Elect. Comm. in Probab. **7** (2002) 181–190

ELECTRONIC COMMUNICATIONS in PROBABILITY

TIGHTNESS OF THE STUDENT T-STATISTIC

PHILIP S. GRIFFIN
Department of Mathematics, Syracuse University,
Syracuse, NY 13244-1150, U.S.A.
email: psgriffi@syr.edu

submitted March 19, 2002 *Final version accepted* September 27, 2002

AMS 2000 Subject classification: 60F05
Tightness, t-statistic, self-normalized sum

Abstract
Let $X, X_1, X_2, \ldots$ be a sequence of nondegenerate, independent and identically distributed random variables and set $S_n = X_1 + \cdots + X_n, V_n^2 = X_1^2 + \cdots + X_n^2$. We answer a question of Götze, Giné and Mason by providing a simple necessary and sufficient condition for tightness of S_n/V_n.

1 Introduction

Let $X, X_1, X_2, \ldots$ be a sequence of nondegenerate, independent and identically distributed random variables and set

$$S_n = X_1 + \cdots + X_n, \quad V_n^2 = X_1^2 + \cdots + X_n^2 \tag{1.1}$$

and

$$t_n = \frac{S_n}{\sqrt{\frac{n}{n-1}\sum_1^n (X_i - \overline{X})^2}}, \tag{1.2}$$

where $\overline{X} = n^{-1}S_n$. Then t_n is the classical Student t-statistic which may be expressed equivalently as

$$t_n = \frac{S_n}{V_n}\sqrt{\frac{n-1}{n - (\frac{S_n}{V_n})^2}}. \tag{1.3}$$

In a beautiful paper Götze, Giné and Mason (1997) solved a long standing conjecture of Logan, Mallows, Rice and Shepp (1973), by proving that t_n, or equivalently the self-normalized sum S_n/V_n (see Proposition 1), is asymptotically standard normal if and only if X is in the domain of attraction of the normal law and $EX = 0$. A key step in their proof was to show that if

$$\frac{S_n}{V_n} \text{ is tight} \tag{1.4}$$

then it is uniformly subgaussian in the sense that

$$\sup_n E \exp\left(t \frac{S_n}{V_n}\right) \le 2\exp(ct^2) \tag{1.5}$$

for all $t \in \mathbb{R}$ and some $c > 0$. This is clearly an important property of the self-normalized sum S_n/V_n which is not shared by scalar normalized sums, *i.e.* sums of the form $(S_n - a_n)b_n^{-1}$ for scalar sequences a_n and b_n. Thus Giné, Götze and Mason asked for precise conditions under which (1.4) holds. In a subsequent paper, Giné and Mason (1998) gave such a characterization for distributions which are in the Feller class. Here we will solve the problem in general.

To describe the result we need to introduce a little notation. For $r > 0$ set

$$G(r) = P(|X| > r), \quad K(r) = r^{-2}E(X^2; |X| \le r), \quad M(r) = r^{-1}E(X; |X| \le r), \tag{1.6}$$

and

$$Q(r) = G(r) + K(r). \tag{1.7}$$

Each of these functions is right continuous with left limits and tends to 0 as r approaches infinity. We can now give an analytic characterization of the two classes of random variables mentioned above. X is in the domain of attraction of the normal law and $EX = 0$ if and only if

$$\limsup_{r\to\infty} \frac{G(r) + |M(r)|}{K(r)} = 0, \tag{1.8}$$

while X is in the Feller class if and only if

$$\limsup_{r\to\infty} \frac{G(r)}{K(r)} < \infty. \tag{1.9}$$

The result of Giné and Mason is that if X is in the Feller class, then (1.4) holds if and only if

$$\limsup_{r\to\infty} \frac{|M(r)|}{K(r)} < \infty. \tag{1.10}$$

The main result of this paper is

Theorem 1 *The following are equivalent:*

$$t_n \textit{ is tight,} \tag{1.11}$$

$$\frac{S_n}{V_n} \textit{ is tight,} \tag{1.12}$$

$$\limsup_{r\to\infty} \frac{|M(r)|}{Q(r)} < \infty. \tag{1.13}$$

Examples of distributions satisfying (1.13) but not (1.9) and (1.10) are easily found, for example any symmetric distribution for which the tail function G is slowly varying.
In the course of the proof of Theorem 1 we also answer the question of when does there exist a centering sequence α_n for which

