HAL
open science

Convergence of self-normalized partial sums processes in C[0,1] and D[0,1]

Alfredas Račkauskas, Charles Suquet

To cite this version:

Alfredas Račkauskas, Charles Suquet. Convergence of self-normalized partial sums processes in C[0,1] and D[0,1]. Publications Irma Lille Nouvelle Serie, 2001, 53 (VI), pp.1-9. hal-00834538

HAL Id: hal-00834538

https://hal.science/hal-00834538v1

Submitted on 16 Jun 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

PUB. IRMA, LILLE 2013
Vol. 53, N° VI

Convergence of self-normalized partial sums processes in $C[0,1]$ and $D[0,1]$*

Alfredas Račkauskas
Vilnius university and Institute of Mathematics and Informatics
Department of Mathematics, Vilnius University
Naugarduko 24, Lt-2006 Vilnius
Lithuania

Charles Suquet
Université des Sciences et Technologies de Lille
Statistique et Probabilités F.R.E. CNRS 2222
Bât. M2, U.F.R. de Mathématiques
F-59655 Villeneuve d'Ascq Cedex France

Abstract

Let $(X_i)_{i\geq 1}$ be an i.i.d. sequence of mean zero random variables, $S_n := X_1 + \cdots + X_n$ and $V_n^2 := X_1^2 + \cdots + X_n^2$. We consider four sequences of partial sums processes: the broken lines with vertices at the points $(k/n, S_k/V_n)$ or $(V_k^2/V_n^2, S_k/V_n)$ and the corresponding random step functions. We prove that each of them converges weakly in $C[0,1]$ or $D[0,1]$ to the Brownian motion *if and only if* X_1 belongs to the *domain of attraction of the normal distribution.* These results contrast with the classical Donsker Prohorov invariance principles where the N.S.C. for such convergences is $\mathbf{E}\, X_1^2 < \infty$.

Résumé

Soit $(X_i)_{i\geq 1}$ une suite de variables aléatoires i.i.d. centrées, $S_n := X_1 + \cdots + X_n$ et $V_n^2 := X_1^2 + \cdots + X_n^2$. Nous considérons quatre suites de processus de sommes partielles : les lignes polygonales de sommets les points $(k/n, S_k/V_n)$ ou $(V_k^2/V_n^2, S_k/V_n)$ et les fonctions aléatoires étagées correspondantes. Nous prouvons que chacune de ces suites converge en loi dans $C[0,1]$ ou $D[0,1]$ vers le mouvement brownien *si et seulement si* X_1 est dans le *domaine d'attraction de la loi normale.* Ces résultats contrastent avec les principes d'invariance de Donsker Prohorov classiques pour lesquels la C.N.S. est $\mathbf{E}\, X_1^2 < \infty$.

*Preprint. Research supported by a cooperation agreement CNRS/LITHUANIA (4714).

Keywords: domain of attraction, functional central limit theorem, invariance principle, self-normalization.
AMS classifications: 60F05, 60B05, 60G17, 60E10.

1 Introduction

Many limit theorems in the theory of probability deal with the asymptotic behavior of sums $S_n = X_1 + \cdots + X_n$ of independent identically distributed mean zero random variables, normalized by some deterministic sequence (b_n). When the X_i's are square integrable, the classical choice of $b_n = \sqrt{n}$ is used to obtain central limit theorems as well as invariance principles in $C[0,1]$ or $D[0,1]$. Self-normalization consists here in replacing b_n by $V_n = (X_1^2 + \cdots + X_n^2)^{1/2}$. There is of course a strong statistical interest in this replacement. For instance the classical Student t-statistic can be easily expressed in term of self-normalized sums. It is now well established that in general self-normalization improves the asymptotic properties of sums of independent random variables.

