Bernoulli **6**(4), 2000, 621–632

Weak dimension-free concentration of measure

SERGEY G. BOBKOV[1] and CHRISTIAN HOUDRÉ[2]

[1]*Department of Mathematics, Syktyvkar University, 167001 Syktyvkar, Russia. E-mail: sbobkov@online.ru*
[2]*Southeast Applied Analysis Center, School of Mathematics, Georgia Institute of Technology, Atlanta GA 30332, USA. E-mail: houdre@math.gatech.edu*

We study a concentration property of product probability measures with respect to the supremum distance. This property is shown to be equivalent to the conditions given by de Haan and Ridder for the stochastic boundedness of centered extreme samples.

Keywords: concentration of measure; extreme samples; isoperimetry

1. Introduction

Let μ be a probability measure on the real line $\mathbb{R}$, and let μ^n be the n-fold tensor product of μ with itself. Given a notion of enlargement $\mathrm{enl}(A)$ for sets $A \subset \mathbb{R}^n$, inequalities of isoperimetric type have the form

$$\mu^n(\mathrm{enl}(A)) \geqslant R^{(n)}(\mu(A)).$$

Moreover, if $R = R^{(n)}$ is dimension-free, such inequalities are often viewed as concentration inequalities. One question of interest which will be addressed here is whether or not such a function (of course, such that $R(p) > p$) exists. Besides the measure, the answer depends in an essential manner on the enlargement which is usually built with the help of a metric, say ρ, by putting

$$\mathrm{enl}(A) = A^h = \{x \in \mathbb{R}^n : \rho(x, a) \leqslant h, \text{ for some } a \in A\},$$

where $h > 0$ is a fixed number (for A compact, A^h is the closed h-neighbourhood of A with respect to ρ). To consider the weakest possible type of enlargement, we equip $\mathbb{R}^n$ with the supremum distance

$$\rho_\infty(x, y) = \sup_{1 \leqslant i \leqslant n} |x_i - y_i|,$$

and consider the value

$$R_h^{(n)}(p) = \inf_{\mu^n(A) \geqslant p} \mu^n(A^h), \tag{1.1}$$

where the infimum is taken over all the Borel sets of measure $\mu^n(A) \geqslant p$, $p \in (0, 1)$. In his work on isoperimetry, Talagrand (1991) made the following observation (see Proposition 5.1 there): if $\inf_n R_h^{(n)}(\frac{1}{2}) > \frac{1}{2}$, then μ has finite exponential moment, that is,

$$\int_{\mathbb{R}} \exp(\varepsilon|x|)\,\mathrm{d}\mu(x) < +\infty,$$

for some $\varepsilon > 0$. In proving this result, he studied the behaviour of $\mu^n(A^h)$ for the cubes A. It turns out that studying the enlargements of the cubes also allows us to find necessary and sufficient conditions for the validity of the concentration inequality $\mu^n(A^h) \geqslant R(\mu(A))$, for some R such that $R(p) > p$. This property turns out to be equivalent to the stochastic boundedness of centered extreme samples. This boundedness was previously studied by de Haan and Ridder (1979) who explicitly described the corresponding class of underlying probability measures μ. For the exponential measure (and for all Lipschitz images of the exponential measure), Talagrand (1995) proved a concentration inequality for a notion of enlargement much smaller than the one defined by the supremum distance (cf. also Bobkov and Ledoux 1997). As we will see, beyond the class of Lipschitz images of the exponential measure, there exist probability distributions still enjoying some concentration property (as defined above).

Definition. *A function U defined on some interval $\Delta \subset \mathbb{R}$ is said to have finite modulus of continuity if, for all (equivalently, for some) $h > 0$,*

$$U^*(h) = \sup\{|U(x) - U(y)| : x,\ y \in \Delta,\ |x - y| \leqslant h\} < +\infty.$$

The function U^* is then called the modulus of continuity generated by U. Clearly, since $U^*(h_1 + h_2) \leqslant U^*(h_1) + U^*(h_2)$, for all $h_1,\ h_2 \geqslant 0$, U has finite modulus of continuity if and only if, for some $a,\ b \geqslant 0$, $|U(x) - U(y)| \leqslant a + b|x - y|$ whenever $x,\ y \in \Delta$.

Now let U_μ be defined as follows:

$$U_\mu(x) = F_\mu^{-1}\left(\frac{1}{1 + \mathrm{e}^{-x}}\right), \qquad x \in \mathbb{R},$$

where $F_\mu(x) = \mu((-\infty, x])$ is the distribution function of the measure μ, and where

$$F_\mu^{-1}(p) = \inf\{x \in \mathbb{R} : F_\mu(x) \geqslant p\}, \qquad p \in (0, 1),$$

is the minimal quantile of order p of μ. The meaning of this definition is that the map U_μ transforms the logistic probability measure ν ($\nu((-\infty, x]) = (1 + \mathrm{e}^{-x})^{-1}$) into the measure μ. The aim of these notes is to prove the following.

