LARGE DEVIATIONS FOR SUMS OF I.I.D. RANDOM COMPACT SETS

Raphaël CERF

CNRS, Université Paris Sud

September 1997

Abstract. We prove a large deviation principle for Minkowski sums of i.i.d. random compact sets in a Banach space, that is the analog of Cramér theorem for random compact sets.

Several works have been devoted to deriving limit theorems for random sets. For i.i.d. random compact sets in $\mathbb{R}^p$, the law of large numbers was initially proved by Artstein and Vitale [1] and the central limit theorem by Cressie [3], Lyashenko [10] and Weil [16]. For generalizations to non compact sets, see also Hess [8]. These limit theorems were generalized to the case of random compact sets in a Banach space by Giné, Hahn and Zinn [7] and Puri and Ralescu [11]. Our aim is to prove a large deviation principle for Minkowski sums of i.i.d. random compact sets in a Banach space, that is, to prove the analog of the Cramér theorem.

We consider a separable Banach space F with norm $|| \quad ||$. We denote by $\mathcal{K}(F)$ the collection of all non empty compact subsets of F. For an element A of $\mathcal{K}(F)$, we denote by $\operatorname{co} A$ the closed convex hull of A. Mazur's theorem [5, p 416] implies that, for A in $\mathcal{K}(F)$, $\operatorname{co} A$ belongs to $\operatorname{co}\mathcal{K}(F)$, the collection of the non empty compact convex subsets of F. The space $\mathcal{K}(F)$ is equipped with the Minkowski addition and the scalar multiplication: for A_1, A_2 in $\mathcal{K}(F)$ and λ a real number,

$$A_1 + A_2 = \{ a_1 + a_2 : a_1 \in A_1,\, a_2 \in A_2 \}, \quad \lambda A_1 = \{ \lambda a_1 : a_1 \in A_1 \}.$$

1991 *Mathematics Subject Classification.* 60D05, 60F10.

Key words and phrases. Cramér theorem, random sets, large deviations.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX

The Hausdorff distance

$$d(A_1,A_2)\,=\,\max\left\{\sup_{a_1\in A_1}\inf_{a_2\in A_2}||a_1-a_2||,\,\sup_{a_2\in A_2}\inf_{a_1\in A_1}||a_2-a_1||\right\}$$

makes $(\mathcal{K}(F),d)$ a complete separable metric space (i.e. a Polish space). We endow $\mathcal{K}(F)$ with the Borel σ–field associated to the Hausdorff topology.

We denote by F^* the topological dual of F and by B^* the unit ball of F^*. The Banach–Alaoglu theorem asserts that B^* endowed with the weak$*$ topology $w*$ is compact [13]. Moreover the space (B^*,w^*) is separable and metrizable. We denote by $M(B^*)$ the set of Borel signed measures on B^* (the σ–field being the σ–field generated by the weak$*$ topology). Let $(\Omega,\mathcal{F},P)$ be a probability space. A random compact set of F is a measurable function from Ω to $\mathcal{K}(F)$ i.e. a random variable with values in $\mathcal{K}(F)$.

We suppose that F is of type $p>1$ i.e. there exists a constant c such that

$$E||\sum_{i=1}^n f_i||^p\,\leq\,c\sum_{i=1}^n E||f_i||^p$$

for any independent random variables $f_1,\cdots,f_n$ with values in F and mean zero. Every Hilbert space is of type 2, the spaces L^p with $1<p<\infty$ are of type $\min(p,2)$. However the space of continuous functions on $[0,1]$ equipped with the supremum norm is of type 1 and not of type p for any $p>1$.

We denote by $\mathbb{N}^*$ the set of positive integers.

Theorem 1. *Let $(A_n)_{n\in\mathbb{N}^*}$ be a sequence of i.i.d. random compact sets of F such that*

$$\forall\alpha\in\mathbb{R}^+\qquad E(\exp(\alpha\sup_{a\in A_1}||a||))\,<\,+\infty\,.$$

For a measure λ of $M(B^*)$ we set

$$\Lambda(\lambda)\,=\,\ln E\exp\left(\int_{B^*}\sup_{a\in A_1}x^*(a)\,d\lambda(x^*)\right)$$

and for a set U belonging to $co\,\mathcal{K}(F)$,

$$\Lambda^*(U)\,=\,\sup_{\lambda\in M(B^*)}\left(\int_{B^*}\sup_{x\in U}x^*(x)\,d\lambda(x^*)\,-\,\Lambda(\lambda)\right).$$

