674-690.
- CSÖRGÖ, M. and HORVÁTH, L. (1986). Approximations of weighted empirical and quantile processes. *Stat. Probab. Letters* **4** 275-280.
- CSÖRGÖ, M. and HORVÁTH, L. (1987a). Nonparametric tests for the changepoint problem. *J. Statist. Planning Inference* **17** 1-9.
- CSÖRGÖ, M. and HORVÁTH, L. (1987b). Detecting change in a random sequence. *J. Multivariate Anal.* **23** 119-130.
- CSÖRGÖ, M. and HORVÁTH, L. (1987c). Asymptotic distributions of pontograms. *Math. Proc. Camb. Phil. Soc.* **101** 131-139.
- CSÖRGÖ, M. and HORVÁTH, L. (1988a). Invariance principles for changepoint problems. *J. Multivariate Anal.* **27** 151-168.
- CSÖRGÖ, M. and HORVÁTH, L. (1988b). On the distributions of L_p norms of weighted uniform empirical and quantile processes. *Ann. Probab.* **16** 142-161.
- CSÖRGÖ, M. and HORVÁTH, L. (1988c). Nonparametric methods for changepoint problems. *Handbook of Statistics*, Vol. 7 403-425, Elsevier Science Publishers B.V. (North-Holland).
- CSÖRGÖ, M. and HORVÁTH, L. (1988d). A note on strong approximations of multivariate empirical processes. *Stoch. Proc. Appl.* **27** 101-109.
- CSÖRGÖ, M. and HORVÁTH, L. (1992). Weighted Approximations in Probability and Statistics (book forthcoming).
- CSÖRGÖ, M., HORVÁTH, L. and SHAO, Q.M. (1991). Convergence of integrals of uniform empirical and quantile processes. In *Technical Report Series of the Laboratory for Research in Statistics and Probability*, Carleton University, Ottawa, **168**. To appear in *Stoch. Proc. Appl.*

- CSÖRGÖ, M., HORVÁTH, L., SHAO, Q.M. and SZYSZKOWICZ, B. (1992). On the weighted asymptotics of partial sums and empirical processes of independent random variables. *A volume dedicated to Doeblin's memory* (to appear).
- CSÖRGÖ, M. and RÉVÉSZ, P. (1975). A new method to prove Strassen type laws of invariance principle. II. *Z. Wahrsch. Verw. Gebiete* **31** 261-269.
- CSÖRGÖ, M. and RÉVÉSZ, P. (1978). Strong approximations of the quantile process. *Ann. Statist.* **6** 882-894.
- CSÖRGÖ, M. and RÉVÉSZ, P. (1979). How big are the increments of a Wiener Process? *Ann. Probab.* **7** 731-737.
- CSÖRGÖ, M. and RÉVÉSZ, P. (1981). *Strong Approximations in Probability and Statistics*. Academic Press, New York.
- CSÖRGÖ, M., SHAO, Q.M. and SZYSZKOWICZ, B. (1991). A note on local and global functions of a Wiener process and some Rényi-type statistics. *Studia Sci. Math. Hungar.* **26** in press.
- DARLING, D.A. and ERDŐS, P. (1956). A limit theorem for the maximum of normalized sums of independent random variables. *Duke Math. J.* **23** 143-145.
- DESHAYES, J. and PICARD, D. (1983). Ruptures de modèles en statistique. Thèses d'état, Université Paris Sud.
- DESHAYES, J. and PICARD, D. (1986). Off-line statistical analysis of change-point models using non parametric and likelihood methods. In *Lecture Notes in Control and Information Sciences* (eds. M. Thoma and A. Wyner) **77: Detection of Abrupt Changes in Signals and Dynamical Systems** (eds. M. Basseville and A. Benveniste 103-168. Springer-Verlag, Berlin.
- DEVROYE, L. and GYÖRFI, L. (1985). *Nonparametric Density Estimation. The L_1 View*. Wiley, New York.
- DONSKER, M. (1951). An invariance principle for certain probability limit theorems. *Four Papers on Probability. Mem. Amer. Math. Soc. No. 6*.
- DONSKER, M. (1952). Justification and extension of Doob's heuristic approach to the Kolmogorov-Smirnov theorems. *Ann. Math. Statist.* **23** 277-283.
- DOOB, J.L. (1949). Heuristic approach to the Kolmogorov-Smirnov theorems. *Ann. Math. Statist.* **20**, 393-403.
- DOOB, J.L. (1984). *Classical Potential Theory and its Probabilistic Counterpart*. Springer-Verlag, New York.
- EASTWOOD, V.R. (1990). Some recent developments concerning asymptotic distributions of pontograms. *Math. Proc. Camb. Phil. Soc.* **108** 559-567.

