AUSTRIAN JOURNAL OF STATISTICS
Volume 35 (2006), Number 2&3, 93–103

Change in the Mean in the Domain of Attraction of the Normal Law

Miklós Csörgő[1], Barbara Szyszkowicz[1], and Qiying Wang[2]
[1]Carleton University, Ottawa, Canada
[2]University of Sydney, Australia

Abstract: Some weighted approximations in probability of self-normalized and Studentized partial sums processes are reviewed and also described in the context of studying the problem of change in the mean of random variables in the domain of attraction of the normal law. This survey of such results constitutes an extended abstract of the talk with the same title that was presented by Miklós Csörgő on July 18, 2005 in Mikulov, based on the joint works M. Csörgő, B. Szyszkowicz, and Q. Wang (2001), (2003) and (2004) by the three of us.

Keywords: Studentized and Self-Normalized Sums, Weighted Approximations in Probability, Brownian Motion, Brownian Bridge.

1 Introduction

Let $X, X_1, X_2, \dots$ be non-degenerate i.i.d. r.v.'s. Put

$$S_n = \sum_{i=1}^{n} X_i\,, \quad \overline{X}_n = S_n/n\,, \quad V_n^2 = \sum_{i=1}^{n} X_i^2\,, \qquad n \in \mathbb{N}\,,$$

and define the classical Student t-statistic

$$\begin{aligned} T_n(X) &= \frac{(1/n^{1/2})\sum_{i=1}^n X_i}{((1/(n-1))\sum_{i=1}^n (X_i - \overline{X}_n)^2)^{1/2}} \\ &= \frac{S_n/V_n}{\sqrt{(n-(S_n/V_n)^2)/(n-1)}}\,. \end{aligned} \tag{1}$$

Assuming $X \stackrel{\mathcal{D}}{=} N(\mu, \sigma^2)$, we obtain Student's ratio (cf. W.S. Gosset $\equiv$ "Student", 1908)

$$\begin{aligned} T_n(X-\mu) &= \frac{(1/n^{1/2})\sum_{i=1}^n (X_i-\mu)/\sigma}{((1/(n-1))\sum_{i=1}^n ((X_i - \overline{X}_n)/\sigma)^2)^{1/2}} \\ &= \frac{\sum_{i=1}^n (X_i-\mu)/V_n}{\sqrt{(n-(S_n/V_n)^2)/(n-1)}} \end{aligned} \tag{2}$$

a so-called t-random variable with $(n-1)$ degrees of freedom.

In view of (1), if $T_n(X)$ or S_n/V_n has an asymptotic distribution, then so does the other, and it is well known that they coincide. Hence, usually and without loss of generality, the limiting distribution of S_n/V_n is studied in the literature.

For example, in Logan et al. (1973) conjecture that "S_n/V_n is asymptotically normal if (and perhaps only if) X is in the domain of attraction of the normal law". It is the "only if" part that has remained open until 1997 for the general case of not necessarily symmetric random variables, when Gine et al. (1997) proved

Theorem A. *The following two statements are equivalent:*

(i) X is in the domain of attraction of the normal law;

(ii) There exists a finite constant μ such that, as $n \to \infty$, $T_n(X-\mu) \xrightarrow{\mathcal{D}} N(0,1)$.

Moreover, if either (i) or (ii) holds, then $\mu = EX$.

Gine et al. (1997) also show that, *if the self-normalized sums S_n/V_n, $n \in \mathbb{N}$, are stochastically bounded, then they are uniformly sub-Gaussian in the sense that*

$$\sup_{n\in\mathbb{N}} \mathrm{E}e^{tS_n/V_n} \leq 2e^{ct^2} \text{ for all } t \in \mathbb{R} \text{ and some } c < \infty \,.$$

This, in turn, implies a basic requirement in the proof of this result that the moments of S_n/V_n converge to those of a $N(0,1)$ r.v. whenever S_n/V_n is asymptotically standard normal.

For characterizations of S_n/V_n being stochastically bounded, we refer to Gine and Mason (1998), and Griffin (2002). Moreover, Chistyakov and Götze (2001) have recently confirmed a second conjecture of Logan et al. (1973) that the Student t–statistic has a non-trivial limiting distribution if and only if X is in the domain of attraction of a stable law with some exponent $\alpha \in [0,2]$.