A rich litterature is devoted to limit theorems for self-normalized sums. Logan, Mallows, Rice and Shepp [11] investigate the various possible limit distributions of self-normalized sums. Giné, Götze and Mason [8] prove that S_n/V_n converges to the Gaussian standard distribution if and only if X_1 is in "DAN", the domain of attraction of the normal distribution (the symetric case was previously treated in Griffin and Mason [10]). Egorov [7] investigates the non identically distributed case. Bentkus and Götze [3] obtain the rate of convergence of S_n/V_n when $X_1 \in DAN$. Griffin and Kuelbs [9] prove the LIL for self-normalized sums when $X_1 \in DAN$. Moderate deviations (of Linnik's type) are studied in Shao [15] and Christiakov and Götze [5]. Large deviations (of Cramér-Chernoff type) are investigated in Shao [14] without moment conditions.

Chuprunov [6] gives invariance principles for various partial sums processes under self-normalization in $C[0,1]$ or $D[0,1]$. Račkauskas and Suquet [13] obtained the weak convergence in the best possible Hölder space of the adaptive broken line of self-normalized partial sums of *symetric* i.i.d. X_i's when $X_1 \in DAN$. In the present contribution we are concerned with invariance principles in the more classical frameworks of $C[0,1]$ or $D[0,1]$ for processes of self-normalized partial sums of i.i.d. X_i's in DAN, without symetry assumption. We prove that these processes converge to the Brownian motion in $C[0,1]$ or $D[0,1]$ if and only if $X_1 \in DAN$. This improves on Chuprunov's result in the i.i.d. case.

Let us fix now the notations used throughout the paper. $X_1, \ldots, X_n, \ldots$ are i.i.d. mean zero random variables. Set $S_0 = 0$, $V_0 = 0$,

$$S_k = X_1 + \cdots + X_k, \quad V_k^2 = X_1^2 + \cdots + X_k^2 \quad \text{for} \quad k = 1, 2, \ldots$$

When $V_n = 0$, we adopt the convention that $S_k/V_n := 0$ $(k \leq n)$. We consider the following processes:

- the standard (non normalized) broken line process

$$\xi_n(t) = S_{[nt]} + (nt - [nt])X_{[nt]+1}, \quad t \in [0,1];$$

- the adaptive broken line process $\zeta_n = (\zeta_n(t), t \in [0,1])$ which is defined by

linear interpolation of the points

$$\Big(V_k^2/V_n^2,\ S_k/V_n\Big), \quad k=0,\dots,n,$$

whenever $V_n>0$ and by $\zeta_n=0$ when $V_n=0$;

- the standard (non normalized) step process

$$W_n(t)=S_{[nt]}, \quad t\in[0,1];$$

- the adaptive step process $Z_n=(Z_n(t),t\in[0,1])$ which whenever $V_n>0$ is constant on each intervall $[V_k^2/V_n^2, V_{k+1}^2/V_n^2)$ and satisfies

$$Z_n\Big(\frac{V_k^2}{V_n^2}\Big)=\frac{S_k}{V_n} \quad k=0,\dots,n,$$

while for $V_n=0$, $Z_n:=0$.

2 Donsker-Prohorov invariance principle under self-normalization

Before stating the results let us remind that $X_1\in DAN$ means that there exists a sequence $\ell_n\uparrow\infty$ such that

$$\ell_n^{-1}n^{-1/2}S_n \xrightarrow{\mathcal{D}} N(0,1). \tag{1}$$

This yields that

$$\ell_n^{-2}n^{-1}V_n^2 \xrightarrow{P} 1. \tag{2}$$

We have moreover for each $\tau>0$,

$$nP(|X_1|>\tau\ell_n\sqrt{n}) \quad\to\quad 0, \tag{3}$$

$$\ell_n^{-2}\mathbf{E}\,(X_1^2;|X_1|\le\tau\ell_n\sqrt{n}) \quad\to\quad 1, \tag{4}$$

$$n\mathbf{E}\,(X_1;|X_1|\le\tau\ell_n\sqrt{n}) \quad\to\quad 0, \tag{5}$$

see for instance Araujo and Giné [1], Chap. 2, Cor. 4.8 (a), Th. 6.17 (i) and Cor. 6.18 (b). Here and in all the paper $(X;E)$ means the product of the random variable X by the indicator function of the event E.