$$\frac{S_n - \alpha_n}{V_n} \text{ is tight.} \tag{1.14}$$

In the case of scalar normalization this reduces to centering at the median of S_n since for any scalar sequence b_n, if $(S_n - \alpha_n)b_n^{-1}$ is tight for some α_n, then it is tight with $\alpha_n = median(S_n)$. For self-normalization this is not the case. We illustrate this by giving an example for which (1.12) holds but (1.14) fails when $\alpha_n = median(S_n)$.
In concluding the introduction we would like to mention that there have been several other interesting lines of investigation into the Student t-statistic. These include large deviation results (Shao (1997)), law of the iterated logarithm results (Griffin and Kuelbs (1991), Giné and Mason (1998)) and Berry-Esseen bounds (Bentkus and Götze (1994)). In addition Chistyakov and Götze (2001) have recently confirmed a second conjecture of Logan, Mallows, Rice and Shepp 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. This last paper contains further references to the literature on self-normalized sums.

2 Preliminaries

We begin by showing that for tightness, and indeed for many asymptotic properties, the behavior of t_n and S_n/V_n are equivalent. In order that S_n/V_n always make sense, we define $S_n/V_n = 0$ if $V_n = 0$.

Proposition 1 *If $EX^2 < \infty$ and $EX \neq 0$ then*

$$\lim_{n\to\infty} \frac{t_n}{S_n/V_n} = \frac{(EX^2)^{1/2}}{(EX^2 - (EX)^2)^{1/2}} \quad a.s. \tag{2.1}$$

If $EX^2 = \infty$ or $EX = 0$ then

$$\lim_{n\to\infty} \frac{t_n}{S_n/V_n} = 1 \quad a.s. \tag{2.2}$$

Proof. If $EX^2 < \infty$ and $EX \neq 0$ then (2.1) follows immediately from (1.3) and the strong law. To prove (2.2) it suffices to show

$$\lim_{n\to\infty} \frac{(n^{-1}S_n)^2}{n^{-1}V_n^2} = 0 \quad a.s. \tag{2.3}$$

If $EX = 0$ this follows immediately from the strong law. Thus we are left to deal with the case $EX^2 = \infty$. Fix $L > 0$, then

$$\begin{aligned} |S_n| &\le \sum_{i=1}^n |X_i| I(|X_i| \le L) + \sum_{i=1}^n |X_i| I(|X_i| > L) \\ &\le nL + V_n \Big(\sum_{i=1}^n I(|X_i| > L)\Big)^{1/2}. \end{aligned} \tag{2.4}$$

Hence

$$\frac{n^{-1}|S_n|}{n^{-1/2}V_n} \le \frac{L}{n^{-1/2}V_n} + \left(\frac{\sum_{i=1}^n I(|X_i| > L)}{n}\right)^{1/2}. \tag{2.5}$$

Thus again by the strong law

$$\limsup_{n\to\infty} \frac{n^{-1}S_n}{n^{-1/2}V_n} \le (P(|X| > L))^{1/2}. \tag{2.6}$$

The result then follows by letting $L \to \infty$. □

The functions defined in (1.6) and (1.7) are defined for $r > 0$. It will be convenient to extend them to $r = 0$ by continuity. Thus set

$$G(0) = P(|X| > 0), K(0) = M(0) = 0. \tag{2.7}$$

We will be particularly interested in the function Q of (1.7) which is in fact continuous. This is most easily seen by observing that

$$Q(r) = r^{-2}E(X^2 \wedge r^2) = r^{-2}\int_0^r 2sG(s)\ ds. \tag{2.8}$$

Taking the right derivative in (2.8) shows that Q is constant on $[0, r_0]$ and strictly decreasing on $[r_0, \infty)$ where

$$r_0 = \inf\{r > 0 : G(r) < G(0)\}. \tag{2.9}$$

Thus for each fixed $\lambda > 0$, we can define a sequence $a_n(\lambda)$ for all $n > (\lambda Q(0))^{-1}$ by

$$Q(a_n(\lambda)) = \frac{1}{\lambda n}. \tag{2.10}$$

Observe that $a_n(\lambda)$ is increasing in both n and λ. For $n > (\lambda Q(0))^{-1}$ set

$$U_n(\lambda) = \sum_{i=1}^n X_i^2 \wedge a_n^2(\lambda). \tag{2.11}$$

Lemma 1 *Fix $\lambda > 0$. For any $\delta \in (0, \lambda^{-\frac{1}{2}})$ and $n > (\lambda Q(0))^{-1}$*

$$P(U_n(\lambda) > \delta^2 a_n^2(\lambda)) \ge \frac{(1-\lambda\delta^2)^2}{1+\lambda}.$$