In view of Theorem A, it is natural to ask whether it could possibly be extended to become a Donsker type functional central limit theorem. In this regard, writing $X \in$ DAN for X being in the domain of attraction of the normal law, we restate here Theorem 1 of Csörgő et al. (2001), and Csörgő et al. (2003) *à la* Theorem A. Consider the sequence $T_{n,t}$ of Student processes in $t \in [0,1]$ on $D[0,1]$ in terms of i.i.d. r.v.'s and notation as in (1), defined as

$$\begin{aligned}\{T_{n,t}(X)\,,\ 0 \leq t \leq 1\} \ &:= \left\{ \frac{(1/n^{1/2})\sum_{i=1}^{[nt]} X_i}{\left((1/(n-1))\sum_{i=1}^{n}(X_i-\bar{X}_n)^2\right)^{1/2}}\,,\ 0 \leq t \leq 1 \right\} \\ &= \left\{ \frac{S_{[nt]}/V_n}{\sqrt{(n-(S_n/V_n)^2)/(n-1)}}\ \ 0 \leq t \leq 1 \right\}. \qquad (3)\end{aligned}$$

Clearly, $T_{n,1}(X) = T_n(X)$, with the latter as in (1).

Theorem 1.1 *As $n \to \infty$, the following statements are equivalent:*

(a) $X \in$ DAN and $EX = \mu$;

(b) $T_{n,t_0}(X-\mu) \xrightarrow{\mathcal{D}} N(0,1)$, for $t_0 \in (0,1]$;

(c) $T_{n,t}(X-\mu) \xrightarrow{\mathcal{D}} W(t)$ on $(D[0,1],\rho)$, where ρ is the sup-norm metric for functions in $D[0,1]$, and $\{W(t), 0 \leq t \leq 1\}$ is a standard Wiener process;

(d) *On the probability space of Theorem 1(d) of Csörgő et al. (2001), and Csörgő et al. (2003) with the there constructed standard Wiener process* $\{W(t), 0 \le t < \infty\}$ *we have*

$$\sup_{0\le t\le 1} |T_{n,t}(X-\mu) - W(nt)/\sqrt{n}| = o_P(1)\,.$$

Proof Clearly, (d) implies (c), (c) implies (b) and, on account of Theorem A, the latter with $t_0 = 1$ implies (a). Hence, one only needs to show that (a) implies (d). For showing this, we refer to Csörgő et al. (2001), Csörgő et al. (2003).

Remark 1.1 What we mean by saying that (d) implies (c) is the following functional central limit theorem. *On account of (d), as* $n \to \infty$*, we have*

$$h\{T_{n,\cdot}(X-\mu)\} \xrightarrow{\mathcal{D}} h\{W(\cdot)\} \tag{4}$$

for all $h : D \to \mathbb{R}$ *that are* $(D, \mathbb{D})$ *measurable and* ρ*-continuous, or* ρ*-continuous except at points forming a set of Wiener measure zero on* $(D, \mathbb{D})$*, where* $\mathbb{D}$ *denotes the* σ*-field of subsets of* D *generated by the finite-dimensional subsets of* D*.*

For studying the probable error of a change in a mean, we need a weighted approximation type extension of Theorem 1.1.

Let Q be the class of positive functions $q(t)$ on $(0, 1]$, i.e., $\inf_{\delta\le t\le 1} q(t) > 0$ for $0 < \delta < 1$, which are nondecreasing near zero, and let

$$I(q,c) = \int_{0+}^{1} t^{-1} \exp(-cq^2(t)/t)dt\,, \quad 0 < c < \infty\,.$$

The next two lemmas are due to Csörgő et al. (1986).

Lemma A *Let* $\{W(t)\,, 0 \le t < \infty\}$ *be a standard Wiener process and* $q(t) \in Q$*. If* $I(q, c) < \infty$ *for some* $c > 0$*, then*

$$\lim_{t\downarrow 0} t^{1/2}/q(t) = 0 \ \textit{ and } \ \limsup_{t\downarrow 0} |W(t)|/q(t) < \infty \quad a.s.$$

Conversely, if $\limsup_{t\downarrow 0} |W(t)|/q(t) < \infty$ *a.s., then* $I(q, c) < \infty$ *for some* $c > 0$*.*

Lemma B *Let* $\{W(t)\,, 0 \le t < \infty\}$ *be a standard Wiener process and* $q(t) \in Q$*. Then* $I(q, c) < \infty$ *for any* $c > 0$*, if and only if*