Let us also recall O'Brien's [12] result: $X_1\in DAN$ if and only if

$$V_n^{-1}\max_{1\le k\le n}|X_k| \xrightarrow{P} 0. \tag{6}$$

Throughout the paper, W denotes a standard Brownian motion.

Theorem 2.1 *The convergence*

$$\frac{\xi_n}{V_n} \xrightarrow{\mathcal{D}} W \tag{7}$$

holds in the space $C[0,1]$ if and only if $X_1\in DAN$.

Theorem 2.2 *The convergence*

$$V_n^{-1}\zeta_n \xrightarrow{\mathcal{D}} W \tag{8}$$

holds in the space $C[0,1]$ *if and only if* $X_1 \in DAN$.

Let us remark that the necessity of $X_1 \in DAN$ in both Theorem 2.1 and Theorem 2.2 follows from Giné, Götze and Mason [8]. Let us notice also that (1) or (8) both exclude the trivial case $P(X_1 = 0) = 1$, so that almost surely $V_n > 0$ for large enough n.

Proof of theorem 2.1. First we prove the convergence of finite dimensional distributions (f.d.d.) of the process $(V_n^{-1}\xi_n)$ to the corresponding f.d.d. of the Wiener process W.

To this aim, consider the process $W_n = (S_{[nt]}, t \in [0,1])$. By (6) applied to the obvious bound

$$\sup_{0\le t\le 1} V_n^{-1}|\xi_n(t) - W_n(t)| \le V_n^{-1} \max_{1\le k\le n} |X_k|,$$

the convergence of f.d.d. of $V_n^{-1}\xi_n$ follows from those of the process $V_n^{-1}W_n$.

Let $0 \le t_1 < t_2 < \cdots < t_d \le 1$ and let (ℓ_n) be the (non random) sequence introduced above. From (1), independence of the $X_i's$ and the fact that $\ell_{[nt]}^2\ell_n^{-2}$ converges to 1 (see e.g. Remark 3.2 in [13]), we get

$$\begin{aligned}&\ell_n^{-1}n^{-1/2}\big(S_{[nt_1]}, S_{[nt_2]} - S_{[nt_1]}, \ldots, S_{[nt_d]} - S_{[nt_{d-1}]}\big) \xrightarrow{\mathcal{D}}\\ &\big(W(t_1), W(t_2) - W(t_1), \ldots, W(t_d) - W(t_{d-1})\big).\end{aligned}$$

Now (2) and the continuity of the map

$$(x_1, x_2, \ldots, x_d) \mapsto (x_1, x_2 + x_1, \ldots, x_d + \cdots + x_1)$$

yields the convergence of f.d.d. of $V_n^{-1}W_n$. The convergence of finite dimensional distributions of the process $V_n^{-1}\xi_n$ is thus established.

To prove the tightness we shall use Theorem 8.3 from Billingsley [4]. Since $\xi_n(0) = 0$, the proof reduces in showing that for all ε, $\eta > 0$ there exist $n_0 \ge 1$ and δ, $0 < \delta < 1$, such that

$$\frac{1}{\delta}P\Big\{\sup_{1\le i\le n\delta} V_n^{-1}|S_{k+i} - S_k| \ge \varepsilon\Big\} \le \eta, \quad n \ge n_0, \tag{9}$$

for all $1 \le k \le n$.

Let us introduce the truncated variables

$$Y_i := \ell_n^{-1}(X_i;\ X_i^2 \le \tau^2\ell_n^2 n), \quad i = 1, \ldots, n,$$

where ℓ_n is defined above and τ will be chosen later. Denote by $\widetilde{S}_k$ and $\widetilde{V}_k$ the corresponding partial sums with their autonormalizing random variables:

$$\widetilde{S}_k = Y_1 + \cdots + Y_k, \quad \widetilde{V}_k = (Y_1^2 + \cdots Y_n^2)^{1/2}, \quad k = 1, \ldots, n.$$