Proof. First observe that for any $n > (\lambda Q(0))^{-1}$

$$EU_n(\lambda) = na_n^2(\lambda)Q(a_n(\lambda)) = \frac{a_n^2(\lambda)}{\lambda} \tag{2.12}$$

and

$$\begin{aligned} EU_n(\lambda)^2 &= nE(X^4 \wedge a_n^4(\lambda)) + 2\binom{n}{2}(E\big(X^2 \wedge a_n^2(\lambda)))^2 \\ &\le na_n^4(\lambda)Q(a_n(\lambda)) + n^2(a_n^2(\lambda)Q(a_n(\lambda)))^2 \\ &= a_n^4(\lambda)(\frac{1}{\lambda} + \frac{1}{\lambda^2}). \end{aligned} \tag{2.13}$$

Thus by a reverse Chebyshev inequality, see Durrett (1996) Exercise 3.8 on page 16, for any $\delta \in (0,\, \lambda^{-\frac{1}{2}})$

$$\begin{aligned} P\big(U_n(\lambda) > \delta^2 a_n^2(\lambda)\big) &= P(U_n(\lambda) > \lambda\delta^2 EU_n(\lambda)) \\ &\geq (1-\lambda\delta^2)^2\, \frac{(EU_n(\lambda))^2}{EU_n(\lambda)^2} \\ &\geq \frac{(1-\lambda\delta^2)^2}{1+\lambda} \end{aligned} \tag{2.14}$$

by (2.12) and (2.13). □

Corollary 1 *Fix* $\lambda > 0$. *For any* $\delta \in (0,\, \lambda^{-\frac{1}{2}})$ *and* $n > (\lambda Q(0))^{-1}$

$$P(V_n > \delta a_n(\lambda)) \geq \frac{(1-\lambda\delta^2)^2}{1+\lambda}. \tag{2.15}$$

Proof. This follows immediately from Lemma 1 since $V_n^2 \geq U_n(\lambda)$ for $n > (\lambda Q(0))^{-1}$. □

Let

$$X_n^* = \max_{1\leq i\leq n} |X_i|.$$

Lemma 2 *Fix* $\lambda > 0$, $L > 0$ *and* $n > (\lambda Q(0))^{-1}$, *then*

$$P(V_n > La_n(\lambda)) \leq \frac{1}{\lambda L^2} + 1 - (1 - \frac{1}{\lambda n})^n. \tag{2.16}$$

Proof. Since $V_n^2 = U_n(\lambda)$ on $\{X_n^* \leq a_n(\lambda)\}$, we have

$$\begin{aligned} P(V_n > La_n(\lambda)) &\leq P(V_n > La_n(\lambda), X_n^* \leq a_n(\lambda)) + P(X_n^* > a_n(\lambda)) \\ &\leq P(U_n(\lambda) > L^2 a_n^2(\lambda)) + P(X_n^* > a_n(\lambda)) \\ &\leq \frac{EU_n(\lambda)}{L^2 a_n^2(\lambda)} + 1 - (1 - G(a_n(\lambda)))^n \\ &\leq \frac{1}{\lambda L^2} + 1 - (1 - \frac{1}{\lambda n})^n. \end{aligned} \tag{2.17}$$