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

Now, more generally than in Theorem 1.1, we also have the following weighted approximations for the sequence of Student processes $\{T_{n,t}, 0 \le t \le 1\}$ of (3) (cf. Theorem 2 of Csörgő et al., 2001),

Theorem 1.2 *Assume that* $EX = \mu$ *and* $X \in$ DAN*. Then, on an appropriate probability space for* $X, X_1, X_2, \ldots$*, a standard Wiener process* $\{W(t)\,, 0 \le t < \infty\}$ *can be constructed in such a way that the following statements hold true.*

(a) Let $q \in Q$. Then, as $n \to \infty$,

$$\sup_{1/n \le t \le 1} \Big| T_{n,t}(X-\mu) - W(nt)/\sqrt{n} \Big| / q(t) = \begin{cases} o_P(1) & \text{if } I(q,c) < \infty \text{ for any } c > 0\,, \\ O_P(1) & \text{if } I(q,c) < \infty \text{ for some } c > 0\,. \end{cases}$$

(b) Let $q \in Q$. Then, as $n \to \infty$

$$\sup_{0 < t \le 1} |T_{n,t}(X-\mu) - W(nt)/\sqrt{n}\,|/q(t) = O_P(1)$$

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

(c) Let $q \in Q$. Then, as $n \to \infty$

$$\sup_{0 < t \le 1} |T_{n,t}(X-\mu) - W(nt)/\sqrt{n}\,|/q(t) = o_P(1)$$

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

The condition that $X \in$ DAN is also necessary for establishing part (c) of Theorem 1.2. Indeed, if part (c) of Theorem 1.2 holds, then, as $n \to \infty$, we have

$$\sup_{0 \le t \le 1} \Big| T_{n,t}(X-\mu) - W(nt)/\sqrt{n}| = o_P(1)\,.$$

Hence Theorem 1.1 via Theorem A implies that $\mathrm{E}X = \mu$ and $X \in$ DAN.

For proof of Theorem 1.2 we refer to Csörgő et al. (2001).

Corollary 1.1 *Let $EX = \mu$ and $X \in$ DAN. Let $q \in Q$, and $\{W(t)\,, 0 \le t \le 1\}$ be a standard Wiener process. Then we have, as $n \to \infty$,*

$$T_{n,t}(X-\mu)/q(t) \xrightarrow{\mathcal{D}} W(t)/q(t) \;\; \text{on } (D[0,1]\,,\rho)$$

if and only if $I(q,c) < \infty$ for any $c > 0$, i.e., $\lim_{t \downarrow 0} |W(t)|/q(t) = 0$ a.s. (cf. Lemma B).

Finally, we state convergence in distribution results for sup-functionals of weighted self-normalized partial sums for the optimal class of weight functions which *coincides with* the class of functions $q \in Q$ satisfying $\limsup_{t \downarrow 0} |W(t)|/q(t) < \infty\ a.s.$ (see Lemma A). Consequently, the results that follow are NOT implied by Corollary 1.1 above, i.e., they cannot be obtained via classical methods of weak convergence. “Tightness” is NOT guaranteed by Lemma A!

Definition for use later on: $\log x = \log(\max\{e, x\})$.

Corollary 1.2 *Let $EX = \mu$ and $X \in$ DAN. Let $q \in Q$ and $\{W(t)\,, 0 \le t < \infty\}$ be a standard Wiener process. If, in addition to $q \in Q$, $q(t)$ is nondecreasing on $(0,1]$, then*

$$\sup_{0 < t \le 1} T_{n,t}(X-\mu)/q(t) \xrightarrow{\mathcal{D}} \sup_{0 < t \le 1} |W(t)|/q(t)$$

if and only if $I(q,c)<\infty$ *for some* $c>0$. *In particular, as* $n\to\infty$, *we have*

$$\sup_{0<t\le 1} T_{n,t}(X-\mu)/(t\log\log(t^{-1}))^{1/2} \xrightarrow{\mathcal{D}} \sup_{0<t\le 1} |W(t)|/(t\log\log(t^{-1}))^{1/2} .$$

Remark 1.2 It is interesting and also of interest to note that the class of the weight functions in Corollary 1.2 is bigger than that in Corollary 1.1. Such a phenomenon was first noticed and proved for weighted empirical and quantile processes by Csörgő et al. (1986) and then by Csörgő and Horváth (1988) for partial sums on assuming $\mathrm{E}|X|^{\upsilon}<\infty$ for some $\upsilon>2$. For more details along these lines when only $\mathrm{E}X^2<\infty$ is assumed, we refer to Szyszkowicz (1991), Szyszkowicz (1996), Szyszkowicz (1997), Csörgő et al. (1999), Csörgő and Horváth (1993), and to Csörgő and Norvaiša (2004).

