ESAIM: PS 17 (2013) 1-12
DOI: 10.1051/ps/2011112

ESAIM: Probability and Statistics
www.esaim-ps.org

LARGE DEVIATIONS FOR QUASI-ARITHMETICALLY SELF-NORMALIZED RANDOM VARIABLES

Jean-Marie Aubry[1] and Marguerite Zani[1]

Abstract. We introduce a family of convex (concave) functions called sup (inf) of powers, which are used as generator functions for a special type of quasi-arithmetic means. Using these means, we generalize the large deviation result on self-normalized statistics that was obtained in the homogeneous case by [Q.-M. Shao, Self-normalized large deviations. *Ann. Probab.* **25** (1997) 285–328]. Furthermore, in the homogenous case, we derive the Bahadur exact slope for tests using self-normalized statistics.

Mathematics Subject Classification. 60F10, 62F05.

Received February 4, 2010. Revised August 17, 2010.

1. Introduction

The use of self-normalization in statistics dates back to Student's t-distribution

$$T_n := \sqrt{n}\frac{\overline{X_n} - m}{S_n}$$

where $\overline{X_n} := \frac{1}{n}\sum_{i=1}^n X_i$ is the empirical mean and $S_n^2 := \frac{1}{n-1}\sum_{i=1}^n \left(X_i - \overline{X_n}\right)^2$ is the empirical unbiaised variance of a Gaussian sample $X_1, \ldots, X_n$ of mean m and unknown variance. This statistic T_n is widely used to test $m = m_0$ against alternative and it is well known that T_n has normal limit distribution when $n \to \infty$ (even when the variables X_i are i.i.d. non Gaussian). Shao [14] proved the associated large deviations results for i.i.d. random variables X_i that are self-normalized by homogenous means, *i.e.* an asymptotic estimate for

$$\mathbb{P}\left(\frac{\overline{X_n}}{M_{p,n}} \geq x\right)^{\frac{1}{n}} \tag{1.1}$$

with $M_{p,n} := \left(\frac{1}{n}\sum_{i=1}^n |X_i|^p\right)^{\frac{1}{p}}$, $p > 1$ and minimal hypotheses on the common law of X_i. There have been several further papers on this subject by Dembo and Shao [6,7], Lai and Shao [10], Shao [15] and a book by De La Pena *et al.* [5]. For related results see also Tchirina [16].

The difficulty here is, of course, that the self-normalized random variables $X_i/M_{p,n}$ are far from independent and the result is indeed quite different from the classical Cramér-Chernoff large deviations [2]; in particular, no

Keywords and phrases. Large deviations, self-normalised statistics, Bahadur exact slope.

[1] Laboratoire d'Analyse et Mathématiques Appliquées (CNRS UMR 8050), Université Paris-Est Créteil, 61 av. du Général de Gaulle, 94010 Créteil Cedex, France; jmaubry@math.cnrs.fr; zani@u-pec.fr

Article published by EDP Sciences

© EDP Sciences, SMAI 2012

https://doi.org/10.1051/ps/2011112 Published online by Cambridge University Press

exponential moment hypothesis on X_i (actually, no moment hypothesis at all) is necessary. It is worth noting that Shao's applications, notably to Student's t-distribution, make only use of the case $p = 2$, *i.e.* the *quadratic* mean self-normalization; besides, the results for other values of $p > 1$ (the so-called *Hölder* or *power* means) present no extra arduousness.

There has however been a renewed interest in recent litterature for *quasi-arithmetic* means, where the power function $x \mapsto x^p$ is replaced by a more general strictly monotonic and continuous *generator function* $\phi : [0, \infty) \to [0, \infty)$. More precisely, by a famous result of Kolmogorov [9] and Nagumo [12], any reasonable notion of mean can be written as

$$M_{\phi,n} := \phi^{-1}\left(\frac{1}{n}\sum_{i=1}^{n}\phi(X_i)\right)$$

for such a function ϕ. Modern applications of quasi-arithmetic means in a probabilistic framework range from Ordered Weighted Averaging operators, introduced in decision theory by Yager [17], to statistics, because of their links with Archimedean copulas [3,4] and covariance functions [13].

Another potential application is the following. Suppose that the i.i.d. variables X_i follow an unknown law $\mu(\theta, c)$, where c is the parameter of interest. In cases where θ is simply a scale parameter (meaning that there exists a function ν such that the law of $X_i/\nu(\theta)$ does not depend on θ), self-normalization by any power mean yields a scale-free statistic that can be used to test for c. Large deviations type results can then be translated in terms of Bahadur asymptotic efficiency, see Section 4. In more general cases, one can try and find ϕ such that $\frac{E(X)}{\phi^{-1}(E(\phi(X)))} = f(c)$ and does not depend on θ. This seems to be a difficult problem, but *ad hoc* examples (starting from ϕ, thus admittedly not very useful) can be constructed to show at least that the class of laws $\mu(\theta, c)$ for which it can be solved is not empty. Suppose that for $\theta > 0$ and $\sqrt{3}/2 < c < 1$, $\mu(\theta, c)$ is the uniform law on $[\theta, \theta + \lambda(\theta, c)]$ where

$$\lambda(\theta, c) = \frac{6\theta - 3c^2(1 + 2\theta) + c\left(3 + \sqrt{3(1-c)\left(3 + 4\theta^2 + c\left(-3 + 4\theta + 4\theta^2\right)\right)}\right)}{4c^2 - 3}.$$

Then for $\phi(x) = x + x^2$ one has indeed $\frac{E(X)}{\phi^{-1}(E(\phi(X)))} = c$ for all θ and the statistic

$$T_n := \frac{\overline{X_n}}{M_{\phi,n}} \tag{1.2}$$

can be used to test for c.