□

Now let

$$T_n(\lambda) = \sum_{i=1}^n X_i I(|X_i| \leq a_n(\lambda)), \quad R_n(\lambda) = \sum_{i=1}^n X_i I(|X_i| > a_n(\lambda))$$

$$J_n(\lambda) = \sum_{i=1}^n I(|X_i| > a_n(\lambda))$$

and set

$$\alpha_n(\lambda) = ET_n(\lambda) = na_n(\lambda)M(a_n(\lambda)). \tag{2.18}$$

Lemma 3 *Fix* $\lambda > 0$, $L > 0$ *and* $n > (\lambda Q(0))^{-1}$, *then*

$$P(|R_n(\lambda)| > LV_n) \leq \frac{1}{\lambda L^2}. \tag{2.19}$$

Proof. Observe that by the Cauchy-Schwartz inequality

$$|R_n(\lambda)| \le V_n\sqrt{J_n(\lambda)}.$$

Thus

$$\begin{aligned} P(|R_n(\lambda)| > LV_n) &\le P(J_n(\lambda) > L^2) \\ &\le \frac{nG(a_n(\lambda))}{L^2} \\ &\le \frac{1}{\lambda L^2}. \end{aligned} \tag{2.20}$$

□

Lemma 4 *Fix $L > 0$, $\lambda > 0$ and $\delta \in (0, \lambda^{-\frac{1}{2}})$. Then for any $n > (\lambda Q(0))^{-1}$*

$$P(|S_n - \alpha_n(\lambda)| > 2LV_n) \le \frac{1}{\lambda}\Big(\frac{1}{L^2\delta^2} + \frac{1}{L}\Big) + 1 - \frac{(1-\lambda\delta^2)^2}{1+\lambda}.$$

Proof. Since

$$S_n = T_n(\lambda) + R_n(\lambda)$$

we have

$$\begin{aligned} \{|S_n - \alpha_n(\lambda)| > 2LV_n\} &\subset \{|T_n(\lambda) - \alpha_n(\lambda)| > LV_n\} \cup \{|R_n(\lambda)| > LV_n\} \\ &\subset \{|T_n(\lambda) - \alpha_n(\lambda)| > L\delta a_n(\lambda)\} \\ &\qquad \cup \{V_n \le \delta a_n(\lambda)\} \cup \{|R_n(\lambda)| > LV_n\}. \end{aligned}$$

Hence by Chebyshev's inequality, (2.15) and (2.19)

$$\begin{aligned} P(|S_n - \alpha_n(\lambda)| > LV_n) &\le \frac{na_n^2(\lambda)K(a_n(\lambda))}{L^2\delta^2 a_n^2(\lambda)} + 1 - \frac{(1-\lambda\delta^2)^2}{1+\lambda} + \frac{1}{\lambda L^2} \\ &\le \frac{1}{\lambda}\Big(\frac{1}{L^2\delta^2} + \frac{1}{L^2}\Big) + 1 - \frac{(1-\lambda\delta^2)^2}{1+\lambda}. \end{aligned}$$

□

Corollary 2 *For any $\lambda > 0$*

$$\limsup_{L\to\infty}\limsup_{n\to\infty} P\big(|S_n - \alpha_n(\lambda)| > LV_n\big) \le \frac{\lambda}{1+\lambda}.$$

Proof. In Lemma 4, let $n \to \infty$, then $L \to \infty$ and finally $\delta \to 0$. □

3 Proofs

We first derive a necessary and sufficient condition for tightness of the centered self-normalized sum then specialize this to the case of centering at 0.

Theorem 2 *Fix a centering sequence α_n. Then the following are equivalent:*

$$\frac{S_n - \alpha_n}{V_n} \text{ is tight,} \tag{3.1}$$

$$\limsup_{n\to\infty} \left|\frac{\alpha_n - na_n(\lambda)M(a_n(\lambda))}{a_n(\lambda)}\right| < \infty \text{ for all } \lambda > 0, \tag{3.2}$$

$$\limsup_{n\to\infty} \left|\frac{\alpha_n - na_n(\lambda)M(a_n(\lambda))}{a_n(\lambda)}\right| < \infty \text{ for all sufficiently small } \lambda > 0. \tag{3.3}$$

Proof. First assume (3.1) holds. For any $n > (\lambda Q(0))^{-1}$

$$\begin{aligned} P(|T_n(\lambda) - &\alpha_n| > 2L^2 a_n(\lambda)) \\ &\le P(|T_n(\lambda) - \alpha_n| > 2L^2 a_n(\lambda), La_n(\lambda) \ge V_n) + P(V_n > La_n(\lambda)) \\ &\le P(|T_n(\lambda) - \alpha_n| > 2LV_n) + P(V_n > La_n(\lambda)) \\ &\le P(|S_n - \alpha_n| > LV_n) + P(|R_n(\lambda)| > LV_n) + P(V_n > La_n(\lambda)). \end{aligned} \tag{3.4}$$