Then we have

$$P\Big\{\sup_{1\le i\le n\delta} V_n^{-1}|S_{k+i}-S_k|\ge\varepsilon\Big\}\le A+B+C, \tag{10}$$

where

$$\begin{aligned} A &:= P\Big\{\sup_{1\le i\le n\delta}|\widetilde{S}_{k+i}-\widetilde{S}_k|\ge\varepsilon\sqrt{n}/2\Big\},\\ B &:= P\{\widetilde{V}_n<\sqrt{n}/2\},\\ C &:= nP\{|X_1|\ge\tau\ell_n\sqrt{n}\}. \end{aligned}$$

Due to (5) we can choose n_1 such that $\sqrt{n}|\mathbf{E}\,Y_1|\le 1/4$ for $n\ge n_1$. Then with $n\ge n_1$ and $\delta\le\varepsilon$ we have

$$\begin{aligned} A &\le P\Big\{\max_{1\le i\le n\delta}\Big|\sum_{j=k+1}^{k+i}(Y_j-\mathbf{E}\,Y_j)\Big|+n\delta|\mathbf{E}\,Y_1|\ge\sqrt{n}\varepsilon/2\Big\}\\ &\le P\Big\{\max_{1\le i\le n\delta}\Big|\sum_{j=k+1}^{k+i}(Y_j-\mathbf{E}\,Y_j)\Big|\ge\sqrt{n}\varepsilon/4\Big\}. \end{aligned}$$

By Chebyshev's inequality and Rosenthal inequality with $p>2$, we have for each $1\le k\le n\delta$

$$\begin{aligned} P\Big\{n^{-1/2}\Big|\sum_{j=k+1}^{k+n\delta}(Y_j-\mathbf{E}\,Y_j)\Big|\ge\frac{\varepsilon}{8}\Big\} &\le \frac{8^p}{\varepsilon^p n^{p/2}}\mathbf{E}\Big|\sum_{j=k+1}^{k+n\delta}(Y_j-\mathbf{E}\,Y_j)\Big|^p\\ &\le \frac{8^p}{\varepsilon^p n^{p/2}}\big[(n\delta)^{p/2}(\mathbf{E}\,Y_1^2)^{p/2}+n\delta\mathbf{E}\,|Y_1|^p\big]. \end{aligned}$$

By (4) we can choose n_2 such that

$$3/4\le\mathbf{E}\,Y_1^2\le 3/2\quad\text{for}\quad n\ge n_2. \tag{11}$$

Then we have $\mathbf{E}\,|Y_1|^p\le 2n^{(p-2)/2}\tau^{p-2}$ and then assuming that $\tau\le\delta^{1/2}$ we obtain

$$\begin{aligned} P\Big\{n^{-1/2}\Big|\sum_{j=k+1}^{k+n\delta}(Y_j-\mathbf{E}\,Y_j)\Big|\ge\frac{\varepsilon}{8}\Big\} &\le \frac{8^p}{\varepsilon^p n^{p/2}}\big[2^{p/2}(n\delta)^{p/2}+\delta n^{p/2}\tau^{p-2}\big]\\ &\le \frac{2\cdot 16^p\delta^{p/2}}{\varepsilon^p}. \end{aligned}$$

Now by Ottaviani inequality we find

$$A\le\frac{\delta\eta}{3}, \tag{12}$$

provided $\delta^{p/2}\le\varepsilon^p/(4\cdot 16^p)$ and $\delta^{(p-2)/2}\le\eta\varepsilon^p/(6\cdot 16^p)$.

Next we consider B. Since $n^{-1}\mathbf{E}\,\widetilde{V}_n^2 = \mathbf{E}\,Y_1^2$ we have by (11) $n^{-1}\mathbf{E}\,\widetilde{V}_n^2 \geq 3/4$, for $n \geq n_2$. Furthermore

$$B \leq P\{n^{-1}|\widetilde{V}_n^2 - \mathbf{E}\,\widetilde{V}_n^2| \geq 1/2\} \leq 4n^{-1}\mathbf{E}\,Y_1^4 \leq 4\tau^2\mathbf{E}\,Y_1^2 \leq \delta\eta/3, \tag{13}$$

provided $n \geq n_2$ and $\tau^2 \leq \delta\eta/18$.

Finally choose n_3 such that $C \leq \eta\delta/3$ when $n \geq n_3$ and join to that the estimates (12, 13) to conclude (9). The proof is completed.