Theorem 1.1. *Let $p \in (0, 1)$. The following properties are equivalent:*

(a) *There exists $h > 0$ such that $\inf_n R_h^{(n)}(p) > p$.*
(b) *There exist $\delta > 0$ and $c > 0$ such that, for all $x \in \mathbb{R}$,*

$$F_\mu(x) - F_\mu(x - \delta) \geqslant cF_\mu(x)(1 - F_\mu(x)). \tag{1.2}$$

(c) *The function U_μ has finite modulus of continuity. In this case, for every $h > 0$, setting $h^* = U_\mu^*(h)$, we have*

$$\inf_n R_{h^*}^{(n)}(p) \geqslant \frac{p}{p + (1-p)\exp(-h)}, \tag{1.3}$$

with equality for $\mu = \nu$. In particular, the following alternative holds: either $\inf_n R_h^{(n)}(p) = p$, *for all* $h > 0$, *or* $\inf_n R_h^{(n)}(p) \to 1$, *as* $h \to +\infty$.

In more probabilistic language, inequality (1.3) can be expressed as follows. Let ξ_n, $n \geqslant 1$, be a sequence of independent random variables defined on some probability space $(\Omega, \mathscr{F}, P)$, with common law μ and associated distribution function F_μ. Let ζ be a logistic random variable (with law ν). Then, the right-hand side of (1.3) is simply $P\{\zeta - m_p(\zeta) \leqslant h\}$, where $m_p(\zeta) = F_\nu^{-1}(p)$ is the quantile of order p of ζ. Let $f : \mathbb{R}^n \to \mathbb{R}$ be an arbitrary Lipschitz function, with Lipschitz constant at most 1 with respect to ρ_∞, and let $\eta = f(\xi_1, \ldots, \xi_n)$. If (1.3) is applied to sets of the form $\{f \leqslant \text{const.}\}$, it is easily seen that

$$P\{\eta - m_p(\eta) \leqslant h^*\} \geqslant P\{\zeta - m_p(\zeta) \leqslant h\},$$

for all $p \in (0, 1)$ and $h > 0$. Furthermore, this can be shown to be equivalent to the following property: there exists a non-decreasing function $U_f : \mathbb{R} \to \mathbb{R}$ with $U_f^* \leqslant U_\mu^*$ such that the random variables η and $U_f(\zeta)$ are identically distributed. Thus, at the level of distributions, all the random variables $f(\xi_1, \ldots, \xi_n)$ where f is ρ_∞-Lipschitz can be viewed as random variables of the form $U(\zeta)$ with $U^* \leqslant U_\mu^*$. A simple consequence of this property is the fact that the variance $\operatorname{var}(f(\xi_1, \ldots, \xi_n))$ can be bounded by a quantity which only depends on U_μ^*. This is in particular true for the functions $f_n(x) = \max\{x_1, \ldots, x_n\}$ and $g_n(x) = \min\{x_1, \ldots, x_n\}$ which play a crucial role below.

Corollary 1.2. *Let* $p \in (0, 1)$. *There exists* $h > 0$ *such that* $\inf_n R_h^{(n)}(p) > p$, *if and only if the random variables* ξ_n *have finite second moment and the following two conditions hold:*

(1a) $\sup_n \operatorname{var} \max\{\xi_1, \ldots, \xi_n\} < +\infty$,
(2a) $\sup_n \operatorname{var} \min\{\xi_1, \ldots, \xi_n\} < +\infty$;

or equivalently, and more generally, if and only if, for any fixed $\alpha \geqslant 1$,

(1b) $\sup_n \mathrm{E}|\max\{\xi_1, \ldots, \xi_n\} - \mathrm{E}\max\{\xi_1, \ldots, \xi_n\}|^\alpha < +\infty$;
(2b) $\sup_n \mathrm{E}|\min\{\xi_1, \ldots, \xi_n\} - \mathrm{E}\min\{\xi_1, \ldots, \xi_n\}|^\alpha < +\infty$.

Moreover, in (1b) *and* (2b) *the second expectations can be replaced by the quantiles of order* $p \in (0, 1)$.

The properties (1a), (2a) can also be written (together but not separately) in a weaker form as follows: for some real numbers a_n and b_n, the random variables

$$\max\{\zeta_1, \ldots, \xi_n\} - a_n, \qquad \min\{\zeta_1, \ldots, \xi_n\} - b_n,$$

are stochastically bounded, or using different terminology, their distributions form a precompact family in the space of all probability measures on $\mathbb{R}$ with respect to the topology of weak convergence. That is,

(1c) $\sup_n P\{|\max\{\zeta_1, \ldots, \xi_n\} - a_n| > h\} \to 0$, as $h \to +\infty$;
(2c) $\sup_n P\{|\min\{\zeta_1, \ldots, \xi_n\} - b_n| > h\} \to 0$, as $h \to +\infty$.

De Haan and Ridder (1979) found, directly in terms of F_μ, several necessary and sufficient conditions for (1c) (cf. their Proposition 2.2.1 and Theorem 3.1), one of which is the following property: there exist x_0 and h_0 such that, for all $x \geqslant x_0$ and $h \geqslant h_0$,

$$1 - F_\mu(x + h) \leqslant c(1 - F_\mu(x)), \qquad \text{for some } c \in (0, 1).$$

When combining the above inequality with a similar inequality for $-\xi_1$, we arrive exactly at (1.2), with possibly another constant depending on μ only.