- EINMAHL, U. (1989). The Darling-Erdős theorem for sums of i.i.d. random variables. *Probab. Theory Related Fields* **82**, no. 2, 241-257.
- ERDÖS, P. and KAC, M. (1946). On certain limit theorems of the theory of probability. *Bull. Amer. Math. Soc.* **52** 292-302.
- ETHIER, S.N. and KURTZ, T.G. (1986). *Markov Processes: Characterization and Convergence*. Wiley, New York.
- FELLER, W. (1944). On the Kolmogorov-Smirnov limit theorems for empirical distributions. *Ann. Math. Statist.* **19** 177-189.
- FELLER, W. (1966). *An Introduction to Probability Theory and its Applications*. Vol. 2. Wiley, New York.
- GLIVENKO, V. (1933). Sulla determinazione empirica delle leggi di probabilità. *Giorn. Ist. Ital. Attuari* **4** 92-99.
- GOMBAY, E. (1991). Sequential changepoint detection. Manuscript.
- GREENWOOD, P.E. and SHIRYAYEV, A.N. (1985), *Contiguity and the Statistical Invariance Principle*. Gordon and Breach Science Publishers.
- HACCOU, P., MELLIS, E., and VAN DE GEER, S. (1985). On the likelihood ratio test for a change point in a sequence of independent exponentially distributed random variables. Centre for Mathematics and Computer Science, Amsterdam, Netherland, Report MS-R8507.
- HÁJEK, J. and ŠIDÁK, Z. (1967). *Theory of Rank Test*. Academic Press, New York.
- HALL, P. (1979). On the invariance principle for U -statistics. *Stochastic Proc. Appl.* **9** 163-174.
- HALL, W.J. and LOYNES, R.M. (1977). On the concept of contiguity. *Ann. Probab.* **5** 278-282.
- HINKLEY, D.V. (1970). Inference about the change-point in a sequence of random variables. *Biometrika*, **57** 1-17.
- HOEFFDING, W. (1948a). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.* **19** 293-325.
- HOEFFDING, W. (1948b). A nonparametric test of independence. *Ann. Math. Statist.* **19** 546-557.
- HOEFFDING, W. (1961). *The Strong Law of Large Numbers for U -Statistics*. Univ. of North Carolina Institute of Statistics Mimeo Series 302.
- HUSE, V.R. (1988). Asymptotic properties for the sequential CUSUM procedure. *Stat. Probab. Letters* **7**.

- FIÓ, K. and McKEAN, H.P.JR. (1965). *Diffusion Processes and their Sample Paths*. Springer-Verlag, Berlin.
- JANSON, S. and WICHURA, M.J. (1983). Invariance principles for stochastic area and related stochastic integrals. *Stoch. Proc. Appl.* **16** 71-84.
- KENDALL, D.G. and KENDALL, W.S. (1980). Alignments in two-dimensional random sets of points. *Adv. Appl. Prob.* **12** 380-424.
- KHMALADZE, E.V. and PARJANADZE, A.M. (1986). Functional limit theorems for linear statistics of sequential ranks. *Probab. Theory Related Fields* **73** 322-334.
- KIEFER, J. (1961). On the large deviations of the empiric distribution function of vector chance variables and a law of the iterated logarithm. *Pacific J. Math.* **11** 649-660.
- KIEFER, J. (1972). Skorohod embedding of multivariate RV's and the sample D.F. *Z. Wahrsh. Verw. Gebiete* **24** 1-35.
- KOMLÓS, J., MAJOR, P. and TUSNÁDY, G. (1975). An approximation of partial sums of independent R.V.'s and the sample D.F.I. *Z. Wahrsch. Verw. Gebiete* **32** 111-131.
- KOMLÓS, J., MAJOR, P. and TUSNÁDY, G. (1976). An approximation of partial sums of independent R.V.'s and the sample D.F.II. *Z. Wahrsch. Verw. Gebiete* **34** 33-58.
- KOLMOGOROV, A.N. (1933). Sulla determinazione empirica di una legge di distribuzione. *Giorn. Ist. Ital. Attuari* **4** 83-91.
- LE CAM, L. (1960). Locally asymptotically normal families of distributions. University of California Publications in Statistics, **3** 37-98.
- LE CAM, L. (1986). *Asymptotic Methods in Statistical Decision Theory*. Springer-Verlag, New York.
- LE CAM, L. and YANG, G.L. (1990). *Asymptotics in Statistics; Some Basic Concepts*. Springer-Verlag, New York.
- LOMBARD, F. (1981). An invariance principle for sequential nonparametric test statistics under contiguous alternatives. *South African Statist. J.* **15** 107-111.
- LOMBARD, F. (1983). Asymptotic distributions of rank statistics in the change-point problem. *South African Statist. J.* **17** 83-105.
- LOMBARD, F. (1986). The changepoint problem for angular data: A nonparametric approach. *Technometrics* **28** 391-397.
- LOMBARD, F. (1987). Rank tests for changepoint problems. *Biometrika* **74** 613-624.