Proof of theorem 2.2.

Due to Theorem 2.1, it suffices to check that $\left\|V_n^{-1}\xi_n - \zeta_n\right\|_\infty$ goes to zero in probability, where $\|f\|_\infty := \sup_{0\leq t\leq 1}|f(t)|$. To this end let us introduce the random change of time θ_n defined as follows. When $V_n > 0$, θ_n is the map from $[0,1]$ *onto* $[0,1]$ which interpolates linearly between the points $(k/n, V_k^2/V_n^2)$, $k = 0, 1, \ldots, n$. When $V_n = 0$, we simply take $\theta_n = I$, the identity on $[0,1]$. With the usual convention $S_k/V_n := 0$ for $V_n = 0$, we always have

$$\zeta_n\big(\theta_n(t)\big) = V_n^{-1}\xi_n(t), \quad 0 \leq t \leq 1. \tag{14}$$

Clearly for each $t \in [0,1]$,

$$\Big|\frac{V_{[nt]}^2}{V_n^2} - \theta_n(t)\Big| \leq \max_{1\leq k\leq n}\frac{X_k^2}{V_n^2}.$$

It follows by (6) that

$$\sup_{0\leq t\leq 1}\Big|\frac{V_{[nt]}^2}{V_n^2} - \theta_n(t)\Big| \xrightarrow{P} 0.$$

By Lemma 3.1 in [13],

$$\sup_{0\leq t\leq 1}\Big|\frac{V_{[nt]}^2}{V_n^2} - t\Big| \xrightarrow{P} 0.$$

So denoting by I the identity on $[0,1]$ we have

$$\|\theta_n - I\|_\infty \xrightarrow{P} 0. \tag{15}$$

Let $\omega(f;\delta) := \sup\{|f(t) - f(s)|;\ |t-s| \leq \delta\}$ denote the modulus of continuity of $f \in C[0,1]$. Then recalling (14) we have

$$\left\|V_n^{-1}\xi_n - \zeta_n\right\|_\infty = \sup_{0\leq t\leq 1}\left|V_n^{-1}\xi_n\big(\theta_n(t)\big) - \zeta_n\big(\theta_n(t)\big)\right| \leq \omega\big(V_n^{-1}\xi_n; \|\theta_n - I\|_\infty\big).$$

It follows that for any $\lambda > 0$ and $0 < \delta \leq 1$,

$$P\big(\left\|V_n^{-1}\xi_n - \zeta_n\right\|_\infty \geq \lambda\big) \leq P\big(\|\theta_n - I\|_\infty > \delta\big) + P\big(\omega(V_n^{-1}\xi_n;\delta) \geq \lambda\big). \tag{16}$$

Now since the Brownian motion has a version in $C[0,1]$, for each positive ε, we can find $\delta \in (0,1]$ such that $P\big(\omega(W;\delta) \geq \lambda\big) < \varepsilon$. As the functional ω is continuous on $C[0,1]$, it follows from Theorem 2.1 that

$$\limsup P\big(\omega(V_n^{-1}\xi_n;\delta) \geq \lambda\big) \leq P\big(\omega(W;\delta) \geq \lambda\big).$$

Hence for $n \geq n_1$ we have $P\big(\omega(V_n^{-1}\xi_n;\delta) \geq \lambda\big) < 2\varepsilon$. Having in mind (15) and (16) we see that the proof is complete.

3 Invariance principle in $D[0,1]$ under self-normalization

We now investigate the convergence in the Skorohod space of the step process $V_n^{-1}W_n$ which jumps at deterministic times $t = k/n$ and of the adaptive step process Z_n which jumps at random times $t = V_k^2/V_n^2$.

Theorem 3.1 *The convergence*

$$V_n^{-1}W_n \xrightarrow{\mathcal{D}} W \tag{17}$$

holds in the space $D[0,1]$ if and only if $X_1 \in DAN$.

Theorem 3.2 *The convergence*

$$Z_n \xrightarrow{\mathcal{D}} W \tag{18}$$

holds in the space $D[0,1]$ if and only if $X_1 \in DAN$.