The description (1.2) is explicit and certainly convenient to use in the case of specific examples of probability distributions μ. However, for our purposes it will be essential to connect (1.2) with the moduli of continuity. When $\mu = \nu$, inequality (1.3) is known (cf. Bobkov 1996, Bobkov and Houdré 1997), and in fact, it is easy to prove (1.3) transporting ν into μ via U_μ. The non-trivial part of Theorem 1.1 will be to show that U_μ has finite modulus of continuity provided that $\inf_n R_h^{(n)}(p) > p$. Property (c) in Theorem 1.1 also allows us to make the following observation: it is possible for the tails $h \to \mu\{x \in \mathbb{R} : |x| > h\}$ to tend to zero exponentially fast (as $h \to +\infty$), or as fast as we want, without the products μ^n satisfying the concentration property (a). Indeed, given a decreasing, continuous function $\varepsilon : [0, +\infty) \to (0, \frac{1}{2}]$ with $\varepsilon(0) = \frac{1}{2}$ and $\varepsilon(h) \to 0$, as $h \to +\infty$, one can construct an even, continuous, strictly increasing function U on $\mathbb{R}$ with $U(0) = 0$ such that $U^* \equiv +\infty$ but such that the measure $\mu = \nu U^{-1}$, the image of ν under U, has the tails bounded by the function $\varepsilon(h)$. Thus, concentration property (a) is not determined by the tail behaviour of μ.

It is of course natural to ask if there exist necessary and sufficient conditions on F_μ for the (stronger) concentration property $\inf_n R_h^{(n)}(p) > p$ when one takes in (1.1) the enlargement A^h with respect to the usual Euclidean metric ρ_2 in $\mathbb{R}^n$. As far as we know, this question is still open. However, if we restrict ourselves in (1.1) to the class of sets $A \subset \mathbb{R}^n$ which are convex or whose complement is convex, the property $\inf_n R_h^{(n)}(p) > p$ will again be equivalent to (1.2) and thus we will obtain the same class of generating distributions μ (cf. Bobkov and Götze 1999). Moreover, for such and only such measures, the variances of all convex ρ_2-Lipschitz functions are bounded by a constant independent of the dimension.

The proofs of Theorem 1.1 and Corollary 1.2 are respectively given in Sections 3 and 4. Here we also discuss property (1b) in Corollary 1.2. We start (Section 2) with characterizations of the concentration property for the distributions of maxima. The paper finishes with some remarks.

2. Concentration of maxima

Below and throughout, let $M_n = \max\{\zeta_1, \ldots, \xi_n\}$.

Lemma 2.1. *The following are equivalent:*

(1a) *For some* $p \in (0, 1)$ *and* $h > 0$, $\inf_n P\{M_n - m_p(M_n) \leqslant h\} > p$.
(1b) *For some* $p \in (0, 1)$ *and* $h < 0$, $\sup_n P\{M_n - m_p(M_n) \leqslant h\} < p$.
(1c) *For a sequence of real numbers* a_n, $\sup_n P\{|M_n - a_n| > h\} \to 0$, *as* $h \to +\infty$.
(1d) *There exists* $\varepsilon > 0$ *such that, for any* $p \in (0, 1)$,

$$\sup_n \mathrm{E}\exp\{\varepsilon(M_n - m_p(M_n))\} < +\infty.$$

(2) *For all (equivalently, for some)* $a \in \mathbb{R}$, *the function* $U(x) = F_\mu^{-1}(1/(1 + \exp(-x)))$ *has finite modulus of continuity in the interval* $x \geqslant a$.

The equivalence between (1c) and (1d) is essentially known and is due to de Haan and Ridder (1979).

Proof. It will be convenient to work with another (equivalent) condition:

(2′) For all (equivalently, for some) $a \in \mathbb{R}$, the function $V(x) = F_\mu^{-1}(\exp(-\exp(-x)))$, $x \geqslant a$, has finite modulus of continuity. In addition, for any $p \in (0, 1)$ and $h > 0$,

$$\sup_n P\{M_n - m_p(M_n) > V_p^*(h)\} \leqslant P\{Z - m_p(Z) > h\}, \tag{2.1}$$

where the random variable Z has distribution $P\{Z \leqslant x\} = \exp(-\exp(-x))$, and where V_p^* is a modulus of continuity generated by V on the interval $[-\log\log(1/p), +\infty)$.

The main step in the proof is the implication (1a) $\Rightarrow$ (2′). Let $p, q \in (0, 1)$ and $h_0 > 0$ be such that

$$\inf_n P\{M_n - m_p(M_n) \leqslant h_0\} \geqslant q > p,$$

that is, such that $F_\mu(F_\mu^{-1}(p^{1/n}) + h_0) \geqslant q^{1/n}$. By the very definition of F_μ^{-1}, this implies

$$F_\mu^{-1}(q^{1/n}) - F_\mu^{-1}(p^{1/n}) \leqslant h_0. \tag{2.2}$$

Putting $a_0 = -\log\log(1/p)$, $b_0 = -\log\log(1/q)$, (2.2) can be rewritten as

$$V(b_0 + \log(n)) - V(a_0 + \log(n)) \leqslant h_0, \tag{2.3}$$

which holds for all $n \geqslant 1$. We need to deduce from (2.3) that

$$V_p^*(h) = \sup\{V(y) - V(x) : a \leqslant x \leqslant y,\ y - x \leqslant h\} < +\infty, \tag{2.4}$$

whenever $a \in \mathbb{R}$ and $h > 0$. Now, fix any real number c such that $1 < c < \exp(b_0 - a_0)$, and let n_0 be any positive integer such that

$$\log(n_0 + 1) - \log(n_0) \leqslant b_0 - a_0, \qquad (\exp(b_0 - a_0) - c)n_0 \geqslant 1. \tag{2.5}$$

Clearly, V is a non-decreasing function on $\mathbb{R}$, hence the property (2.4) does not depend on a and h, so we can let $a = a_0 + \log(n_0)$, $h = h_0$. Thus, in order to prove (2.4), it can be assumed that $a_0 + \log(n_0) \leqslant x \leqslant y \leqslant x + h_0$.