- LOMBARD, F. and MASON, D.M. (1985). Limit theorems for generalized sequential rank statistics. *Z. Wahrsch. verw. Gebiete* **70** 395-410.
- MAJOR, P. (1976a). The approximation of partial sums of independent r.v.'s. *Z. Wahrsch. verw. Gebiete* **35** 213-220.
- MAJOR, P. (1976b). Approximation of partial sums of i.i.d.r.v's when the summands have only two moments. *Z. Wahrschein. verw. Gebiete* **35** 221-230.
- MAJOR, P. (1979). An improvement of Strassen's invariance principle. *Ann. Probab.* **7** 55-61.
- MASON, D.M. (1981). On the use of a statistic based on sequential ranks to prove limit theorems for simple linear rank statistics. *Ann. Statist.* **9** 424-436.
- MATTHEWS, D.A., FAREWELL, V.T., and PYKE, R. (1985). Asymptotic score statistics processes and tests for constant hazard against a change-point alternative. *Ann. Statist.* **13** 583-591.
- OOSTERHOFF, J. and VAN ZWET, W.R. (1975). A note on contiguity and Hellinger distance. In: J. Jurechkova, ed., *Contributions to Statistics*. Reidel, Dordrecht 157-166.
- O'REILLY, N. (1974). On the weak convergence of empirical processes in sup-norm metrics. *Ann. Probab.* **2** 642-651.
- PAGE, E.S. (1954). Continuous inspection schemes. *Biometrika* **41** 100-115.
- PAGE, E.G. (1955). A test for a change in a parameter occurring at an unknown point. *Biometrika* **42** 523-526.
- PARDZHANADZE, A.M. and KHMALADZE, E.V. (1986). On the asymptotic theory of statistics of sequential ranks. *Theory Probab. Appl.* **31** 669-682.
- PARENT, E.A., JR. (1965). Sequential ranking procedures. Doctoral Dissertation, Stanford Univ.
- PARZEN, E. (1979). Nonparametric statistical data modeling. *J. Amer. Statist. Assoc.* **74** 105-131.
- PARZEN, E. (1992). Comparison change analysis. In *Nonparametric Statistics and Related Topics* (A.K.Md.E. Saleh, ed.) 3-15, North Holland, Amsterdam.
- PETTITT, A.N. (1979). A non-parametric approach to the changepoint problem. *Appl. Statist.* **28**, 126-135.
- PETTITT, A.N. (1980). A simple cumulative sum type statistic for the change-point problem with zero-one observations. *Biometrika* **67** 79-84.
- PETTITT, A.N. (1980a). Some results on estimating a change-point using nonparametric type statistics. *J. Statist. Comput. Simul.*, **11** 261-272.