Here again, the necessity of $X_1 \in DAN$ in both Theorems 3.1 and 3.2 follows from Giné, Götze and Mason [8].

Proof of Theorem 3.1. We have to prove tightness of the process W_n only. To this end we shall use Theorem 15.6 from Billingsley [4]. According to this theorem we have to find a nondecreasing continuous on $[0,1]$ function F such that

$$P(|W_n(t) - W_n(t_1)| \geq \lambda, |W_n(t_2) - W_n(t)| \geq \lambda) \leq \lambda^{-2\gamma}(F(t_2) - F(t_1))^{\alpha}, \tag{19}$$

for all $0 \leq t_1 \leq t \leq t_2$ and all $n \geq 1$, where $\gamma \geq 0$ and $\alpha > 1$.

Let us denote by $I_\lambda(t_1, t, t_2)$ the probability in the left hand side of (19). If $t_2 - t_1 \leq 1/n$ then either $t, t_1 \in [(k-1)/n, k/n)$ or $t, t_2 \in [(k-1)/n, k/n)$ for some $k = 1, \ldots, n$. In both cases $I_\lambda(t_1, t, t_2) = 0$, so we may and do assume that $t_2 - t_1 > 1/n$. By Markov inequality we have

$$I_\lambda(t_1, t, t_2) \leq \lambda^{-4}\mathbf{E}\, V_n^{-4}(S_{[nt]} - S_{[nt_1]})^2(S_{[nt_2]} - S_{[nt]})^2. \tag{20}$$

To treat this bound let us remark that due to independence and identical distribution of the X_i's, we have with $d := \sharp\{i_1, i_2, i_3, i_4\}$

$$\mathbf{E}\,\frac{X_{i_1}X_{i_2}X_{i_3}X_{i_4}}{V_n^4} = \begin{cases} \mathbf{E}\, V_n^{-4}X_1^2X_2^2 & \text{when} \quad d = 2; \\ \mathbf{E}\, V_n^{-4}X_1X_2X_3^2 & \text{when} \quad d = 3; \\ \mathbf{E}\, V_n^{-4}X_1X_2X_3X_4 & \text{when} \quad d = 4. \end{cases}$$

After some combinatorics, this leads to the upper bound

$$I_\lambda(t_1, t, t_2) \leq c\frac{(t_2 - t_1)^2}{\lambda^4}\Big[n^2\mathbf{E}\,\frac{X_1^2X_2^2}{V_n^4} + n^3\mathbf{E}\,\frac{X_1X_2X_3^2}{V_n^4} + n^4\mathbf{E}\,\frac{X_1X_2X_3X_4}{V_n^4}\Big].$$

It remains to check that the square brackets in the above estimate is bounded uniformly in n.

Since $X_1 \in DAN$, we have by (1) and (2) the stochastic boundedness of the sequence (S_n/V_n). Now from Lemmas 2.1, 2.2 and 2.4 in Giné, Götze and Mason [8], it is easily seen that for some constant K depending only on the distribution of X_1,

$$\begin{aligned} \mathbf{E}\, V_n^{-4} X_1^2 X_2^2 &\leq Kn^{-2}; \\ \mathbf{E}\, V_n^{-4} X_1 X_2 X_3^2 &\leq Kn^{-3}; \\ \mathbf{E}\, V_n^{-4} X_1 X_2 X_3 X_4 &\leq Kn^{-4}. \end{aligned}$$

The proof is complete.

It is worth noticing that without the hypothesis $X_1 \in DAN$, the stochastic boundedness of the sequence (S_n/V_n) is sufficient to obtain the tightness of the process $V_n^{-1} W_n$.