2 Change in the Mean in the Domain of Attraction of the Normal Law

We are to illustrate how self-normalized weighted approximations can be made use of in change-point analysis when studying the problem of change in the mean of independent observations.

Let $X_1,\ldots,X_n$, $n\ge 1$, be non-degenerate independent real-valued r.v.'s with finite means. Suppose we wish to test the *nonparametric* "no-change in the mean" null hypothesis

$$H_0 : \mathrm{E}X_1=\mathrm{E}X_2=\cdots=\mathrm{E}X_n$$

against the "at most one change in the mean" (AMOC) alternative hypothesis

$$\begin{aligned} H_A :\quad & \text{there is an integer } k^*,\, 1\le k^*<n\,, \text{ such that} \\ & \mathrm{E}X_1=\cdots=\mathrm{E}X_{k^*}\ne \mathrm{E}X_{k^*+1}=\cdots=\mathrm{E}X_n\,. \end{aligned}$$

The hypothesized time k^* of at most one change in the mean is usually unknown. Hence, given *chronologically ordered* independent observables $X_1,X_2,\ldots,X_n$, $n\ge 1$, in order to test H_0 versus H_A, from a *non-parametric* point of view it appears to be reasonable to compare the sample mean $(X_1+\cdots+X_k)/k=:S_k/k$ at any time $1\le k<n$ to the sample mean $(X_{k+1}+\cdots+X_n)/(n-k)=:(S_n-S_k)/(n-k)$ after time $1\le k<n$ via functionals in k of the family of the standardized statistics

$$\begin{aligned} V_n(k) &:= \left(n\frac{k}{n}\left(1-\frac{k}{n}\right)\right)^{1/2}\left(\frac{S_k}{k}-\frac{S_n-S_k}{n-k}\right) \\ &= \frac{1}{\left(\frac{k}{n}\left(1-\frac{k}{n}\right)\right)^{1/2}}\left(\frac{S_k}{n^{1/2}}-\left(\frac{k}{n}\right)\frac{S_n}{n^{1/2}}\right), \quad 1\le k<n\,. \end{aligned} \tag{5}$$

For instance, one would want to reject H_0 in favor of H_A for large observed values of

$$V_n := \max_{1\le k<n} |V_n(k)|\,. \tag{6}$$

On the other hand, when assuming for example that the independent observables $X_1, \ldots, X_n$, $n \geq 1$, are $N(\mu, \sigma^2)$ random variables, then we find ourselves modelling and testing for a *parametric* shift in the mean AMOC problem. It is, however, easy to check that, when the variance σ^2 is known, then

$$-2 \log \Lambda_k = \frac{1}{\sigma^2}(V_n(k))^2 , \tag{7}$$

where Λ_k is the *likelihood ratio statistic* if the change in the mean occurs at $k^* = k$. Hence, the *maximally selected likelihood ratio statistic* $\max_{1 \leq k < n}(-2 \log \Lambda_k)$ will be large if and only if V_n of (5) is large.

A similar conclusion holds true if the variance σ^2 is an unknown but constant nuisance parameter (cf. Gombay and Horváth, 1994; Gombay and Horváth, 1996a; Gombay and Horváth, 1996b; and Csörgő and Horváth (1997), Section 1.4, and references therein). Namely in this case the maximally selected likelihood ratio statistic

$$\max_{1 \leq k < n} (-2 \log \Lambda_k)$$

will be large if and only if

$$\hat{V}_n := \max_{1 \leq k < n} \frac{1}{\hat{\hat{\sigma}}_k} |V_n(k)| \tag{8}$$

is large, where

$$\hat{\hat{\sigma}}_k^2 := \frac{1}{n} \left\{ \sum_{1 \leq i \leq k} \left(X_i - \frac{S_k}{k} \right)^2 + \sum_{k < i \leq n} \left(X_i - \frac{S_n - S_k}{n - k} \right)^2 \right\} . \tag{9}$$