Thus by (3.1), (2.16) and (2.19), for any $\lambda > 0$

$$\limsup_{L\to\infty} \limsup_{n\to\infty} P(|T_n(\lambda) - \alpha_n| > 2L^2 a_n(\lambda)) \le 1 - e^{-\frac{1}{\lambda}}. \tag{3.5}$$

On the other hand by Chebyshev's inequality

$$P(|T_n(\lambda) - \alpha_n(\lambda)| > L^2 a_n(\lambda)) \le \frac{na_n^2(\lambda)K(a_n(\lambda))}{L^4 a_n^2(\lambda)} \le \frac{1}{\lambda L^4}. \tag{3.6}$$

Thus for a fixed $\lambda > 0$, if $L^4 > \lambda^{-1}\exp(\lambda^{-1})$, then by (3.5) and (3.6), for all n sufficiently large

$$|\alpha_n - \alpha_n(\lambda)| \le 3L^2 a_n(\lambda).$$

Hence (3.2) holds.

That (3.2) implies (3.3) is trivial. Finally assume (3.3) holds. Fix $\lambda > 0$ sufficiently small that

$$c(\lambda) =: \sup_{n > (\lambda Q(0))^{-1}} \frac{|\alpha_n - na_n(\lambda)M(a_n(\lambda))|}{a_n(\lambda)} < \infty.$$

Observe that for any $L > 0$ and any $n > (\lambda Q(0))^{-1}$

$$\begin{aligned} \{|S_n - \alpha_n| > 2LV_n\} &\subset \{|S_n - \alpha_n(\lambda)| > 2LV_n - c(\lambda)a_n(\lambda)\} \\ &\subset \{|S_n - \alpha_n(\lambda)| > LV_n\} \cup \{LV_n \le c(\lambda)a_n(\lambda)\}. \end{aligned} \tag{3.7}$$

Now if L is large enough that $c(\lambda)L^{-1} < \lambda^{-\frac{1}{2}}$, then by Corollary 1

$$\limsup_{n\to\infty} P(V_n \le c(\lambda)L^{-1}a_n(\lambda)) \le 1 - \frac{(1 - \lambda(c(\lambda)L^{-1})^2)^2}{1+\lambda}. \tag{3.8}$$

Hence by Corollary 2, (3.7) and (3.8), for all λ sufficiently small

$$\limsup_{L\to\infty} \limsup_{n\to\infty} P(|S_n - \alpha_n| > 2LV_n) \le \frac{\lambda}{1+\lambda} + 1 - \frac{1}{1+\lambda}.$$

Thus (3.1) follows by letting $\lambda \downarrow 0$. □

Theorem 3 *The following are equivalent:*

$$\frac{S_n}{V_n} \textit{ is tight,} \tag{3.9}$$

$$\limsup_{r\to\infty} \frac{|M(r)|}{Q(r)} < \infty. \tag{3.10}$$

Proof. Assume (3.10). Then with $\alpha_n = 0$ we have for $n > (\lambda Q(0))^{-1}$

$$|\frac{\alpha_n - na_n(\lambda)M(a_n(\lambda))}{a_n(\lambda)}| = \frac{1}{\lambda}\frac{|M(a_n(\lambda))|}{Q(a_n(\lambda))},$$

so (3.9) holds by Theorem 2.
Conversely assume (3.10) fails, so

$$\frac{|M(r_k)|}{Q(r_k)} \to \infty \tag{3.11}$$

for some $r_k \to \infty$. Set

$$n_k = \max\{n : nQ(r_k) \le 1\}. \tag{3.12}$$

Then for $n_k > (Q(0))^{-1}$ we have that

$$r_k = a_{n_k}(\lambda_k) \tag{3.13}$$

where $1 \le \lambda_k \le 1 + n_k^{-1}$. Now for such k

$$\begin{aligned} |n_k a_{n_k}(\lambda_k)M(a_{n_k}(\lambda_k)) - n_k a_{n_k}(1)M(a_{n_k}(1))| &\le a_{n_k}(\lambda_k)n_k G(a_{n_k}(1)) \\ &\le a_{n_k}(\lambda_k), \end{aligned}$$

while by (3.11)

$$n_k|M(a_{n_k}(\lambda_k))| = \frac{|M(a_{n_k}(\lambda_k))|}{\lambda_k Q(a_{n_k}(\lambda_k))} = \frac{|M(r_k)|}{\lambda_k Q(r_k)} \to \infty.$$