So the natural question arises, whether Shao's asymptotic results on (1.1) can be extended using a more general quasi-arithmetic mean $M_{\phi,n}$. In this paper, we show large deviations for statistics of type (1.2). The main drawback and technical difficulty is that now $M_{\phi,n}$ is no longer homogenous (indeed the only homogeneous quasi-arithmetic means are the power means and the geometric mean, see [8] (p. 68)). For this reason, we need to restrict ourselves to a subclass of quasi-arithmetic means where the generator function ϕ enjoys some extra properties, studied in Section 2. The main theorem on self-normalized large deviations is then stated and proved in Section 3. In Section 4, we show how large deviations results can be used to determine Bahadur exact slope for self-normalized statistics.

2. Infimum and supremum of powers

We shall apply Shao's method in [14] to a more general class of self-normalized large deviations, where the power function $x \mapsto x^p$ used as a generator function for the normalizing mean is replaced by a more general convex function ϕ or, equivalently, $x \mapsto x^{\frac{1}{p}}$ is replaced by the concave function $\psi := \phi^{-1}$. The first step of our approach consists in writing ψ as an infimum of concave power functions, which can be done only in certain cases. Alternatively, the condition can be expressed in terms of ϕ written as a supremum of convex power functions.

https://doi.org/10.1051/ps/2011112 Published online by Cambridge University Press

2.1. **Characterization of admissible generator functions**

In the following, $\partial\omega(x)$ denotes the superdifferential (resp. subdifferential) at x of the concave (resp. convex) function ω; $a \vee b$ denotes the maximum of a and b.

Proposition 2.1. *For $\psi : [0,\infty) \to [0,\infty)$, the following are equivalent.*

(i) *There exists $\mu : [1,\infty) \to (0,\infty]$ not identically $= \infty$ such that for all $x > 0$,*

$$\psi(x) = \inf_{p\geq 1} \mu(p) x^{\frac{1}{p}}. \tag{2.1}$$

(ii) *The function $\omega : x \mapsto \log(\psi(\exp(x)))$ is proper closed concave on $\mathbb{R}$ and $\partial\omega(x) \subset [0,1]$ for all $x \in \mathbb{R}$.*

Proof. Let us start by assuming (i). Then, writing $\alpha := \frac{1}{p}$, $y := \log(x)$ and $\vartheta(\alpha) := -\log(\mu(\frac{1}{\alpha}))$, we turn (2.1) into

$$\omega(y) = \inf_{0<\alpha\leq 1} \alpha y - \vartheta(\alpha). \tag{2.2}$$

So ω is the concave conjugate of $\tilde{\vartheta}$, the concave hull of ϑ restricted to $(0,1]$; the closure of $\tilde{\vartheta}$ has support $\subset [0,1]$ hence the conclusion.

Conversely, suppose that (ii) holds. Let ϑ be the concave conjugate of ω, it is a proper closed concave function with support $\subset [0,1]$ because of the hypothesis on the superdifferential of ω. Inverting the Fenchel-Legendre transformation, we can write $\omega(y) = \inf_{0\leq\alpha\leq 1} \alpha y - \vartheta(\alpha)$, which by closedness is equivalent to (2.2), hence (i) with $\mu(p) := \exp(-\vartheta(\frac{1}{p}))$. □

Remarks. (i) clearly implies that ψ itself is nondecreasing, concave and that $\psi(0) = 0$ but this is not enough (take $\psi(x) = x + \sqrt{x}$ for instance). The example $\mu(p) = 1$ for all $p \geq 1$ shows however that ψ is not necessarily (strictly) increasing. Possible infinite values for μ mean that the infimum on p is actually taken on the support $S_\mu := [1,\infty)\backslash\mu^{-1}(\{\infty\})$.

When a function ψ satisfies the hypothesis of Proposition 2.1, we shall say that it is an *inf(imum) of powers* with weight μ. Note that we can (and will) always assume, without loss of generality, that $\mu(p) = \exp(-\vartheta(\frac{1}{p}))$ for a proper closed concave function ϑ; in particular μ is continuous on its support. We shall say in that case that μ is a *regular* weight and remark that $\inf_{p\geq 1} \mu(p) > 0$. If furthermore ψ is differentiable, then $\mu(p)$ can be parametrized by the contact point $s > 0$, $p(s) = \frac{\psi(s)}{s\psi'(s)}$ and $\mu(p(s)) = \psi(s) s^{-\frac{1}{p(s)}}$ (see examples below).

Proposition 2.2. *For $\phi : [0,\infty) \to [0,\infty]$, the following are equivalent.*

(i) *There exists $\nu : [1,\infty) \to [0,\infty)$ not identically $= 0$ such that for all $x > 0$,*

$$\phi(x) = \sup_{p\geq 1} \nu(p) x^p. \tag{2.3}$$

(ii) *The function $\omega : x \mapsto \log(\phi(\exp(x)))$ is proper closed convex on $\mathbb{R}$ and $\partial\omega(x) \subset [1,\infty)$ for all $x \in \mathbb{R}$ where $\phi(x) < \infty$.*

The proof is very similar to that of Proposition 2.1 and is omitted (note that ϕ can take infinite values).

Examples. Some inf of powers with parametrization.

$\psi(x)$	$\omega(x)$	$p(s)$	$\mu(p(s))$
$\log(1+x)$	$\log\log(1+\mathrm{e}^x)$	$\frac{(1+s)\log(1+s)}{s}$	$s^{\frac{s}{(1+s)\log(1+s)}} \log(1+s)$
$\sqrt{1+x}-1$	$\log(\sqrt{1+\mathrm{e}^x}-1)$	$\frac{2\left(1+s-\sqrt{1+s}\right)}{s}$	$s^{\frac{s}{2\left(-1-s+\sqrt{1+s}\right)}} \left(\sqrt{1+s}-1\right)$
$1-\mathrm{e}^{-x}$	$\log\left(1-\mathrm{e}^{-\mathrm{e}^x}\right)$	$\frac{\mathrm{e}^s-1}{s}$	$(1-\mathrm{e}^{-s}) s^{\frac{s}{\mathrm{e}^s-1}}$

https://doi.org/10.1051/ps/2011112 Published online by Cambridge University Press

2.2. **Bijectivity**

Let us give the necessary and sufficient conditions under which an infimum (resp. supremum) of powers is bijective, in which case its converse is a supremum (resp. infimum) of powers.