These conclusions, and further examples as well in Csörgő and Horváth (1988) [Section 2], and in Csörgő and Horváth (1997) [Section 1.4] that are based on Gombay and Horváth (1994), Gombay and Horváth (1996a), Gombay and Horváth (1996b) show that under the null hypothesis H_0 a large number of parametric and nonparametric modelling of AMOC problems result in the same test statistic, namely that of (5), or its variant in (8). Consequently, if the underlying distribution is not known, the just mentioned test statistics should continue to work just as well when testing for H_0 versus H_A as above. Furthermore, Brodsky and Darkhovsky (1993) argue quite convincingly in their Section 1.2 that detecting changes in the mean (mathematical expectation) of a random sequence constitutes one basic situation to which other changes in distribution can be conveniently reduced. Thus V_n and $\hat{V}_n$ gain a somewhat focal role in change-point analysis in general as well. Studying the asymptotic behavior of these statistics is clearly of interest and, as we will now see, also interesting.

Here we are to deal with testing H_0 in the i.i.d. case against the AMOC alternative H_A as before, i.e., we consider testing for

$$H_0' : X_1, X_2, \ldots, X_n \text{ being a random sample on } X \text{ with a finite mean } \mathrm{E}X = \mu$$

versus H_A, via the family of the standardized statistics $\{V_n(k)\, 1 \leq k < n\}$ (cf. (5)). Let $S_0 = 0$, and for $n \geq 1$ define the sequence of tied-down partial sums processes

$$Z_n(t) := \begin{cases} (S_{[(n+1)t]} - [(n+1)t] S_n / n) / n^{1/2} , & 0 \leq t < 1 , \\ 0 , & t = 1 . \end{cases} \tag{10}$$

Given H_0', and assuming also that $\sigma^2 := \text{Var}\, X < \infty$, as $n \to \infty$, an application of Donsker's theorem yields

$$\begin{aligned} \frac{1}{\sigma} Z_n(t) &\xrightarrow{\mathcal{D}} W(t) - tW(1) \;\text{ on } (D[0,1],\rho) \\ &\stackrel{\mathcal{D}}{=} B(t)\,, \quad 0 \le t \le 1\,, \end{aligned} \tag{11}$$

where the mean zero Gaussian process $\{B(t)\,, 0 \le t \le 1\}$ with covariance function

$$\mathrm{E}B(s)B(t) = s \wedge t - st$$

is called *Brownian bridge* (or tied-down Brownian motion).

In view of (5), we are interested in exploring further possibilities of weak convergence for the standardized sequence of stochastic processes

$$\left\{ \frac{1}{(t(1-t))^{1/2}} Z_n(t)\,, \quad 0 < t < 1 \right\}. \tag{12}$$

An application of Donsker's theorem immediately yields that, as $n \to \infty$, under H_0' with finite positive variance $\sigma^2 = \text{Var}\, X$ we have (cf. (11))

$$\frac{1}{\sigma} Z_n(t)/(t(1-t))^{1/2} \xrightarrow{\mathcal{D}} B(t)/(t(1-t))^{1/2} \text{ on } (D[a,b],\rho)\,, \tag{13}$$

where $0 < a < b < 1$, $t \in [a,b]$ and $B(\cdot)$ is a Brownian bridge.

This however is not a satisfactory solution for testing H_0' versus H_A even if the variance σ^2 were known. This is due to the arbitrariness of choosing $a, b \in (0,1)$ which, in turn, may exclude observing possible changes in the mean "near 0" with $t \in (0,a)$ and/or "near n" with $t \in (b,1)$. We note also that,

$$\sup_{0<t<1} \frac{1}{\sigma} |Z_n(t)|/(t(1-t))^{1/2}$$

and, naturally, also the standardized statistics V_n and $\hat{V}_n$ (cf. (6) and (8)) converge in distribution to ∞ as $n \to \infty$ *even if the null assumption of no change in the mean were true*. Hence, in order to secure nondegenerate limiting behavior under H_0', we seek appropriate renormalizations.