Now define a sequence n_k, $k \geqslant 1$, recursively in the following way: let n_1 be the largest integer such that $a_0 + \log(n_1) \leqslant x$; and if $k \geqslant 1$, let n_{k+1} be the largest integer such that $a_0 + \log(n_{k+1}) \leqslant b_0 + \log(n_k)$. Then $n_0 \leqslant n_k < n_{k+1}$, for all $k \geqslant 1$, since $a_0 + \log(n_k + 1) \leqslant b_0 + \log(n_k)$ which holds due to (2.5) and since $n_k \geqslant n_0$.

Denote by K the smallest k such that $b_0 + \log(n_k) \geqslant y$. By construction, the intervals

$\Delta_k = [a_0 + \log(n_k),\ b_0 + \log(n_k)]$, $1 \leqslant k \leqslant K$, cover the interval $[x, y]$. Therefore, using (2.3), we obtain

$$V(y) - V(x) \leqslant \sum_{k=1}^{K} V(b_0 + \log(n_k)) - V(a_0 + \log(n_k)) \leqslant Kh_0.$$

Our aim is now to find an estimate of K depending on $y - x$; we would then have an estimate for $V(y) - V(x)$ in terms of $y - x$. Denote by $[u]$ the integer part of a real u. Then $n_{k+1} = [\exp(b_0 - a_0)n_k]$, hence

$$n_{k+1} \geqslant \exp(b_0 - a_0)n_k - 1 \geqslant cn_k,$$

since $n_k \geqslant n_0$ and since $\exp(b_0 - a_0)n_0 \geqslant cn_0 + 1$. By induction, it is easy to see that $n_k \geqslant c^{k-1}n_1$, that is, $\log(n_k) \geqslant (k-1)\log(c) + \log(n_1)$. Thus, the inequality $b_0 + \log(n_k) \geqslant y$ follows from $b_0 + (k-1)\log(c) + \log(n_1) \geqslant y$. The last inequality can be rewritten as

$$k \geqslant 1 + \frac{y - b_0 - \log(n_1)}{\log(c)}. \tag{2.6}$$

By the very definition of n_1, we also have $\log(n_1 + 1) > x - a_0$, and since $\log(n_1 + 1) - \log(n_1) \leqslant b_0 - a_0$, we have the estimate $\log(n_1) > x - b_0$. Therefore, (2.6) is fulfilled if we take k such that $k \geqslant 1 + (y - x)/\log(c)$. Hence

$$K \leqslant 2 + \frac{y - x}{\log(c)}.$$

We thus have proved (2.4) and the first part of (2′).

To prove the second part of (2′), fix $p \in (0, 1)$, $h > 0$, and set $r = P\{Z - m_p(Z) \leqslant h\}$, $a = -\log\log(1/p)$, $b = -\log\log(1/r)$. Then, as easily verified, $b - a = h$. As previously seen, inequalities of the form

$$P\{M_n - m_p(M_n) \leqslant V_p^*(h)\} \geqslant r \tag{2.7}$$

are equivalent to

$$V(b + \log(n)) - V(a + \log(n)) \leqslant V_p^*(h)$$

(this is (2.3) with $V_p^*(h)$ instead of h_0). Since $b - a = h$, the above inequality holds true by the very definition of V_p^*. It just remains to note that (2.7) and (2.1) coincide.

(2′) ⇒ (1a). Let $p = \exp(-\exp(-a))$. Then, as shown in the previous steps, (1a) holds if and only if (2.4) holds for some $b > a$, some $h > 0$ and all $n \geqslant 1$. But (2.4) holds for all $b > a$ and $h > 0$ since $V(b + \log(n)) - V(a + \log(n)) \leqslant V_p^*(b - a)$.

(2) ⇒ (2′). Note that $U(x) = V(T(x))$, where $T(x) = -\log\log(1 + \mathrm{e}^{-x})$. Then T is an increasing bijection from $\mathbb{R}$ to $\mathbb{R}$, and has a finite Lipschitz constant on every interval $[a, +\infty)$, and similarly for its inverse T^{-1}. Therefore, U has finite modulus of continuity on $[a, +\infty)$ if and only if V has finite modulus of continuity on $[a, +\infty)$.