For a non convex set U in $\mathcal{K}(F)$ we set $\Lambda^*(U)=+\infty$.
Then the law of the random set $S_n=(A_1+\cdots+A_n)/n$ satisfies a large deviation principle with rate function Λ^* i.e. for any subset $\mathcal{U}$ of $\mathcal{K}(F)$

$$-\inf_{U\in\,\mathrm{interior}(\mathcal{U})}\Lambda^*(U)\,\leq\,\liminf_{n\to\infty}\frac{1}{n}\ln P(S_n\in\mathcal{U})$$
$$\leq\,\limsup_{n\to\infty}\frac{1}{n}\ln P(S_n\in\mathcal{U})\,\leq\,-\inf_{U\in\,\mathrm{closure}(\mathcal{U})}\Lambda^*(U)$$

(where interior($\mathcal{U}$) and closure($\mathcal{U}$) are the interior and the closure of $\mathcal{U}$ with respect to the Hausdorff topology).

Remark. The rate function Λ^* is a "good" rate function i.e. it is lower semicontinuous and its level sets $\{\, U \in \mathcal{K}(F) : \Lambda^*(U) \leq \lambda \,\}$, $\lambda \in \mathbb{R}^+$, are compact.

We first collect several results which are the main ingredients to prove theorem 1.

An embedding theorem. To a compact convex subset A of F we associate its support function $s_A : B^* \to \mathbb{R}$ defined by

$$\forall x^* \in B^* \qquad s_A(x^*) \;=\; \sup \{\, x^*(x) : x \in A \,\}\,.$$

We denote by $\mathcal{C}(B^*, w*)$ the set of continuous functions from B^* endowed with the weak$*$ topology to $\mathbb{R}$. With the uniform norm $|| \quad ||_\infty$, $\mathcal{C}(B^*, w*)$ is a separable Banach space (for f in $\mathcal{C}(B^*, w*)$, $||f||_\infty = \sup_{x^* \in B^*} |f(x^*)|$). Whenever A is compact, its support function s_A belongs to $\mathcal{C}(B^*, w*)$. The map $s : \operatorname{co}\mathcal{K}(F) \to \mathcal{C}(B^*, w*)$ has the following properties. For any A_1, A_2 in $\operatorname{co}\mathcal{K}(F)$ and t in $\mathbb{R}^+$,

$$s_{A_1} = s_{A_2} \iff A_1 = A_2, \quad A_1 \subset A_2 \iff s_{A_1} \leq s_{A_2},$$
$$s_{A_1 + A_2} = s_{A_1} + s_{A_2}\,, \quad s_{tA_1} = t s_{A_1}\,,$$

and finally $d(A_1, A_2) \;=\; ||s_{A_1} - s_{A_2}||_\infty$. Hence $\operatorname{co}\mathcal{K}(F)$ is algebraically and topologically isomorphic to its image under s, $s(\operatorname{co}\mathcal{K}(F))$, which is a subset of the separable Banach space $\mathcal{C}(B^*, w*)$. This embedding theorem was used in [1] and [7] to prove limit theorems for random sets. In the context of normed spaces, this theorem is due to Rådström [12] and Hörmander [9].

A general Cramér theorem. We state here a slightly weakened version of the general Cramér theorem (see either [4, theorem 3.1.6 and corollary 3.1.7] or [6, theorem 6.1.3]).

Let E be a separable Banach space and let E_1 be a closed convex subset of E. Let $(X_n)_{n \in \mathbb{N}^*}$ be a sequence of i.i.d. random variables defined on $(\Omega, \mathcal{F}, P)$ with values in E_1 and set $S_n = (X_1 + \cdots + X_n)/n$. Suppose that for any λ in the dual E^*

$$\Lambda_E(\lambda) \;=\; \ln E \exp(\lambda(X_1)) \;<\; +\infty$$

and that the laws of $(S_n)_{n \in \mathbb{N}^*}$ are exponentially tight i.e. for any positive L there exists a compact subset K_L of E_1 such that $\limsup_{n \to \infty} (1/n) \ln P(S_n \notin K_L) \leq -L$.
Then the law of S_n satisfies a large deviation principle with rate function

$$\Lambda_E^*(x) \;=\; \sup_{\lambda \in E^*} \big(\lambda(x) - \Lambda_E(\lambda)\big)$$

i.e. for any subset U of E

$$-\inf_{x\in \,\mathrm{interior}(U)} \Lambda^*_E(x) \;\leq\; \liminf_{n\to\infty} \frac{1}{n} \ln P(S_n \in U)$$
$$\leq\; \limsup_{n\to\infty} \frac{1}{n} \ln P(S_n \in U) \;\leq\; -\inf_{x\in \,\mathrm{closure}(U)} \Lambda^*_E(x)$$