Let $\mathcal{Q}^*$ be the class of *positive functions* $q^* : (0,1) \to (0,\infty)$, i.e., $\inf_{\delta \le t \le 1-\delta} q^*(t) > 0$ for all $\delta \in (0,1/2)$, which are *non-decreasing near 0* and *non-increasing near 1*. For $q^* \in \mathcal{Q}^*$, we define the integral (cf. definition of $I(q,c)$ earlier)

$$I_0^1(q^*,c) := \int_{0+}^{1-} (t(1-t))^{-1} \exp(-cq^{*2}(t)/(t(1-t)))dt\,, \quad c > 0\,. \tag{14}$$

With $Z_n(\cdot)$ as in (10), define the sequence of partial sums processes

$$T_n(t) := n^{1/2} Z_n(t) = \begin{cases} S_{[(n+1)t]} - [(n+1)t]S_n/n\,, & 0 \le t < 1\,, \\ 0\,, & t = 1\,. \end{cases} \tag{15}$$

Theorem 2.1 *Let $q^* \in Q^*$. Assume H_0' and that $X \in$ DAN. Then, on an appropriate probability space for $X, X_1, X_2, \ldots$, a sequence of Brownian bridges $\{B_n(t)\,, 0 \le t \le 1\}$ can be constructed such that, as $n \to \infty$,*

$$\sup_{0<t<1} \left| \frac{1}{\hat{\sigma}_{[nt+1]}} T_n(t) - B_n(t) \right| \Big/ q^*(t) = o_P(1)\,, \tag{16}$$

if and only if $I_0^1(q^*, c) < \infty$ for all $c > 0$, where

$$\hat{\sigma}_k^2 := n\hat{\hat{\sigma}}_k^2 = \begin{cases} \sum_{1\le i\le k}(X_i - \hat{X}_k)^2 + \sum_{k<i\le n}(X_i - X_k^*)^2\,, & \text{if } 1 \le k < n \\ \sum_{1\le i\le n}(X_i - \hat{X}_n)^2\,, & \text{if } k = n\,, \end{cases} \tag{17}$$

with

$$\hat{X}_k = \frac{1}{k}\sum_{1\le j\le k} X_j \quad \text{and} \quad X_k^* = \frac{1}{n-k}\sum_{k<j\le n} X_j\,.$$

Proof We refer to pages 507–509 of Csörgő, Szyszkowicz, and Wang (2004) and note in passing that, assuming H_0' and $\mathrm{E}X^2 < \infty$, (16) is established on pages 74 and 75 of Csörgő and Horváth (1997).

Remark 2.1 We note that when proving Theorem 2.1, we may assume without loss of generality that $\mathrm{E}X = 0$. Hence, by part (c) of Theorem 1.2, mutatis mutandis, there exist two independent standard Wiener processes $\{W^{(1)}(t)\,, 0 \le t < \infty\}$ and $\{W^{(2)}(t)\,, 0 \le t < \infty\}$ such that, as $n \to \infty$,

$$\sup_{0<t\le 1/2} |S_{[nt]}/V_n - W^{(1)}(nt)/\sqrt{n}|/q^*(t) = o_P(1)\,,$$

and

$$\sup_{1/2\le t\le 1} |(S_n - S_{[nt]})/V_n - W^{(2)}(n-nt)/\sqrt{n}|/q^*(t) = o_P(1)\,.$$

These two conclusions constitute the initial steps of the proof of Theorem 2.1 (cf. pages 507–509 of Csörgő et al., 2004).

Remark 2.2 With $q^* \in Q^*$, the integral condition for having (16) amounts to saying that the latter statement holds true if and only if (cf. Csörgő et al., 1986 and Lemma B)

$$\lim_{t\downarrow 0} |B(t)|/q^*(t) = \lim_{t\uparrow 1} |B(t)|/q^*(t) = 0 \quad a.s.$$

As a direct consequence of Theorem 2.1, we conclude weighted weak convergence in sup-norm metric as follows.

Corollary 2.1 *Let $\{B(t)\,, 0 \le t \le 1\}$ be a Brownian bridge, $q^* \in Q^*$ and $I_0^1(q^*, c) < \infty$ for all $c > 0$. Assume H_0'. If $X \in$ DAN, then as $n \to \infty$,*

$$\frac{1}{\hat{\sigma}_{[nt+1]}} T_n(t)/q^*(t) \xrightarrow{\mathcal{D}} B(t)/q^*(t) \ \text{ on } (D[0,1], \rho)\,. \tag{18}$$

For the notion of weak convergence of (18) *à la* Remark 1.1, we refer to Remark 3.4 of Csörgő et al. (2004).