Proposition 2.3. *Let ψ be an inf of powers with regular weight μ. Then the following are equivalent.*

(i) $\lim_{p\to\infty} \mu(p) = \infty$;
(ii) *ψ is unbounded;*
(iii) *ψ is a bijection:* $[0,\infty) \to [0,\infty)$ *and* $\psi^{-1}(y) = \sup_{p\geq 1} \mu(p)^{-p} y^p$.

Proof. Suppose (i), fix any $C < \infty$ and pick p_0 such that $\mu(p) \geq C$ when $p \geq p_0$. Then choose $x := 1 \vee \left(\frac{C}{\inf_{p>1}\mu(p)}\right)^{p_0}$. We thus have $\inf_{p\geq p_0} \mu(p) x^{\frac{1}{p}} \geq C$ and $\inf_{1<p\leq p_0} \mu(p) x^{\frac{1}{p}} \geq (\inf_{p>1}\mu(p)) x^{\frac{1}{p_0}} \geq C$ so $\psi(x) \geq C$. Conversely, since $\mu(p) = \exp(-\vartheta(\frac{1}{p}))$ with ϑ concave, if μ does not have infinite limit at ∞ then it has to be bounded. In that case ψ is also bounded (by the same constant).

Being concave and taking value 0 at 0, ψ unbounded implies that it is strictly increasing and surjective on $[0,\infty)$, hence it is a bijection. Furthermore we have $x \geq \psi^{-1}(y) \iff y \leq \psi(x) \iff \forall p > 1, y \leq \mu(p) x^{\frac{1}{p}} \iff \forall p > 1, x \geq \mu(p)^{-p} y^p \iff x \geq \sup_{p>1} \mu(p)^{-p} y^p$. The converse is trivial. □

Proposition 2.4. *Let ϕ be a sup of powers with regular weight ν. Then the following are equivalent.*

(i) $\lim_{p\to\infty} \nu(p)^{\frac{1}{p}} = 0$;
(ii) *ϕ is finitely-valued;*
(iii) *ϕ is a bijection:* $[0,\infty) \to [0,\infty)$ *and* $\phi^{-1}(x) = \inf_{p\geq 1} \nu(p)^{-\frac{1}{p}} x^{\frac{1}{p}}$.

The proof is very similar to that of Proposition 2.3 and is omitted.

2.3. **Operations**

We shall now see that the class of functions satisfying the inf (resp. sup) of powers property is "large", as it is under minor conditions stable by powers, composition, and products. We give the proofs only in the inf case.

Proposition 2.5. *If ψ is an inf (resp. sup) of powers, then so is $\psi^\alpha : x \mapsto \psi(x)^\alpha$ for $0 < \alpha \leq 1$ (resp. for $\alpha \geq 1$).*

Proof. Take $\mu^\alpha(p) := \mu(\alpha p)^\alpha$ if $p > \frac{1}{\alpha}$, $\mu^\alpha(p) := \infty$ else. □

Proposition 2.6. *If ψ_1 and ψ_2 are inf (resp. sup) of powers, then so is $\psi_1 \circ \psi_2$.*

Proof. Using Proposition 2.5,

$$\begin{aligned}
\psi_1(\psi_2(x)) &= \inf_{p_1>1} \mu_1(p_1)\psi_2(x)^{\frac{1}{p_1}} \\
&= \inf_{p_2>p_1>1} \mu_1(p_1)\mu_2\left(\frac{p_2}{p_1}\right)^{\frac{1}{p_1}} x^{\frac{1}{p_2}} \\
&= \inf_{p_2>1} \underbrace{\inf_{p_1\in(1,p_2)} \mu(p_1)\mu_2\left(\frac{p_2}{p_1}\right)^{\frac{1}{p_1}}}_{=:\mu(p_2)} x^{\frac{1}{p_2}}.
\end{aligned}$$

□

Proposition 2.7. *If ψ_1 and ψ_2 are inf of powers with $p_1 \leq \frac{1}{\alpha_1} \Rightarrow \mu_1(p_1) = \infty$ and $p_2 \leq \frac{1}{\alpha_2} \Rightarrow \mu_2(p_2) = \infty$ for some $\alpha_1 + \alpha_2 \leq 1$, then $\psi_1\psi_2$ is also an inf of powers. If ϕ_1 and ϕ_2 are sup of powers (no extra condition), then so is $\phi_1\phi_2$.*

https://doi.org/10.1051/ps/2011112 Published online by Cambridge University Press

Proof. We can write

$$\psi_1(x)\psi_2(x) = \inf_{p_1>\frac{1}{\alpha_1}} \inf_{p_2>\frac{1}{\alpha_2}} \mu_1(p_1)\mu_2(p_2) x^{\frac{1}{p_1}+\frac{1}{p_2}}$$
$$= \inf_{p>1} \underbrace{\inf_{\frac{1}{p_1}+\frac{1}{p_2}=\frac{1}{p}} \mu_1(p_1)\mu_2(p_2)}_{=:\mu(p)} x^{\frac{1}{p}}. \qquad \square$$