- PICARD, D. (1985). Testing and estimating change-points in time series. *Adv. Appl. Prob.*, **17** 126-135.
- PROKHOROV, YU. V. (1956). Convergence of random processes and limit theorems in probability theory. *Theory Probab. Appl.* **1** 157-214.
- PYKE, R. and SHORACK, G.R. (1968). Weak convergence of a two-sample empirical process and a new approach to Chernoff-Savage theorem. *Ann. Math. Statist.* **39** 755-771.
- RÉNYI, A. (1953). On the theory of order statistics. *Acta. Math. Acad. Sci. Hung.* **4** 191-232.
- REYNOLDS, M.R., JR. (1975). A sequential signed-rank test for symmetry. *Ann. Statist.* **3** 382-400.
- ROUSSAS, G.G. (1972). *Contiguity of Probability Measures: Some Applications in Statistics*. Cambridge University Press.
- SCHECHTMAN, E. and WOLFE, D.E. (1981). Distribution-free tests for the change-point problem. Technical Report, The Ohio State University.
- SCHEFFÉ, H. (1947). A useful convergence theorem for probability distributions. *Ann. Math. Statist.* **18** 434-438.
- SEN, A. and SRIVASTAVA, M.S. (1975). On tests for detecting changes in mean. *Ann. Statist.* **3** 98-108.
- SERFLING, R.J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley, New York.
- SHABAN, S.A. (1980). Change-point problem and two-phase regression: an annotated bibliography. *Int. Statist. Rev.* Vol. 48, No. 1, 83-93.
- SHEPP, L.A. (1966). Radon-Nikodym derivatives of Gaussian measures. *Ann. Math. Statist.* **37** 321-354.
- SHORACK, G.R. and WELLNER, J.A. (1986). *Empirical Processes with Applications to Statistics*. Wiley, New York.
- SIEGMUND, D. (1985). *Sequential Analysis: Tests and Confidence Intervals*. Springer, New York.
- SIEGMUND, D. (1986). Boundary crossing probabilities and statistical applications. *Ann. Statist.* **14** 361-404.
- SKOROHOD, A.V. (1956). Limit theorems for stochastic processes. *Theory Probab. Appl.* **1** 261-290.
- SKOROHOD, A.V. (1961). *Studies in the Theory of Random Processes*, Kiev Univ., Addison-Wesley, Reading, Mass., 1965 (translation).

- SMIRNOV, N.V. (1939). On the estimation of the discrepancy between empirical curves of distribution for two independent samples. *Bull. Math. de l'Université de Moscow* 2 (Fasc. 2).
- STRASSEN, V. (1964). An invariance principle for the law of iterated logarithm. *Z. Wahrsch. verw. Gebiete* 3 211-226.
- SZYSZKOWICZ, B. (1991a). Changepoint problem and contiguous alternatives. *Statist. and Probab. Letters* 11 299-308.
- SZYSZKOWICZ, B. (1991b). Weak convergence of empirical type processes under contiguous and changepoint alternatives. In *Tech. Rep. Ser. Lab. Res. Stat. Prob.* No. 164 Carleton U.-U. Ottawa.
- SZYSZKOWICZ, B. (1991c). Empirical type processes and contiguity. *C.R. Math. Rep. Acad. Sci. Canada* XIII, No. 4 161-166.
- SZYSZKOWICZ, B. (1991d). Weighted stochastic processes under contiguous alternatives. *C.R. Math. Rep. Acad. Sci. Canada* XIII, No. 5 211-216.
- SZYSZKOWICZ, B. (1992a). On $\|\cdot/q\|$ -metric convergence and contiguous alternatives. In *Tech. Rep. Ser. Lab. Res. Stat. Prob.* No. 191 Carleton U.-U. Ottawa.
- SZYSZKOWICZ, B. (1992b). L_p -approximation of weighted partial sum processes. In *Tech. Rep. Ser. Lab. Res. Stat. Prob.* No. 191 Carleton U. - U. Ottawa. To appear in *Stoch. Proc. Appl.*
- SZYSZKOWICZ, B. (1992c). L_p -functionals of weighted partial sum processes. *C.R. Math. Rep. Acad. Sci. Canada* XIV, No. 1 31-36.
- SZYSZKOWICZ, B. (1992d). Asymptotic distributions of weighted pontograms under contiguous alternatives. In *Tech. Rep. Ser. Lab. Res. Stat. Prob.* No. 194 Carleton U.-U. Ottawa. To appear in *Math. Proc. Camb. Phil. Soc.* (1992).
- SZYSZKOWICZ, B. (1992e). Asymptotic distributions of weighted compound Poisson bridges. In *Tech. Rep. Ser. Lab. Res. Stat. Prob.* No. 194 Carleton U.-U. Ottawa. To appear in *Math. Proc. Camb. Phil. Soc.*
- WOLFE, D.E. and SCHECHTMAN, E. (1984). Nonparametric statistical procedures for the changepoint problem. *J. Statist. Plann. Inference* 9 389-396.
- WITTING, H. and NÖLLE, G. (1970). *Angewandte mathematische Statistik*. Teubner, Stuttgart.

- YAO, Y.C. and DAVID, R.A. (1984). The asymptotic behaviour of the likelihood ratio statistic for testing a shift in mean in a sequence of independent normal variates. Center for Stochastic Processes, Dept. of Statistics, University of North Carolina, Chapel Hill, Technical Report No. 80.
- ZACKS, S. (1983). Survey of classical and Bayesian approaches to the change-point problem: fixed sample and sequential procedures of testing and estimation. *Recent Advances in Statistics*. New York, Academic Press, 245-269.

END

2/6 05 193

FIN