Via (18), appropriate functionals converge in distribution. For example, whenever $I_0^1(q^*, c) < \infty$ for all $c > 0$ with $q^* \in Q^*$, as $n \to \infty$, we have

$$\sup_{0<t<1} \frac{1}{\hat{\sigma}_{[nt+1]}} |T_n(t)|/q^*(t) \xrightarrow{\mathcal{D}} \sup_{0<t<1} |B(t)|/q^*(t)\,. \tag{19}$$

On the other hand, further work yields the following result (cf. Corollary 5.2 in Csörgő et al., 2004):

Corollary 2.2 *Let* $\{B(t)\,, 0 \le t \le 1\}$ *be a Brownian bridge, assume* H_0' *and that* $X \in$ DAN. *If, in addition to* $q^* \in Q^*$, $q^*(t)$ *is nondecreasing on* $(0, 1/2]$ *and nonincreasing on* $[1/2, 1)$, *and* $I_0^1(q^*, c) < \infty$ *for some* $c > 0$, *then, as* $n \to \infty$,

$$\sup_{0<t<1} \frac{1}{\hat{\sigma}_{[nt+1]}} |T_n(t)|/q^*(t) \xrightarrow{\mathcal{D}} \sup_{0<t<1} |B(t)|/q^*(t) \tag{20}$$

and, in particular, with

$$q_1^*(t) = \begin{cases} (t \log\log(t^{-1}))^{1/2}\,, & \textit{if } t \in (0, 1/2]\,, \\ ((1-t)\log\log(1-t)^{-1})^{1/2}\,, & \textit{if } t \in [1/2, 1)\,, \end{cases}$$

$$\sup_{0<t<1} \frac{1}{\hat{\sigma}_{[nt+1]}} |T_n(t)|/q_1^*(t) \xrightarrow{\mathcal{D}} \sup_{0<t<1} |B(t)|/q_1^*(t)\,. \tag{21}$$

Large values of the statistics on the left hand sides in (2.15), (20) and (21) respectively, indicate a change in the mean, and hence, based on these corollaries, rejection of H_0' can be quantified accordingly. *For tables* and further weight functions along these lines we refer to Orasch and Pouliot (2004).

Remark 2.3 The finiteness of the integral $I_0^1(q^*, c)$ for some $c > 0$ with $q^* \in Q^*$ so that (20) should hold true amounts to saying that the latter statement obtains for such $q^* \in Q^*$ if and only if (cf. Csörgő et al., 1986 and Lemma A)

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

and

$$\limsup_{t \uparrow 1} |B(t)|/q^*(t) < \infty \quad \text{a.s.}$$

Thus, even though we no longer have weak convergence as in (18), we still have (20). In other words, while (18) implies (19), it does not imply (20) in general, nor (21) in particular. The here presented Theorem 2.1 and Corollaries 2.1, 2.2 constitute first steps in studying the problem of change in the mean in the domain of attraction of the normal law with possibly infinite variance.

Remark 2.4 Theorem 2.1 and Corollaries 2.1, 2.2 continue to hold true if $\hat{\sigma}_{[nt+1]}$ is replaced by $\hat{\sigma}_n$ (cf. (8)). In case of $\mathrm{E}X^2 < \infty$ in Theorem 2.1 $\{\hat{\sigma}_k^2/n\}$ is a weakly uniformly consistent sequence of estimators for $\sigma^2 := \text{Var}\, X$. Namely, given H_0' and that $\mathrm{E}X^2 < \infty$, as $n \to \infty$ we have (cf. Lemma 2.1.1 of Csörgő and Horváth, 1997

$$\max_{1 \le k \le n} \left| \hat{\sigma}_k^2/n - \sigma^2 \right| = o_P(1)\,. \tag{22}$$

When testing H_0' versus H_A in case of a $N(\mu, \sigma^2)$ random sample, the Gombay and Horváth (1994), (1996a, b) based Section 1.4 of Csörgő and Horváth (1997) suggests the use of the pooled sample sums of squares $\hat{\sigma}_k^2$ in lieu of $\hat{\sigma}_n^2$ (cf. (8)) in general as well, for the sake of gaining more powerful tests in this regard. Hence is the continued initial use of $\hat{\sigma}_{[nt+1]}$ in (16) as well in the general nonparametric context of assuming H_0' together with $\mathrm{E}X^2 < \infty$. And hence is also the initial use of $\hat{\sigma}_{[nt+1]}$ for self-normalization in Theorem 2.1 and Corollaries 2.1, 2.2 under H_0' with $X \in$ DAN, even though we no longer have anything like (22) any more when $\mathrm{E}X^2 = \infty$.