Moreover the condition

$$\forall \alpha \in \mathbb{R}^+ \qquad E(\exp(\alpha ||X_1||)) \;<\; +\infty$$

automatically ensures that the laws of $(S_n)_{n\in\mathbb{N}^*}$ are exponentially tight and that Λ^*_E is a "good" rate function (it is lower semicontinuous and its level sets are compact, see either [4, theorem 3.3.11] or [6, exercise 6.2.21]).

Distance to the convex hull. We introduce next a quantity which measures the non–convexity of a set. Let A belong to $\mathcal{K}(F)$, its inner radius is [15]

$$r(A) = \sup_{a\in \,\mathrm{co} A} \inf \{\, R : \exists a_1, \cdots, a_s \in A,\ a \in \mathrm{co}\,\{\, a_1, \cdots, a_s \,\},\ ||a - a_i|| \leq R,\ 1 \leq i \leq s \,\}\,.$$

Obviously, $r(A)$ is zero if and only if A is convex. For any A, $r(A) \leq 2||A|| = 2\sup_{a\in A} ||a||$.

In [11], Puri and Ralescu extended a result of Cassels [2] and proved the following inequality: for any $A_1, \cdots, A_n$ in $\mathcal{K}(F)$

$$d(A_1 + \cdots + A_n, \mathrm{co}\, A_1 + \cdots + \mathrm{co}\, A_n) \;\leq\; c^{1/p} \big(r(A_1)^p + \cdots + r(A_n)^p \big)^{1/p}\,.$$

Of course, the exponent p is related to the fact that F is a Banach space of type p and the constant c is the one appearing in the functional inequality (see the definition just before theorem 1).

Proof of theorem 1. We suppose first that the sets $(A_n)_{n\in\mathbb{N}^*}$ are convex. We apply the general Cramér theorem with $E = \mathcal{C}(B^*, w^*)$, $E_1 = s(\mathrm{co}\,\mathcal{K}(F))$ and the sequence of random functions $(s_{A_n})_{n\in\mathbb{N}^*}$. By the Riesz representation theorem [14], the topological dual of E is the set $M(B^*)$ of the signed Borel measures on (B^*, w^*). By the hypothesis of theorem 1,

$$\forall \alpha \in \mathbb{R}^+ \qquad E(\exp \alpha ||s_{A_1}||_\infty) \;=\; E(\exp(\alpha \sup_{a\in A_1} ||a||)) \;<\; +\infty$$

so that the law of $(s_{A_1} + \cdots + s_{A_n})/n$ satisfies a large deviation principle with rate function Λ^*_E (defined on E). We push back this large deviation principle to the space $\mathrm{co}\,\mathcal{K}(F)$ with the help of the homeomorphism s. Since for any U in $\mathrm{co}\,\mathcal{K}(F)$, $\Lambda^*(U) = \Lambda^*_E(s_U)$ (where

Λ^* is the rate function on $\mathcal{K}(F)$ defined in theorem 1), we obtain that for any $\mathcal{U}$ included in $\mathrm{co}\,\mathcal{K}(F)$

$$-\inf_{U\in \,\mathrm{interior_{co}}\,(\mathcal{U})}\Lambda^*(U)\;\le\;\liminf_{n\to\infty}\frac{1}{n}\ln P(S_n\in\mathcal{U})$$
$$\le\;\limsup_{n\to\infty}\frac{1}{n}\ln P(S_n\in\mathcal{U})\;\le\;-\inf_{U\in\,\mathrm{closure_{co}}\,(\mathcal{U})}\Lambda^*(U)\,,$$

where $\mathrm{interior_{co}}\,(\mathcal{U})$ and $\mathrm{closure_{co}}\,(\mathcal{U})$ are the interior and the closure of $\mathcal{U}$ for the topology induced by the Hausdorff metric on $\mathrm{co}\,\mathcal{K}(F)$.