Proof of Theorem 3.2. We shall exploit Theorem 3.1 through the identity

$$Z_n(u) = V_n^{-1} W_n(\theta_n^{-1}(u)), \quad u \in [0,1],$$

where

$$\theta_n^{-1}(u) := \sup\{t \in [0,1];\ \theta_n(t) = u\}, \quad u \in [0,1],$$

is a generalized inverse of the random change of time introduced in the proof of Theorem 2.2. Observe that θ_n^{-1} is a non decreasing càdlàg function mapping $[0,1]$ into $[0,1]$. By a classical result on the random changes of time in $D[0,1]$ (see e.g. Billingsley [4] p. 145 and Th. 4.4 *ibidem*), the problem is easily reduced to check the convergence in probability to zero of $\left\|\theta_n^{-1} - I\right\|_\infty$. To this end we note that

$$\begin{aligned} \left\|\theta_n^{-1} - I\right\|_\infty &= \sup_{t\in[0,1]} \left|\theta_n^{-1}\big(\theta_n(t)\big) - \theta_n(t)\right| \\ &\leq \sup_{t\in[0,1]} \left|\theta_n^{-1}\big(\theta_n(t)\big) - t\right| + \left\|\theta_n - I\right\|_\infty . \end{aligned} \tag{21}$$

Now for each $t \in [0,1]$ we have

$$0 \leq \theta_n^{-1}\big(\theta_n(t)\big) - t \leq \frac{L_n}{n},$$

where L_n is the discrete random variable equal to the maximum length of runs of *consecutive* occurences among the events $\{X_i = 0\}$ $(1 \leq i \leq n)$. Put $p := P(X_1 = 0)$. Since the case $p = 0$ is obvious and the hypothesis $X_1 \in DAN$ excludes the trivial case $p = 1$, we are left with the case $0 < p < 1$ for which the elementary estimate

$$P\Big(\frac{L_n}{n} > \varepsilon\Big) \leq np^{n\varepsilon - 1} = o(1),$$

gives the required convergence in probability of $\left\|\theta_n^{-1} \circ \theta_n - I\right\|_\infty$. Together with (15) and (21), this makes the proof complete.

References

[1] Araujo, A. and Giné, E. (1980) *The Central Limit Theorem for Real and Banach Valued Random Variables.* Wiley, New York.

[2] Bentkus, V., Bloznelis, M. and Götze F. (1996) A Berry-Eséen bound for student's statistic in the non i.i.d. case *J.Theoret.Probab.* **9**, 765–796.

[3] Bentkus, V. and Götze, F. (1996) The Berry-Esseen bound for student's statistic *Annal. Probab.* **24**, 491–503

[4] Billingsley, P. (1968) *Convergence of probability measures.* Wiley, New York.

[5] Christiakov, G. P. and Götze, F. (1999) Moderate deviations for self-normalized sums. Preprint 99-048, SFB 343, Univ. Bielefeld.

[6] Chuprunov, A. N. (1997) On convergence of random polygonal lines under Student-type normalizations. *Theory of Probab. and Appl.* **41**, 756–761.

[7] Egorov, V. A. (1996) On the asymptotic behavior of self-normalized sums of random variables. *Theory of Probab. and Appl.* **41**, 643–650.

[8] Giné, E., Götze, F. and Mason, D. (1997) When is the Student t-statistic asymptotically standard normal? *Ann. Probab.* **25**, 1514–1531.

[9] Griffin, P. S. and Kuelbs, J. (1989) Self-normalized laws of the iterated logarithm. *Ann. Probab.* **17**, 1571–1601.

[10] Griffin, P. S. and Mason, D. M. (1991) On the asymptotic normality of self-normalized sums. *Proc. Cambridge Phil. Soc.* **109**, 597–610.

[11] Logan, B. F., Mallows, C. L., Rice, S. O. and Shepp, L. A. (1973) Limit distributions of self-normalized sums. *Ann. Probab.* **1**, 788–809.

[12] O'Brien, G. L. (1980) A limit theorem for sample maxima and heavy branches in Galton-Watson trees. *J. Appl. Probab.* **17**, 539–545.

[13] Račkauskas, A. and Suquet, Ch. (2000) Invariance principle for self-normalized sums of symmetric random variables, *Pub. IRMA Lille* **51-III**.

[14] Shao, Q.-M. (1997) Self-normalized large deviations, *Ann. Probab.* **25**, 285–328.

[15] Shao, Q.-M. (1999) A Cramér type large deviation result for Student's t-statistic. *J. Theoret. Probab.* **12**, 387-398.