(1a) ⇒ (1b). Simply note (recalling (2.2)) that, for all $0 < p < q < 1$ and all $h > 0$,

$$P\{M_n - m_p(M_n) < h\} \geqslant q \Rightarrow P\{M_n - m_q(M_n) \leqslant -h\} < p. \tag{2.8}$$

$(2') \Rightarrow$ (1c). Let $a_n = m_p(M_n)$. Then (2.1) implies that $\sup_n P\{M_n - a_n > h\} \to 0$, as $h \to 0$, so we need to estimate the left deviations $\sup_n P\{M_n - a_n < h\}$. Take $a_n = m_q(M_n)$ with fixed (but arbitrary) $q \in (0, 1)$. Inserting in (2.8) $V_p^*(h) + \varepsilon$ instead of h ($\varepsilon > 0$), and letting $\varepsilon \to 0^+$, gives, for all $p \in (0, q)$ and for all $h > 0$:

$$P\{M_n - m_p(M_n) \leqslant V_p^*(h)\} \geqslant q \Leftrightarrow \tag{2.9}$$

$$P\{M_n - m_q(M_n) \leqslant -h'\} < p, \qquad \text{for all } h' > V_p^*(h). \tag{2.10}$$

If p is chosen so that $q = P\{Z - m_p(Z) \leqslant h\}$, that is, $\log\log(1/p) - \log\log(1/q) = h$, then (2.9) is true, thanks to (2.4), hence (2.10) holds. It remains to note that $p \to 0$ as $h \to +\infty$, and since $V_p^*(h)$ is finite, we conclude that $P\{M_n - a_n \leqslant -h'\} \to 0$ as $h' \to +\infty$.

(1c) $\Rightarrow$ (1a). Without loss of generality, we may prove (1a) for $p = \frac{1}{2}$. By assumption, there exists h_0 such that $P\{|M_n - a_n| > h_0\} < \frac{1}{2}$, for all $n \geqslant 1$. Hence, $|m_p(M_n) - a_n| \leqslant h_0$. Therefore,

$$\sup_n P\{M_n - m_p(M_n) > h\} \leqslant \sup_n P\{M_n - a_n > h - h_0\} \to 0, \qquad \text{as } h \to +\infty.$$

Finally, it is clear that $(2')$ implies (1d) which in turn implies (1a). Thus, Lemma 2.1 is proved. $\square$

3. Proof of Theorem 1.1

(a) $\Rightarrow$ (c). Assume that there exist $h > 0$ and $0 < p < q < 1$, such that $\inf_n R_h^{(n)}(p) \geqslant q$. Thus, for all integers $n \geqslant 1$ and for all Borel sets $A \subset \mathbb{R}^n$ with $\mu^n(A) \geqslant p$, we have

$$\mu^n(A^h) \geqslant q. \tag{3.1}$$

Applying (3.1) first to the cubes $A_n(p) = (-\infty, F_\mu^{-1}(p^{1/n})]^n$, and since $\mu^n(A_n(p)) = F_\mu^n(F_\mu^{-1}(p^{1/n})) \geqslant p$ and $A_n(p)^h = (-\infty, F_\mu^{-1}(p^{1/n}) + h]^n$, gives

$$P\{M_n - m_p(M_n) \leqslant h\} = F_\mu^n(F_\mu^{-1}(p^{1/n}) + h) = \mu^n(A_n(p)^h) \geqslant q,$$

that is, property (1a) of Lemma 2.1 is fulfilled. Therefore, so is property (2): the function U_μ has finite modulus of continuity on the interval $[0, +\infty)$. Now, apply (3.1) to the cubes $B_n(p) = [-F_\mu^{-1}(1 - p^{1/n}), +\infty)^n$. By applying the same argument to the random variables $-\xi_n$, $n \geqslant 1$, we see that U_μ has finite modulus of continuity on $(-\infty, 0]$, and therefore on the whole real line.

(c) $\Rightarrow$ (a). It is known that, for the measure ν with $\nu((-\infty, x]) = 1/(1 + \exp(-x))$,

$$\nu^n(A^h) \geqslant \frac{p}{p + (1-p)\exp(-\mathrm{h})}, \tag{3.2}$$

whenever $\nu^n(A) \geqslant p$, with equality at the standard half-spaces $A = \{x : x_1 \leqslant \text{const.}\}$ – different proofs of (3.2) can also be found in Bobkov (1996) and in Bobkov and Houdré (1997, Corollary 15.3). Introduce the function $i(x_1, \ldots, x_n) = (U_\mu(x_1), \ldots, U_\mu(x_n))$ which transforms ν^n into μ^n. Let $h > 0$, $h^* = U_\mu^*(h)$. Now observe the following inclusion: for any set $A \subset \mathbb{R}^n$,

$$(i^{-1}(A))^h \subset i^{-1}(A^{h^*}). \tag{3.3}$$

Indeed, if $x \in (i^{-1}(A))^h$, then for some $y \in i^{-1}(A)$ we have $\rho_\infty(x, y) \leqslant h$, that is, $|x_k - y_k| \leqslant h$, for all $1 \leqslant k \leqslant n$. Since $i(y) = (U_\mu(y_1), \ldots, U_\mu(y_n))A$ and since $|U_\mu(x_k) - U_\mu(y_k)| \leqslant h^*$, we obtain $\rho_\infty(i(x), i(y)) \leqslant h$, and therefore $i(x) \in A^{h^*}$, hence $x \in i^{-1}(A^{h^*})$. Now combine (3.2) and (3.3) to prove (1.3). Let $\nu^n(A) = \mu^n(i^{-1}(A)) \geqslant p$; then

$$\mu^n(A^{h^*}) = \nu^n(i^{-1}(A^{h^*})) \geqslant \nu^n((i^{-1}(A))^h) \geqslant \frac{p}{p + (1-p)\exp(-h)}.$$

To complete the proof of Theorem 1.1, it remains to see the equivalence of (b) and (c). A quantitative version of this equivalence is given by the following lemma.

