1 Introduction

Let us consider a sequence of identically distributed random variables $Y_1, Y_2, \ldots$ with distribution function $F$. The empirical process is defined as

$$
\sum_{j=1}^{N} \left(1\{Y_j \leq x\} - F(x)\right)
$$

indexed by the domain of $F$. Normalizing (1.1) with $N^{-1}$ yields the difference of the empirical distribution function $F_N(x)$ and $F(x)$, which by Glivenko-Cantelli theorem converges almost surely to 0, uniformly in $x$. The investigation of weak convergence to a non-degenerated limit was initiated by Donsker (1952). He studied the empirical process of independent uniformly distributed random variables and proved weak convergence to a Brownian Bridge $(B^0_t)_{0 \leq t \leq 1}$ that is a centered Gaussian process with covariance $E B^0_s B^0_t = s \wedge t - st$. Although it turned out that some parts of his proof were incorrect, see Dudley (1999, Theorem 1.1.1) for a revised proof, this result marked the beginning of a modern empirical process theory. Subsequently, a number of authors studied the empirical process under different types of dependence, see e.g. Billingsley (1968, Theorem 22.2) for functions of $\phi$-mixing processes, Deo (1973) for strong-mixing, and Prieur (2002) for $s$-weakly dependent sequences. Besides this progress an interpretation of (1.1) as a two-parameter empirical process became popular in the 1970s. Replacing $N$ by the integer part of $Nt$, $0 \leq t \leq 1$, yields the so called sequential empirical process $(R_N(x,t))$, given by

$$
R_N(x,t) = \sum_{j=1}^{\lfloor Nt \rfloor} \left(1\{Y_j \leq x\} - F(x)\right).
$$

For a fixed $x$ this is the partial sum process for the functional $1\{\cdot \leq x\} - F(x)$; for a fixed $t$ we get the empirical process of the first $\lfloor Nt \rfloor$ observations. Especially for $t = 1$ we have $N^{-1}R_N(x,1) = F_N(x) - F(x)$. Müller (1970) and Kiefer (1972) studied (1.2) initially for independent observations. They proved among others weak convergence to a Gaussian process $(K(x,t))$ called the Kiefer-Müller process. This process can be understand as a Brownian bridge in the first and a Brownian motion in the second parameter. Under weak dependence one still gets a Gaussian limit but with a slightly different covariance structure. The case of strong mixing observations was handled by Berkes and Philipp (1977). Berkes and Horváth (2001) studied the sequential empirical process for GARCH processes. $S$-mixing sequences were treated by Berkes et al. (2009). In this thesis we study $(R_N(x,t))$ under long-range dependence. The paper by Dehling and Taqqu (1989) provides the basis for our work. Their result will be discussed below.
1 Introduction

Figure 1.1: Yearly minimal water level of the Nile river for the years 622–1284 (663 observations) measured at the Roda gauge near Cairo. The dataset is from R package longmemo.

The investigation of long-range dependence was motivated by hydrological data, see Figure 1.1 for an illustration. In the mid of the last century the hydrologist Harold Hurst (1951) studied the yearly discharge of the main Nile River at Aswan and of Lake Albert on the border between Uganda and the Democratic Republic of the Congo. His aim was to determine the optimal size of a reservoir so that a default outflow is guaranteed. Therefor he used the so-called R/S statistic: let $S_i$ be the cumulated sum of the yearly discharge, $S_0 := 0$, and let $\hat{\sigma}$ be the sample standard deviation. If the yearly outflow is assumed to be equal to the sample mean, $S_i - \frac{i}{n}S_n$ describes the difference of the cumulative in- and output. To avoid a flooding and to ensure a minimum storage it is reasonable to consider the normalized gap between maximum and minimum of these differences as a test statistic, that is,

$$\max_{0 \leq i \leq n}(S_i - \frac{i}{n}S_n) - \min_{0 \leq i \leq n}(S_i - \frac{i}{n}S_n) \over \hat{\sigma}.$$  

(1.3)