3. Self-normalized large deviations

Let ψ be a bijective inf of powers with weight μ (*cf.* Prop. 2.3) and let $\phi := \psi^{-1}$ extended to $\mathbb{R}$ by even symmetry ($\phi(x) := \phi(|x|)$). Let $(X_i)_{i\geq 1}$ be a sequence of independent random variables having the same law as an arbitrary $X \not\equiv 0$. Their empirical mean

$$\overline{X_n} := \frac{1}{n}\sum_{i=1}^n X_i$$

will be normalized by the quasi-arithmetic mean

$$M_{\phi,n} := \psi\left(\frac{1}{n}\sum_{i=1}^n \phi(X_i)\right).$$

Theorem 3.1. *Suppose that there exist $1 < \pi_0 < \pi_1 < \infty$ such that $\mu(p) < \infty$ only if $p \in S_\mu := [\pi_0, \pi_1]$. Then for any $x > \begin{cases} \frac{0\vee\mathbb{E}X}{\psi(\mathbb{E}\phi(X))} & \text{if } \mathbb{E}|X| < \infty \\ 0 & \text{if } \mathbb{E}|X| = \infty, \end{cases}$*

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{\overline{X_n}}{M_{\phi,n}} \geq x\right)^{\frac{1}{n}} = \sup_{c>0,p\in S_\mu} \inf_{t\geq 0} \mathbb{E}\exp\left(t\left(cX - x\frac{\mu(p)}{p}\left(\phi(X) + (p-1)c^{\frac{p}{p-1}}\right)\right)\right).$$

Proof. For short we shall always write $q := \frac{p}{p-1}$, whenever $p > 1$ is considered. We repeatedly use the fact that $x \mapsto \frac{x^p}{p}$ and $y \mapsto \frac{y^q}{q}$ are convex conjugate functions and that for any $x, y \geq 0$ we have $xy = \inf_{b>0} \frac{(b^{-1}x)^p}{p} + \frac{(by)^q}{q}$.

Since ψ is an inf of powers, we obtain for $x \geq 0$

$$\mathbb{P}\left(\frac{\overline{X_n}}{M_{\phi,n}} \geq x\right) = \mathbb{P}\left(\sum_{i=1}^n X_i \geq \inf_{p\in S_\mu} x\mu(p) n^{\frac{1}{q}} \left(\sum_{i=1}^n \phi(X_i)\right)^{\frac{1}{p}}\right)$$
$$= \mathbb{P}\left(\sum_{i=1}^n X_i \geq \inf_{b>0,p\in S_\mu} x\mu(p)\left(\frac{b^{-p}}{p}\sum_{i=1}^n \phi(X_i) + n\frac{b^q}{q}\right)\right)$$
$$= \mathbb{P}\left(\sup_{b>0,p\in S_\mu} \sum_{i=1}^n \left(X_i - x\mu(p)\left(\frac{b^{-p}}{p}\phi(X_i) + \frac{b^q}{q}\right)\right) \geq 0\right)$$
$$= \mathbb{P}\left(\sup_{c>0,p\in S_\mu} \sum_{i=1}^n \left(cX_i - x\frac{\mu(p)}{p}(\phi(X_i) + (p-1)c^q)\right) \geq 0\right). \tag{3.1}$$

Then we can finish along the general lines of [14], with a significant complication due to the non-homogeneity of ϕ. We prove the lower bound in Section 3.1 and the upper bound in Section 3.2. $\square$

Remark. The hypothesis on the support of μ excludes some interesting examples. It could be replaced by a decay hypothesis on the tail distribution of X; this aspect will be studied in a subsequent paper.

https://doi.org/10.1051/ps/2011112 Published online by Cambridge University Press

3.1. **Lower bound**

Suppose first $\mathbb{E}|X| < \infty$. Since we look for one-sided large deviations in (3.1) we need to have

$$\mathbb{E}\left(cX - x\frac{\mu(p)}{p}(\phi(X) + (p-1)c^q)\right) < 0. \tag{3.2}$$

Observe that, by Young's inequality, with $\beta(p) := \frac{\mu(p)^{1-p}}{p}\left(\frac{0\vee\mathbb{E}X}{x}\right)^p$ we have for all values of $c > 0$ and $p \in S_\mu$

$$c\frac{0\vee\mathbb{E}X}{x} - \mu(p)\frac{c^q}{q} - \beta(p) \le 0 \tag{3.3}$$

while the hypothesis on x combined to (2.1) implies $x > \frac{0\vee\mathbb{E}X}{\mu(p)(\mathbb{E}\phi(X))^{\frac{1}{p}}}$, hence

$$\beta(p) - \frac{\mu(p)}{p}\mathbb{E}\phi(X) < 0. \tag{3.4}$$

Adding up (3.3) and (3.4) yields precisely (3.2) and we obtain

$$\liminf_{n\to\infty}\ \mathbb{P}\left(\frac{\overline{X_n}}{M_{\phi,n}} \ge x\right)^{\frac{1}{n}} \ge \inf_{t\ge 0}\ \mathbb{E}\exp\left(t\left(cX - x\frac{\mu(p)}{p}(\phi(X) + (p-1)c^q)\right)\right). \tag{3.5}$$

If $\mathbb{E}|X| = \infty$ and $x > 0$ then the right-hand side of (3.5) is zero, so the inequality is trivial.

Remark. So far we used no assumption on ϕ, except that it takes non-negative values on the range of X. When $\phi = \psi^{-1}$, this function is convex and Jensen's inequality implies that the lower bound on x is ≤ 1.

3.2. **Upper bound**

Now we use the fact that $\phi := \psi^{-1}$. Let $K > 3$ be fixed and let $[\pi_0', \pi_1'] := S_\mu' := \left\{p \in S_\mu : \frac{\mu(p)}{p} \le C\right\}$, where C is determined by the following (proof in Appendix 4.2):

Lemma 3.1. *There exists $C < \infty$ such that* $\inf_{t\ge 0}\ \mathbb{E}\exp\left(t(KX - xC\phi(X))\right) \le \frac{1}{K}$.