Lemma 3.1. *Given $\delta > 0$ and $c > 0$, assume that*

$$F_\mu(x) - F_\mu(x - \delta) \geqslant cF_\mu(x)(1 - F_\mu(x)), \qquad \text{for all } x \in \mathbb{R}. \tag{3.4}$$

Then $U_\mu^*(h) \leqslant \delta$ with $h = \log(1 + c)$. Conversely, if, for some positive h and δ, $U_\mu^*(h) \leqslant \delta$, then (3.4) holds with $c = 1 - \mathrm{e}^{-h}$.

Proof. Recall that $U_\mu(x) = F_\mu^{-1}(F_\nu(x))$, $x \in \mathbb{R}$. Define the values $U_\mu(-\infty)$ and $U_\mu(+\infty)$ in the usual limiting sense. Define also the function $U_\mu^{-1} : \mathbb{R} \to [-\infty, +\infty]$ by $U_\mu^{-1}(a) = F_\nu^{-1}(F_\mu(a))$ so that $F_\mu(a) = F_\nu(U_\mu^{-1}(a))$, for all $a \in \mathbb{R}$. It is straightforward to verify that, for all $z \in \mathbb{R}$,

$$U_\mu(U_\mu^{-1}(z)) \leqslant z \leqslant U_\mu^{-1}(U_\mu(z)). \tag{3.5}$$

To prove the lemma, we first assume that $U_\mu^*(h) \leqslant \delta$ and derive (3.4). Fix $x \in \mathbb{R}$ and assume that $0 < F(x - \delta) \leqslant F(x) < 1$ (otherwise (3.4) is immediate). Thus, the value $a = U_\mu^{-1}(x - \delta)$ is finite. By the assumption and by the left inequality in (3.5) with $z = x - \delta$, we obtain

$$U_\mu(a + h) - \delta \leqslant U_\mu(a) \leqslant x - \delta,$$

that is, $x \geqslant U_\mu(a + h)$. Taking U_μ^{-1} of both sides and applying the right-hand inequality in (3.5), we obtain

$$U_\mu^{-1}(x) \geqslant a + h = U_\mu^{-1}(x - \delta) + h. \tag{3.6}$$

We also need the following trivial inequalities for the logistic distribution:

$$1 - \mathrm{e}^{v-u} \leqslant \frac{F_\nu(u) - F_\nu(v)}{F_\nu(u)(1 - F_\nu(u))} \leqslant \mathrm{e}^{u-v} - 1, \qquad \text{for all } u \geqslant v. \tag{3.7}$$

Applying (3.6) and the left-hand inequality in (3.7) with $u = U_\mu^{-1}(x)$, $v = U_\mu^{-1}(x) - h$, we obtain that

$$\begin{aligned}F_\mu(x) - F_\mu(x-\delta) &= F_\nu(U_\mu^{-1}(x)) - F_\nu(U_\mu^{-1}(x-\delta))\\ &\geqslant F_\nu(U_\mu^{-1}(x)) - F_\nu(U_\mu^{-1}(x) - h)\\ &\geqslant (1-\mathrm{e}^{-h})F_\nu(U_\mu^{-1}(x))(1-F_\nu(U_\mu^{-1}(x)))\\ &= (1-\mathrm{e}^{-h})F_\mu(x)(1-F_\mu(x)).\end{aligned}$$

This proves the second assertion of Lemma 3.1. The first assertion can be proved in a similar way, with the help of the right-hand inequality in (3.7). Actually it is also proved, as Lemma 4.5 in Bobkov and Götze (1999), for the related two-sided exponential distribution (instead of ν), with $h = \log(1 + c/2)$. □

Corollary 3.2. *In Theorem* 1.1, *for every* $h > 0$, *the concentration inequality* (1.3) *holds with*

$$h^* = \left(\frac{h}{\log(1+c)} + 1\right)\delta,$$

where c *and* δ *are from* (1.2).

Indeed, by the first assertion of Lemma 3.1, $U_\mu^*(\log(1+c)) \leqslant \delta$. Hence, for any integer $k \geqslant 1$, we have $U_\mu^*(k\log(1+c)) \leqslant k\delta$. Taking $k = [h/\log(1+c)] + 1$, we obtain $h \leqslant k\log(1+c)$, so that $U_\mu^*(h) \leqslant (h/\log(1+c)+1)\delta$.

4. Concentration of maxima and minima in L^α-norm

Proof of Corollary 1.2.
Sufficiency. Assume that (1b) is true. Then, for the sequence $a_n = \mathrm{E}M_n$, or $a_n = m_p(M_n)$ (as stated at the end of Corollary 1.2), Chebyshev's inequality implies that $\sup_n P\{|M_n - a_n| > h\} \to 0$ as $h \to +\infty$. Thus, property (1c) of Lemma 2.1 is fulfilled, so the function U_μ has finite modulus of continuity on the interval $[0, +\infty)$. Assumption (2b) is just (1b) for the sequence $(-\xi_n)$, $n \geqslant 1$. Hence, again by Lemma 2.1, the function U_μ has finite modulus of continuity on the interval $(-\infty, 0]$. As a result, U_μ has finite modulus on the whole real line. It now remains to make use of Theorem 1.1.