Consequently

$$\frac{n_k a_{n_k}(1)|M(a_{n_k}(1))|}{a_{n_k}(\lambda_k)} \to \infty.$$

Since $a_{n_k}(1) \le a_{n_k}(\lambda_k)$ it then follows that $n_k|M(a_{n_k}(1))| \to \infty$. Thus we conclude that (3.2) fails with $\alpha_n = 0$ and $\lambda = 1$. Hence by Theorem 2, (3.9) fails. □

Theorem 1 follows immediately from Proposition 1 and Theorem 3. We conclude by giving an example showing that it is possible for (1.12) to hold, but for (1.14) to fail with $\alpha_n =$ *median*(S_n).

Example 1 *Let $X > 0$ have distribution given by*

$$G(r) = \frac{1}{\ln r}, \quad r \ge e. \tag{3.14}$$

Since G is slowly varying, it follows from Darling (1952) that

$$S_n/X_n^* \xrightarrow{p} 1 \text{ and } V_n/X_n^* \xrightarrow{p} 1. \tag{3.15}$$

Consequently $S_n/V_n^* \xrightarrow{p} 1$ *and in particular* (1.12) *holds. Set*

$$b_n(\lambda) = e^{\lambda n}. \tag{3.16}$$

Observe that for any $\lambda_1 > \lambda_2 > 0$

$$\frac{b_n(\lambda_1)}{b_n(\lambda_2)} \to \infty. \tag{3.17}$$

Now fix $\lambda_1 < (\ln 2)^{-1}$. *Then*

$$P(S_n > b_n(\lambda_1)) \geq P(X_n^* > b_n(\lambda_1)) = 1 - (1 - \frac{1}{\lambda_1 n})^n \to 1 - e^{-\lambda_1^{-1}} > \frac{1}{2}. \tag{3.18}$$

Hence the median m_n *of* S_n *satisfies*

$$\limsup_{n\to\infty} \frac{m_n}{b_n(\lambda_1)} \geq 1. \tag{3.19}$$

Now fix $\lambda_2 \in (0, \lambda_1)$. *Then for any* $\lambda_3 \in (0, \lambda_2)$.

$$\begin{aligned} P(V_n \leq b_n(\lambda_2)) &\geq P(V_n \leq b_n(\lambda_2), \frac{V_n}{X_n^*} \leq \frac{b_n(\lambda_2)}{b_n(\lambda_3)}) \\ &\geq P(X_n^* \leq b_n(\lambda_3)) - P(\frac{V_n}{X_n^*} > \frac{b_n(\lambda_2)}{b_n(\lambda_3)}) \\ &\to e^{-\lambda_3^{-1}}. \end{aligned} \tag{3.20}$$

by (3.15) *and* (3.17). *Thus by* (3.17), (3.19) *and* (3.20), m_n/V_n *is not tight. Since, as we have already observed,* (1.12) *holds, it then follows that* (1.14) *must fail when* $\alpha_n = m_n$.

References

Bentkus, V., and Götze, F., (1994). The Berry-Esseen bound for student's statistic. *Ann. Probab.* 24,491–503.

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

Darling, D.A., (1952). The influence of the maximum term in the addition of independent random variables. *Trans. Amer. Math. Soc.* 73, 95–107.

Durrett, R., (1996). Probability:Theory and Examples , Second Edition , (Duxbury Press).

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

Giné, E. and Mason, D.M., (1998). On the LIL for self-normalized sums of IID random variables. *J. Theor. Prob.* 11, 351–370.

Griffin, P.S., and Kuelbs, J., (1991). Some extensions of the LIL via self-normalizations. *Ann. Probab.* 19, 380–395.

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.

Shao, Qi-Man., (1997). Self-normalized large deviations. *Ann. Probab.* 25, 285–328.