We deduce from (3.1) that $\mathbb{P}\left(\frac{\overline{X_n}}{M_{\phi,n}} \ge x\right) \le I_1 + I_2 + I_3$, with

$$I_1 := \mathbb{P}\left(\sup_{c>K, p\in S_\mu}\sum_{i=1}^n\left(cX_i - x\frac{\mu(p)}{p}(\phi(X_i) + (p-1)c^q)\right) \ge 0\right)$$
$$I_2 := \mathbb{P}\left(\sup_{0<c\le K, p\in S_\mu'}\sum_{i=1}^n\left(cX_i - x\frac{\mu(p)}{p}(\phi(X_i) + (p-1)c^q)\right) \ge 0\right)$$
$$I_3 := \mathbb{P}\left(\sup_{0<c\le K, p\in S_\mu\setminus S_\mu'}\sum_{i=1}^n\left(cX_i - x\frac{\mu(p)}{p}(\phi(X_i) + (p-1)c^q)\right) \ge 0\right).$$

https://doi.org/10.1051/ps/2011112 Published online by Cambridge University Press

We start with I_1. Let $K'(p) := \frac{xK^{q-1}\mu(p)}{2q}$: the restriction $c > K$ implies that $K'(p) < x\frac{\mu(p)}{2p}(p-1)c^{q-1}$, so

$$\begin{aligned}
&\sup_{c>K, p\in S_\mu} \sum_{i=1}^n \left(cX_i - x\frac{\mu(p)}{p}(\phi(X_i) + (p-1)c^q) \right) \geq 0 \\
\Rightarrow &\sup_{c>K, p\in S_\mu} \sum_{i=1}^n \left(|X_i|\mathbf{1}_{|X_i|>K'(p)} + K'(p) - x\frac{\mu(p)}{p}\left(\frac{\phi(X_i)}{c} + (p-1)c^{q-1} \right) \right) \geq 0 \\
\Rightarrow &\sup_{p\in S_\mu} \sum_{i=1}^n |X_i|\mathbf{1}_{|X_i|>K'(p)} - x\frac{\mu(p)}{p} \inf_{c>K} \sum_{i=1}^n \left(\frac{\phi(X_i)}{c} + \frac{1}{2}(p-1)c^{q-1} \right) \geq 0 \\
\Rightarrow &\sup_{p\in S_\mu} \sum_{i=1}^n |X_i|\mathbf{1}_{|X_i|>K'(p)} - x\mu(p) \inf_{b>K^{\frac{1}{p}}} \left(\frac{1}{p}\left(\frac{(\sum_{i=1}^n \phi(X_i))^{\frac{1}{p}}}{b} \right)^p + \frac{1}{q}\left(\left(\frac{n}{2}\right)^{\frac{1}{q}} b \right)^q \right) \geq 0
\end{aligned}$$

by Young's inequality,

$$\Rightarrow \sup_{p\in S_\mu} \sum_{i=1}^n |X_i|\mathbf{1}_{|X_i|>K'(p)} - x\mu(p)\left(\sum_{i=1}^n \phi(X_i) \right)^{\frac{1}{p}} \left(\frac{n}{2}\right)^{\frac{1}{q}} \geq 0$$

by Proposition 2.3,

$$\Rightarrow \sup_{p\in S_\mu} \sum_{i=1}^n |X_i|\mathbf{1}_{|X_i|>K'(p)} - x\left(\sum_{i=1}^n |X_i|^p \right)^{\frac{1}{p}} \left(\frac{n}{2}\right)^{\frac{1}{q}} \geq 0$$

and using Hölder's inequality,

$$\Rightarrow \sup_{p\in S_\mu} \sum_{i=1}^n \mathbf{1}_{|X_i|>K'(p)} - \frac{n}{2}x^q \geq 0.$$

Finally we obtain that

$$\begin{aligned}
I_1 &\leq \mathbb{P}\left(\sup_{p\in S_\mu} \sum_{i=1}^n \left(\mathbf{1}_{|X_i|>K'(p)} - \frac{x^q}{2} \right) \geq 0 \right) \\
&\leq \mathbb{P}\left(\sum_{i=1}^n \left(\mathbf{1}_{|X_i|>\inf_{p\in S_\mu} K'(p)} - \tau \right) \geq 0 \right)
\end{aligned}$$

with $\tau := \min(\frac{x^{\frac{\pi_0}{\pi_0-1}}}{2}, 1) > 0$. Applying the Chernoff bound on binomial random variables, we deduce that

$$\limsup_{n\to\infty} I_1^{\frac{1}{n}} \leq \left(\frac{e}{\tau}\mathbb{P}\left(|X| > \inf_{p\in S_\mu} K'(p) \right) \right)^\tau.$$

Since $\inf_{p\in S_\mu} \mu(p) > 0$, when $K \to \infty$ the quantity $\inf_{p\in S_\mu} K'(p) \to \infty$ and therefore $\limsup_{n\to\infty} I_1^{\frac{1}{n}} \to 0$.

https://doi.org/10.1051/ps/2011112 Published online by Cambridge University Press

Let us now bound I_2. We discretize $c \in [0,K]$ and $p \in S'_\mu$ with a step Δ to be chosen later (Lem. 3.2). Let T_l denote the interval $[{\pi_0}' + l\Delta, {\pi_0}' + (l+1)\Delta]$ and let p_l realize $\inf_{p\in T_l} \frac{\mu(p)}{p}$.

$$I_2 \leq \sum_{j=1}^{\frac{K}{\Delta}} \sum_{l=0}^{\frac{{\pi_1}'-{\pi_0}'}{\Delta}-1} \mathbb{P}\left(\sum_{i=1}^{n}\left(j\Delta X_i - x\frac{\mu(p_l)}{p_l}\left(\Phi(X_i) + \inf_{p\in T_l}(p-1)((j-1)\Delta)^q\right)\right) \geq 0\right)$$
$$\leq \sum_{j=1}^{\frac{K}{\Delta}} \sum_{l=0}^{\frac{{\pi_1}'-{\pi_0}'}{\Delta}-1} \inf_{t\geq 0}\left(\mathbb{E}\exp\left(t\left(j\Delta X - x\frac{\mu(p_l)}{p_l}\left(\Phi(X) + \inf_{p\in T_l}(p-1)((j-1)\Delta)^q\right)\right)\right)\right)^n.$$