Necessity. As before, let $M_n = \max\{\xi_1, \ldots, \xi_n\}$. Let ζ_n, $n \geqslant 1$, be a sequence of independent random variables with common (logistic) distribution ν, and let $Z_n = \max\{\zeta_1, \ldots, \zeta_n\}$. Since U_μ transforms ν into μ, M_n and $U_\mu(Z_n)$ are identically distributed. Therefore,

$$\mathrm{E}|M_n - M_n'|^\alpha = \mathrm{E}|U_\mu(Z_n) - U_\mu(Z_n')|^\alpha, \tag{4.1}$$

where (M_n', Z_n') is an independent copy of (M_n, Z_n). By Theorem 1.1, there exist constants $a, b \geqslant 0$ such that $|U_\mu(x) - U_\mu(y)| \leqslant a + b|x-y|$ whenever $x, y \in \mathbb{R}$. Thus, for all $\alpha \geqslant 1$, the left-hand side of (4.1) is bounded as $n \to \infty$, if the same is true for Z_n instead of M_n. That is to say, we have reduced our attempt at a proof to the case $\mu = \nu$. So, one may assume that $\xi_n = \zeta_n$, and that $M_n = Z_n$, for all n. In this special case, U_μ is the identity function. Therefore, applying a remark following the statement of Theorem 1.1 to the functions $f_n(x) = \max_{1 \leqslant k \leqslant n} x_k$, there exist Lipschitz functions $U_n : \mathbb{R} \to \mathbb{R}$, with Lipschitz constants at most 1, such that the random variables M_n and $U_n(\xi_1)$ are identically distributed (of course, in this particular case, this is easily verified directly). Therefore, for all $n \geqslant 1$,

$$\mathrm{E}|M_n - M_n'|^\alpha = \mathrm{E}|U_n(\xi_1) - U_n(\xi_1')|^\alpha \leqslant \mathrm{E}|\xi_1 - \xi_1'|^\alpha = C_\alpha,$$

where ξ_1' is an independent copy of ξ_1. Now, by Hölder's inequality,

$$\mathrm{E}|M_n - \mathrm{E}M_n|^\alpha = \mathrm{E}|\mathrm{E}'(M_n(\omega) - M_n'(\omega'))|^\alpha \leqslant \mathrm{E}|M_n - M_n'|^\alpha \leqslant C_\alpha,$$

where E' is taken with respect to the random variable M_n'. This proves (1b). Property (2b) is proved in a similar way, taking into account that ν is symmetric about 0. In order to prove the last statement on the quantiles, one can apply (1.3) to the cubes $\{x : x_i \leqslant \text{const.}, \text{ for all } i \leqslant n\}$. This gives

$$P\{M_n - m_p(M_n) > h^*\} \leqslant P\{\zeta_1 - m_p(\zeta_1) > h\},$$
$$P\{M_n - m_p(M_n) < -h^*\} \leqslant P\{\zeta_1 - m_p(\zeta_1) < -h\},$$

for all $p \in (0, 1)$, $h > 0$. Since $h^* \leqslant a + bh$, these inequalities immediately imply that

$$\sup_n \mathrm{E}|M_n - m_p(M_n)|^\alpha < +\infty.$$

Corollary 1.2 follows. □

One may wonder how to express (1b), the concentration property of maxima in L^α-norm, separately from (2b). Using Corollary 1.2, one can derive the following description.

Corollary 4.1. *Let* $\mathrm{E}|\xi_1|^\alpha < +\infty$, $\alpha \geqslant 1$. *The following are equivalent:*

(1) $\sup_n \mathrm{E}|M_n - \mathrm{E}M_n|^\alpha < +\infty$.

(2) *There exist* a, $b \geqslant 0$ *such that, for all* $\frac{1}{2} \leqslant p \leqslant q < 1$,

$$F_\mu^{-1}(q) - F_\mu^{-1}(p) \leqslant a + b \log \frac{1-p}{1-q}. \tag{4.2}$$

Corollary 4.1 (generalizing a statement of the authors in the case $\alpha = 1$) and its elegant proof were kindly indicated to us by a referee. As also mentioned to us, these arguments also apply to more general norms.

Recall that $F_\mu^{-1}(p)$ is the minimal quantile of order p of F_μ. Thus, condition (4.2) expresses the fact that the function $U_\mu(x) = F_\mu^{-1}(1/(1+\exp(-x)))$ sending ν into μ has a finite modulus of continuity on the interval $[0, +\infty)$. This implies that $\mathrm{E}(\xi_1^+)^\alpha < +\infty$, and moreover that $\mathrm{E}\exp(\varepsilon\xi_1) < +\infty$, for some $\varepsilon > 0$ (as usual, $x^+ = \max\{x, 0\}$). However, it says nothing about the behaviour of F_μ at $-\infty$.