To get an idea of the asymptotic behavior of the R/S statistic let us consider the functional $\phi(f) = \sup_{t \in [0,1]}(f(t) - tf(1)) - \inf_{t \in [0,1]}(f(t) - tf(1))$ defined on the space of càdlàg functions on $[0,1]$. If the yearly discharges are realizations of some identically distributed random variables $Y_1, Y_2, \ldots$, with expectation $\mu$ and finite variance, the weak
For independent observations we have weak convergence of \( n^{-1/2} \sum_{i=1}^{\lfloor nt \rfloor} (Y_i - \mu) \) to a Brownian motion \((B_t)_{t \in [0,1]}\) so that the continuous mapping theorem yields 
\[
n^{-1/2} R/S \xrightarrow{d} \sup_{t \in [0,1]} (B_t - tB_1) - \inf_{t \in [0,1]} (B_t - tB_1).
\]
Consequently, if the yearly discharges are assumed to be independent, the R/S statistic would grow like \( \sqrt{n} \) but Hurst observed that the growth rate behaves more like \( n^H \) with \( H > 1/2 \). He conjectured that this behavior is typical for long-term records of natural events and to confirm this, he applied the R/S statistic additionally on 75 different data sets of temperature, rainfall, annual growth rings of trees and similar phenomena, see Hurst (1951, pp. 784). It turned out that the average value of \( H \) is 0.72 instead of 0.5 what one would expect under independence. Moreover, the same behavior appeared for data sets which Hurst simulated by using coins and a deck of cards, respectively. Besides he found out that “to toss 10 coins 100 times required about 35 min, whereas shuffling and cutting for 100 cards required 20 min”, see Hurst (1951, p.781). Over the years many authors found further examples for long-range dependent behavior in finance, physics, biology, and even in linguistics, see Rooch (2012, pp. 6) for a detailed list of references.

Dehling and Taqqu (1989, Theorem 1.1) studied the sequential empirical process for subordinated data \((G(X_j))\), where the underlying Gaussian process exhibits long-range dependence. They could show that under some technical assumptions 
\[
(dN^-1 R_N(x,t)) \xrightarrow{d} (J(x)Z(t))
\] in the càdlàg space \( D([-\infty, \infty] \times [0,1]) \) equipped with the sup-norm and the sigma algebra generated by the open balls. In contrast to the case of independent and weak
Introduction

dependent observations the normalizing factor \( d_N \) grows faster than \( \sqrt{N} \). Furthermore, the limiting process is given by the product of a deterministic function \( J(x) \) and a stochastic process \( (Z(t)) \). The latter belongs to the class of Hermite processes. Each of these processes has various representations as a \( m \)-fold stochastic integral. For \( m = 1 \) the Hermite process is a fractional Brownian motion, the long-range dependent analogue of a Brownian motion, see Figure 1.2. If \( m > 1 \), Hermite processes are non Gaussian. All this phenomena explain why functional limit theorems under long-range dependence are sometimes called non-central limit theorems.

The aim of this thesis is to establish some new non-central limit theorems for the sequential empirical process under long-range dependence. For this purpose we study (1.2) with respect to stronger metrics, multivariate data, and more general index sets. Most of the definitions and techniques we use, such as multivariate long-range dependence, subordinated Gaussian processes, and Hermite polynomials, will be introduced in Chapter 2. More specific concepts are handled in the respective chapters. In Chapter 3 we show that (1.4) still holds in an appropriate subspace of \( D([-\infty, \infty] \times [0, 1]) \) equipped with a non-uniform sup-norm. We use this result to establish the asymptotic behavior of \( U \)-statistics with an unbounded kernel by applying a modified functional delta method. In Chapter 4 we study (1.2) for multivariate subordinated data. To clarify the idea of the proof we distinguish between one- and \( p \)-dimensional subordination. In the first case the limiting process is again given by the product of a deterministic function and an Hermite process. The latter case yields a sum of various dependent processes of this kind. Chapter 5 handles the sequential empirical process in a more general setup. We replace the index set \( \mathbb{R}^p \times [0, 1] \) by \( \mathcal{F} \times [0, 1] \), where \( \mathcal{F} \) is a class of square-integrable
functions. If the elements of $\mathcal{F}$ fulfill some uniformly bounded moment condition and if additionally an entropy condition holds then we have weak convergence to a process similar to that one established in Chapter 4.