By our choice of C in Lemma 3.1, for any $0 \leq l < \frac{{\pi_1}'-{\pi_0}'}{\Delta}$ we have $\frac{\mu(p_l)}{p_l} \leq C$. Let $Y_{j,l}$ be standard Gaussian independent random variables, $\delta > 0$ and

$$\xi_{j,l} := \delta Y_{j,l} + j\Delta X - x\frac{\mu(p_l)}{p_l}\left(\Phi(X) + (p_l-1)(j\Delta)^{\frac{p_l}{p_l-1}}\right).$$

Lemma 3.2. *There exists $\Delta > 0$ such that for all j, l, the infimum over $t \geq 0$ of $\exp(t\xi_{j,l})$ is reached for $t_{j,l} \in \left[0, \frac{1}{\Delta K^{\frac{\pi_0}{\pi_0-1}+1}}\right]$.*

Let us now define $D(j,l) := (p_l-1)(j\Delta)^{\frac{p_l}{p_l-1}} - \inf_{p\in T_l}(p-1)((j-1)\Delta)^q$.

Lemma 3.3. *There exists a constant $C' < \infty$ (depending only on π_0 and π_1) such that $\max_{j,l} D(j,l) \leq C' K^{\frac{\pi_0}{\pi_0-1}} \log(K)\Delta$.*

Using Lemma 3.2,

$$I_2 \leq \sum_{j=1}^{\frac{K}{\Delta}} \sum_{l=0}^{\frac{\pi_1-\pi_0}{\Delta}-1} \exp\left(nxt_{j,l}\frac{\mu(p_l)}{p_l}D(j,l)\right)\left(\mathbb{E}\exp\left(t_{j,l}\xi_{j,l}\right)\right)^n$$
$$\leq \frac{K(\pi_1-\pi_0)}{\Delta^2}\exp\left(n\frac{xC}{\Delta K^{\frac{\pi_0}{\pi_0-1}+1}}\max_{j,l} D(j,l)\right)$$
$$\times \sup_{c>0, p\in S'_\mu}\inf_{t\geq 0}\left(\mathrm{e}^{\frac{(t\delta)^2}{2}}\mathbb{E}\exp\left(t\left(cX - x\frac{\mu(p)}{p}(\Phi(X)+(p-1)c^q)\right)\right)\right)^n$$

and from Lemma 3.3 we deduce that

$$\limsup_{n\to\infty} I_2^{\frac{1}{n}} \leq \exp\left(xC''\frac{\log(K)}{K}\right)\sup_{c>0,p\in S'_\mu}\inf_{t\geq 0}\mathrm{e}^{\frac{(t\delta)^2}{2}}\mathbb{E}\exp\left(t\left(cX - x\frac{\mu(p)}{p}(\Phi(X)+(p-1)c^q)\right)\right).$$

The bound on I_3 is simpler: since $p \in S_\mu \backslash S'_\mu$,

$$I_3 \leq \mathbb{P}\left(\sum_{i=1}^{n} KX_i - xC\phi(X_1)\right) \leq \inf_{t\geq 0}\left(\mathbb{E}\exp\left(t(KX - xC\phi(X))\right)\right)^n \leq \left(\frac{1}{K}\right)^n$$

by Lemma 3.1, hence $\limsup_{n\to\infty} I_3^{\frac{1}{n}} \to 0$ when $K \to \infty$.

https://doi.org/10.1051/ps/2011112 Published online by Cambridge University Press

Putting all together, the upper bound follows by letting $K \to \infty$: for any $\delta > 0$ we have

$$\limsup_{n\to\infty} \mathbb{P}\left(\frac{\overline{X_n}}{M_{\phi,n}} \geq x\right)^{\frac{1}{n}} \leq \sup_{c>0, p\in S_\mu} \inf_{t\geq 0} \mathrm{e}^{\frac{(t\delta)^2}{2}} \mathbb{E}\exp\left(t\left(cX - x\frac{\mu(p)}{p}(\Phi(X) + (p-1)c^q)\right)\right).$$

Combining this with the following completes the proof of Theorem 2.1.

Lemma 3.4 (straightforward adaptation from [14], Lem. 2.1)**.** *For any random variable X we have*

$$\lim_{\delta\searrow 0} \sup_{b\geq 0} \inf_{t\geq 0} \mathrm{e}^{\frac{(t\delta)^2}{2}} \mathbb{E}\exp\left(t\left(cX - x\frac{\mu(p)}{p}(\Phi(X) + (p-1)c^q)\right)\right) = \sup_{b\geq 0} \inf_{t\geq 0} \mathbb{E}\exp\left(t\left(cX - x\frac{\mu(p)}{p}(\Phi(X) + (p-1)c^q)\right)\right)$$

for $x > \frac{0\vee\mathbb{E}X}{\psi(\mathbb{E}\phi(X))}$. Moreover, the convergence is uniform in $x \in [\alpha, 1]$ for any $\frac{0\vee\mathbb{E}X}{\psi(\mathbb{E}\phi(X))} < \alpha < 1$.

4. An application to Bahadur exact slope

4.1. Bahadur results

Let us recall here some basic facts about Bahadur exact slopes of test statistics. For a reference, see [1,11]. Consider a sample $X_1, \ldots, X_n$ having common law μ_θ depending on a parameter $\theta \in \Theta$. To test $(\mathcal{H}_0) : \theta \in \Theta_0$ against the alternative $(\mathcal{H}_1) : \theta \in \Theta_1 := \Theta\backslash\Theta_0$, we use a test statistic S_n, large values of S_n rejecting the null hypothesis. The p-value of this test is by definition $H_n(S_n)$, where

$$H_n(t) := \sup_{\theta\in\Theta_0} P_\theta(S_n \geq t).$$

The Bahadur exact slope $c(\theta)$ of S_n is then given by the following relation

$$c(\theta) = -2\liminf_{n\to\infty} \frac{1}{n}\log\left(H_n(S_n)\right). \tag{4.1}$$

Quantitatively, for $\theta \in \Theta_1$, the larger $c(\theta)$ is, the faster S_n rejects H_0.