Proof of Corollary 4.1. When $\xi_i \geqslant 0$, the proof is immediate since then $0 \leqslant \min\{\xi_1, \ldots, \xi_n\} \leqslant \xi_1$, which implies (2b). In general, one can observe that

$$M_n^+ - |\xi_1| \leqslant M_n \leqslant M_n^+,$$

so that

$$|M_n - \mathrm{E}M_n| \leqslant |M_n^+ - \mathrm{E}M_n^+| + |\xi_1| + \mathrm{E}|\xi_1|.$$

Therefore, $\sup_n \mathrm{E}|M_n - \mathrm{E}M_n|^\alpha < +\infty$ if and only if $\sup_n \mathrm{E}|M_n^+ - \mathrm{E}M_n^+|^\alpha < +\infty$. On the other hand, $M_n^+ = \max\{\zeta_1^+, \ldots, \xi_n^+\}$ corresponds to non-negative random variables. Hence, by the previous step, the sequence $\mathrm{E}|M_n^+ - \mathrm{E}M_n^+|^\alpha$ is bounded if and only if U_μ has finite modulus of continuity on $[0, +\infty)$. This finishes the proof. $\square$

In the case $\alpha = 1$, condition (2) together with $\mathrm{E}|\xi_1| < +\infty$ can equivalently be written as one property:

$$\int_0^1 p^n(1-p)\,\mathrm{d}F_\mu^{-1}(p) = O\left(\frac{1}{n}\right), \qquad \text{as } n \to \infty.$$

In turn, this leads to another formulation.

Corollary 4.2. *Let ξ be a random variable with values in* $(0, 1)$ *and with distribution function F_ξ. Then,*

$$\mathrm{E}\xi^n = O\left(\frac{1}{n}\right), \qquad \text{as } n \to \infty,$$

if and only if there exist a, $b \geqslant 0$ such that, for all $\frac{1}{2} \leqslant p \leqslant q < 1$,

$$\int_p^q \frac{1}{1-t}\,\mathrm{d}F_\xi(t) \leqslant a + b \log \frac{1-p}{1-q}.$$

5. Concluding remarks

Let (X, ρ, μ) be a metric space equipped with a Borel probability measure μ. As in Section 1, define the open h-neighbourhood of a set $A \subset X$ by

$$A^h = \{x \in \mathbb{R}^n : \rho(x, a) < h \text{ for some } a \in A\}, \qquad h > 0,$$

and the associated ('integral') isoperimetric function

$$R_h(p) = \inf_{\mu(A) \geqslant p} \mu(A^h), \qquad 0 < p < 1,$$

where the infimum is taken over all Borel sets of measure $\mu(A) \geqslant p$. With this notation, one can easily prove the following statement which contrasts with the $p-1$ alternative of Theorem 1.1, and deals with n fixed.

Proposition 5.1. $R_h(p) \to 1$ *as* $h \to +\infty$ *whenever* $p \in (0, 1)$.

This statement remains true if the metric is replaced by a pseudo-metric ρ such that

$\rho(x, y) < +\infty$, for almost all (x, y) with respect to measure $\mu \otimes \mu$. In Section 1, the concentration property $\inf_n R_h^{(n)}(p) > p$ could also have been expressed as $R_h(p) > p$, for the space $X = \mathbb{R}^\infty$ equipped with the pseudo-metric

$$\rho_\infty(x, y) = \sup_{i \geqslant 1} |x_i - y_i|,$$

and with the product measure μ^∞. In this case, $\rho(x, y) = +\infty$, for almost all (x, y) with respect to μ^∞, whenever the measure μ does not have compact support on the real line.

Acknowledgements

The research of the first author was supported in part by the Alexander von Humboldt Foundation and the Russian Foundation for Fundamental Research, grant 96-01-00201. This author enjoyed the hospitality of the School of Mathematics, Georgia Institute of Technology, while part of the research was being carried out.

The research of the second author was supported in part by an NSF Postdoctoral Fellowship and by National Science Foundation grant no. DMS-9803239. This author enjoyed the hospitality of the Steklov Mathematical Institute (St Petersburg branch) and of the Department of Mathematics, University of Syktyvkar, while part of the research was being carried out.

We thank the referees for pointing out to us the work of de Haan and Ridder and for other important comments including Corollary 4.1, with $\alpha > 1$, and its elegant proof.

References

Bobkov, S.G. (1996) Extremal properties of half-spaces for log-concave distributions. *Ann. Probab.*, **24**, 35–48.

Bobkov S.G. and Götze, F. (1999) Discrete isoperimetric and Poincaré-type inequalities. *Probab. Theory Related Fields*, **114**, 245–277.

Bobkov, S.G. and Houdré, C. (1997) *Some Connections between Sobolev-Type Inequalities and Isoperimetry*, *Mem. Amer. Math. Soc.* 616. Providence, RI: American Mathematical Society.

Bobkov, S.G. and Ledoux, M. (1997) Poincaré inequalities and Talagrand's concentration phenomenon for the exponential distribution. *Probab. Theory Related Fields*, **107**, 383–400.

de Haan, L. and Ridder, G. (1979) Stochastic compactness of sample extremes. *Ann. Probab.*, **7**, 290–303.

Talagrand, M. (1991) A new isoperimetric inequality and the concentration of measure phenomenon. In J. Lindenstrauss and V.D. Milman (eds), *Geometric Aspects of Functional Analysis*, Lecture Notes in Math. 1469, pp. 94–124. Berlin: Springer-Verlag.

Talagrand, M. (1995) Concentration of measure and isoperimetric inequalities in product spaces. *Inst. Hautes Études Sci. Publ. Math.*, **81**, 73–205.

Received June 1996 and revised January 2000