References

Brodsky, B. E., and Darkhovsky, B. S. (1993). *Nonparametric Methods in Change-Point Problems*. Dordrecht: Kluwer.

Chistyakov, G. P., and Götze, F. (2001). Limit distributions of studentized means. *Preprint*.

Csörgő, M., Csörgő, S., Horváth, L., and Mason, D. M. (1986). Weighted empirical and quantile processes. *The Annals of Probability*, *14*, 31-85.

Csörgő, M., and Horváth, L. (1988). Nonparametric methods for changepoint problems. In P. R. Krishnaiah and C. R. Rao (Eds.), *Quality Control and Reliability* (Vol. 7, p. 403-425). Amsterdam: Elsevier.

Csörgő, M., and Horváth, L. (1993). *Weighted Approximations in Probability and Statistics*. New York: Wiley.

Csörgő, M., and Horváth, L. (1997). *Limit Theorems in Change-Point Analysis*. New York: Wiley.

Csörgő, M., and Norvaiša, R. (2004). Weighted invariance principle for Banach space valued random variables. *Lietuvos Matematikos Rinkinis*, *44*, 139–175.

Csörgő, M., Norvaiša, R., and Szyszkowicz, B. (1999). Convergence of weighted partial sums when the limiting distribution is not necessarily Radon. *Stochastic Processes and their Applications*, *81*, 81-101.

Csörgő, M., Szyszkowicz, B., and Wang, Q. (2001). Donsker's theorem and weighted approximations for self-normalized partial sums processes. *Technical Report Series of the Laboratory for Research in Statistics and Probability, Carleton University, Ottawa*, *360*.

Csörgő, M., Szyszkowicz, B., and Wang, Q. (2003). Donsker's theorem for self-normalized partial sums processes. *The Annals of Probability*, *31*, 1228-1240.

Csörgő, M., Szyszkowicz, B., and Wang, Q. (2004). On weighted approximations and strong limit theorems for self-normalized partial sums processes. In L. Horváth and B. Szyszkowicz (Eds.), *Asymptotic Methods in Stochastics* (Vol. 44, p. 489-521). AMS, Providence, Rhode Island: Fields Institute Communications.

Gine, E., Götze, F., and Mason, D. M. (1997). When is the Student t-statistic asymptotically normal? *The Annals of Probability*, *25*, 1514-1531.

Gine, E., and Mason, D. M. (1998). On the LIL for self-normalized sums of IID random variables. *Journal of Theory of Probability*, *11*, 351-370.

Gombay, E., and Horváth, L. (1994). An application of the maximum likelihood test to the change-point problem. *Stochastic Processes and their Applications*, *50*, 161-171.

Gombay, E., and Horváth, L. (1996a). Applications for the time and change and the power function in change-point models. *Journal of Statistical Planning and Inference*, *52*, 43-66.

Gombay, E., and Horváth, L. (1996b). On the rate of approximations for maximum likelihood test in chanage-point models. *Journal of Multivariate Analysis*, *56*, 120-152.

Griffin, P. S. (2002). Tightness of the Student t-statistic. *Electronic Communications in Probability*, *7*, 181-190.

Logan, B. F., Mallows, C. L., Rice, S. O., and Shepp, L. A. (1973). Limit distributions of self-normalized sums. *The Annals of Probability*, *1*, 788–809.

Orasch, M., and Pouliot, W. (2004). Tabulating weighted sup-norm functionals used in change-point analysis. *Journal of Statistical Computation and Simulation*, *74*, 249-276.

"Student". (1908). The probable error of a mean. *Biometrika*, *6*, 1-25.

Szyszkowicz, B. (1991). Weighted stochastic processes under contiguous alternatives. *Comptes Rendus Mathématiques de l'Académie des Sciences, La Société Royale du Canada*, *13*, 211-216.

Szyszkowicz, B. (1996). Weighted approximations of partial sum processes in $d[0,\infty)$. i. *Studia Scientiarum Mathematicarum Hungarica*, *31*, 323-353.

Szyszkowicz, B. (1997). Weighted approximations of partial sum processes in $d[0,\infty)$. ii. *Studia Scientiarum Mathematicarum Hungarica*, *33*, 305-320.

Author's address:

Miklós Csörgő
Laboratory for Research on Statistics and Probability
Carleton University
Ottawa, Ontario
1125 Colonel By Drive
Canada K1S 5B6

E-mail: mcsorgo@math.carleton.ca