A theorem of Bahadur (Thm. 7.2. in [1]) gives the following characterization of $c(\theta)$: suppose that $\lim_n n^{-1/2} S_n =: b(\theta)$ for any $\theta \in \Theta_1$, and that $\lim_n n^{-1}\log\left(H_n(n^{1/2}t)\right) =: -F(t)$ under any $\theta \in \Theta_0$. If F is continuous on an interval I containing $b(\Theta_1)$, then $c(\theta)$ is given by:

$$c(\theta) = 2F(b(\theta)). \tag{4.2}$$

4.2. The homogeneous case

For the sake of simplicity, we assume now that we are in the homogeneous case, *i.e.* $M_{\phi,n} = M_{p,n} := \left(\frac{1}{n}\sum_{i=1}^n |X_i|^p\right)^{\frac{1}{p}}$, with $p > 1$. Proofs can be adapted to the general case, adding suitable assumptions on ϕ.

We want to test whether the sample $X_1, \ldots, X_n$ has exponential distribution $\mathcal{E}(\beta)$, where $\beta > 0$ is unknown (null hypothesis H_0), against the alternative that it has gamma distribution $\gamma(\theta, \beta)$ for some $\theta > 1$ and $\beta > 0$ (alternative hypothesis H_1). To achieve this we use the statistic

$$T_n = \frac{\overline{X_n}}{M_{p,n}}.$$

Under H_0, T_n is scale-free, therefore we can assume $\beta = 1$.

https://doi.org/10.1051/ps/2011112 Published online by Cambridge University Press

From Theorem 1.1 of Shao [14],

$$\lim_{n\to\infty} \frac{1}{n} \log P(T_n \geq x) = -I(x)$$

where

$$I(x) = -\log\left(\sup_{c>0}\inf_{t\geq 0} \mathbb{E}\exp\left(t\left(cX - x\left(\frac{X^p}{p} + \frac{c^q}{q}\right)\right)\right)\right) \tag{4.3}$$

for any $x > \frac{\mathbb{E}X}{(\mathbb{E}X^p)^{1/p}} = \Gamma(p+1)^{-\frac{1}{p}}$. In order to apply Bahadur results, it remains to study the continuity of I on its domain. This is the object of the following lemma:

Lemma 4.1. *The rate function I is continuous on $\left(\Gamma(p+1)^{-\frac{1}{p}}, 1\right)$.*

Proof of Lemma 4.1. Set for any $(t,c,x) \in [0,\infty)\times(0,\infty)\times(T_0,+\infty)$

$$H(t,c,x) = \mathbb{E}\exp\left(t\left(cX - x\left(\frac{X^p}{p} + \frac{c^q}{q}\right)\right)\right).$$

Denote by

$$\Phi(u) = cu - x\left(\frac{u^p}{p} + \frac{c^q}{q}\right).$$

The function Φ is convex on $\mathbb{R}^+$ and reached its maximum at $u_0 = (\frac{c}{x})^{1/p}$. Moreover $\Phi(u_0) > 0$ if and only if $x < 1$. Since X has full support on $\mathbb{R}^+$,

$$P\left(cX - x\left(\frac{X^p}{p} + \frac{c^q}{q}\right) > 0\right) > 0.$$

From Lemmas 1 and 3 of [2], we know that for any (c,x), $\inf_{t\geq 0} H(t,c,x)$ is reached at some unique and finite $t = t(c,x) > 0$. The function $t(c,x)$ is given by the equation

$$\frac{\partial}{\partial t} H(t,c,x) = 0$$

more precisely, $\frac{\partial}{\partial t}H$ being $\mathcal{C}^1$ on its domain, by the Implicit Function theorem t is $\mathcal{C}^1$ at each point (c,x). Thus the same can be said of

$$\underline{H}(c,x) = \inf_{t\geq 0} H(t,c,x) = H(t(c,x),c,x).$$

Since $H \to 0$ when $c \to \infty$, we also know that $\sup_{c>0} H$ is reached, therefore $\sup_{c>0} \underline{H}(c,x)$ is reached for some $c = c(x)$ and this function is continuous on $\left(\Gamma(p+1)^{-\frac{1}{p}}, 1\right)$. So is $I(x) = -\log\left(\underline{H}(c(x),x)\right)$. □

Suppose that we want to test H_0 against the alternative that the sample has $\Gamma(\theta,\beta)$ distribution for $\theta > 1$ and $\beta > 0$. Under the alternative,

$$T_n \to T(\theta,\beta) = \theta\left(\frac{\Gamma(\theta)}{\Gamma(\theta+p)}\right)^{1/p}.$$

For $\theta > 1$, $T(\theta,\beta) > \Gamma(p+1)^{-\frac{1}{p}}$ and the Bahadur exact slope is therefore

$$c(\theta) = -\log\left(\sup_{c>0}\inf_{t\geq 0} \mathbb{E}\exp\left(t\left(cX - \theta\left(\frac{\Gamma(\theta)}{\Gamma(\theta+p)}\right)^{1/p}\left(\frac{X^p}{p} + \frac{c^q}{q}\right)\right)\right)\right).$$

https://doi.org/10.1051/ps/2011112 Published online by Cambridge University Press

Appendix A. Technical lemmata

Proof of Lemma 3.1. Take any $p > 1$ such that $\mu(p) < \infty$. Then $\phi(z) \geq \mu(p)^{-p} z^p$ and in particular $Kz - x\phi(z) \to -\infty$ when $z \to \infty$, so $\mathbb{E}\exp(KX - x\phi(X)) < \infty$. Observe also that for any $z > 0$, $Kz - xC\phi(z) \to -\infty$ when $C \to \infty$. By Lebesgue's dominated convergence theorem, it follows that $\mathbb{E}\exp(KX - xC\phi(X)) \to 0$ when $C \to \infty$. □

Proof of Lemma 3.2. It is easily checked that $\xi_{j,k}$ satisfies the same conditions as ξ_j in [14], in particular, thanks to the hypothesis on x, $-\infty \leq \mathbb{E}\xi_{j,l} < 0$. Therefore we know that there exists $t_{j,l} \geq 0$ such that

$$\begin{aligned} 1 &\geq \inf_{t \geq 0} \mathbb{E}\exp\left(t\xi_{j,l}\right) = \mathbb{E}\exp\left(t_{j,l}\xi_{j,l}\right) \\ &\geq \mathbb{E}\exp\left(t_j\left(\delta Y_{j,l} - x\frac{\mu(p_l)}{p_l}\left(\Phi(X_{j,l}) + (p_l - 1)(j\Delta)^{\frac{p_l}{p_l - 1}}\right)\right)\right) \end{aligned}$$

choosing B large enough so $\mathbb{P}(X \geq B) \geq \frac{1}{e}$,

$$\geq \exp\left(\frac{(t_{j,l}\delta)^2}{2} - 1 - t_{j,l}xC\left(\phi(B) + (\pi_1 - 1)K^{\frac{\pi_0}{\pi_0 - 1}}\right)\right) =: \exp(R(t_{j,l})).$$

It follows that $t_{j,l}$ has to be smaller than t^+, the largest root of R, so the lemma is proved with $\Delta := \frac{1}{t^+ K^{\frac{\pi_0}{\pi_0 - 1} + 1}}$. □

Proof of Lemma 3.3. We have

$$D(j,l) = \underbrace{(p_l - 1)(j\Delta)^{\frac{p_l}{p_l - 1}} - (p_l - 1)((j-1)\Delta)^{\frac{p_l}{p_l - 1}}}_{D_1} + \underbrace{(p_l - 1)((j-1)\Delta)^{\frac{p_l}{p_l - 1}} - \inf_{p \in T_l}(p - 1)((j-1)\Delta)^q}_{D_2}$$

with

$$D_1 \leq (p_l - 1) \sup_{c \in [j\Delta, (j-1)\Delta]} \frac{\partial}{\partial c} c^{\frac{p_l}{p_l - 1}} \Delta \leq \pi_1 K^{\frac{1}{\pi_0 - 1}} \Delta$$

and

$$D_2 \leq \sup_{p \in T_l} \left|\frac{\partial}{\partial p}(p - 1)((j-1)\Delta)^q\right| \Delta \leq K^{\frac{\pi_0}{\pi_0 - 1}} \log(K)\left(1 + \frac{1}{(\pi_0 - 1)^2}\right)\Delta.$$ □

References

[1] R.R. Bahadur, Some limit theorems in statistics. CBMS-NSF Regional Conference Series in *Appl. Math.* SIAM, Society for Industrial and Applied Mathematics, Philadelphia, Pa. (1971).

[2] H. Chernoff, A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Statist.* **23** (1952) 493–507.

[3] E. Cuvelier and M. Noirhomme-Fraiture, An approach to stochastic process using quasi-arithmetic means, in *Recent advances in stochastic modeling and data analysis*, World Scientific (2007) 2–9.

[4] E. Cuvelier and M. Noirhomme-Fraiture, Parametric families of probability distributions for functional data using quasi-arithmetic means with Archimedean generators, in *Functional and operatorial statistics, Contrib. Statist.* Springer (2008) 127–133.

[5] V.H. de la Peña, T.L. Lai and Q.-M. Shao, Self-normalized processes: Limit theory and statistical applications, in *Probab. Appl. (New York)*. Springer-Verlag, Berlin (2009).

https://doi.org/10.1051/ps/2011112 Published online by Cambridge University Press

[6] A. Dembo and Q.-M. Shao, Self-normalized large deviations in vector spaces, in *High dimensional probability (Oberwolfach, 1996)*, Birkhäuser, Basel. *Progr. Probab.* **43** (1998) 27–32.
[7] A. Dembo and Q.-M. Shao, Large and moderate deviations for Hotelling's T^2-statistic. *Electron. Comm. Probab.* **11** (2006) 149–159.
[8] G.H. Hardy, J.E. Littlewood and G. Pólya, *Inequalities*. Cambridge University Press, 2d ed. (1952).
[9] A. Kolmogoroff, Sur la notion de la moyenne. *Rendiconti Accad. d. L. Roma* **12** (1930) 388–391.
[10] T.L. Lai and Q.-M. Shao, Self-normalized limit theorems in probability and statistics, in *Asymptotic theory in probability and statistics with applications*, Int. Press, Somerville, MA *Adv. Lect. Math. (ALM)* **2** (2008) 3–43.
[11] Y. Nikitin, *Asymptotic efficiency of non parametric tests*. Cambridge University Press (1995).
[12] M. Nagumo, Über eine Klasse der Mittelwerte. *Japan. J. Math.* **7** (1930) 71–79.
[13] E. Porcu, J. Mateu and G. Christakos, Quasi-arithmetic means of covariance functions with potential applications to space-time data. *J. Multivar. Anal.* **100** (2009) 1830–1844.
[14] Q.-M. Shao, Self-normalized large deviations. *Ann. Probab.* **25** (1997) 285–328.
[15] Q.-M. Shao, A note on the self-normalized large deviation. *Chinese J. Appl. Probab. Statist.* **22** (2006) 358–362.
[16] A.V. Tchirina, Large deviations for a class of scale-free statistics under the gamma distribution. *J. Math. Sci.* **128** (2005) 2640–2655.
[17] R.R. Yager, On ordered weighted averaging aggregation operators in multicriteria decisionmaking. *IEEE Trans. Systems Man Cybernet.* **18** (1988) 183–190.

https://doi.org/10.1051/ps/2011112 Published online by Cambridge University Press