In the general case, where the sets $(A_n)_{n\in\mathbb{N}^*}$ are not necessarily convex, we set $S_n = (A_1+\cdots+A_n)/n$ and $S_n^{\mathrm{co}} = (\mathrm{co}\,A_1+\cdots+\mathrm{co}\,A_n)/n$. We will use the following lemma.

Lemma 2. *For any $\delta>0$*

$$\lim_{n\to\infty}\frac{1}{n}\ln P(d(S_n,S_n^{\mathrm{co}})\ge\delta)\;=\;-\infty\,.$$

Proof. We apply first the inequality of Puri and Ralescu:

$$P(d(S_n,S_n^{\mathrm{co}})>\delta)\;\le\;P(c^{1/p}(r(A_1)^p+\cdots+r(A_n)^p)^{1/p}\ge n\delta)\,.$$

Notice that this inequality requires the assumption that the space F is of type p.
Let α be a positive real number. We have

$$\begin{aligned}
P(c^{1/p}(r(A_1)^p+\cdots+r(A_n)^p)^{1/p}\ge n\delta) &= P(r(A_1)^p+\cdots+r(A_n)^p)\ge(n\delta)^p/c)\\
&\le P((r(A_1)+\cdots+r(A_n))/n*\max_{1\le k\le n}r(A_k)^{p-1}/n^{p-1}\ge\delta^p/c)\\
&\le P(\max_{1\le k\le n}r(A_k)^{p-1}>\alpha n^{p-1})\;+\;P(r(A_1)+\cdots+r(A_n)\ge n\delta^p/(\alpha c))\\
&\le nP(r(A_1)>\alpha^{1/(p-1)}n)\;+\;P(r(A_1)+\cdots+r(A_n)\ge n\delta^p/(\alpha c))\,.
\end{aligned}$$

Since $r(A_1)\le 2||A_1||=2\sup_{a\in A_1}||a||$, then $\Lambda_r(t)=\ln E\exp tr(A_1)$ is finite for any t in $\mathbb{R}$. By the classical Cramér theorem in $\mathbb{R}$, the sequence $(r(A_1)+\cdots+r(A_n))/n$ satisfies a large deviation principle with rate function $\Lambda_r^*(u)=\sup_t(ut-\Lambda_r(t))$. In addition for any u,t in $\mathbb{R}$, $P(r(A_1)\ge u)\le\exp-(ut-\Lambda_r(t))$ whence $P(r(A_1)\ge u)\le\exp-\Lambda_r^*(u)$. Thus

$$P(d(S_n,S_n^{\mathrm{co}})\ge\delta)\;\le\;n\exp-\Lambda_r^*(\alpha^{1/(p-1)}n)\;+\;\exp-n\Lambda_r^*(\delta^p/(\alpha c))\,.$$

Because Λ_r is finite everywhere we have that $\lim_{u\to+\infty}\Lambda_r^*(u)/u=+\infty$ (see [4] or [6]). It follows that

$$\limsup_{n\to\infty}\frac{1}{n}\ln P(d(S_n,S_n^{\mathrm{co}})\ge\delta)\;\le\;-\Lambda_r^*(\delta^p/(\alpha c))\,,$$

and this is true for any $\alpha > 0$. Letting α go to zero, we obtain the claim. $\square$

We now prove the lower bound for the large deviation principle of theorem 1. Let $\mathcal{U}$ be a subset of $\mathcal{K}(F)$. Let U belong to interior$(\mathcal{U})$ (if interior$(\mathcal{U}) \cap \mathrm{co}\,\mathcal{K}(F)$ is empty, there is nothing to prove). Then there exists $\delta > 0$ such that $\{\, V \in \mathcal{K}(F) : d(U,V) < \delta \,\} \subset \mathcal{U}$. We have then

$$\begin{aligned} P(S_n \in \mathcal{U}) \,\geq\, P(d(S_n,U) < \delta) \,&\geq\, P(d(S_n^{\mathrm{co}},U) < \delta/2,\, d(S_n,S_n^{\mathrm{co}}) < \delta/2) \\ &\geq\, P(d(S_n^{\mathrm{co}},U) < \delta/2) \,-\, P(d(S_n,S_n^{\mathrm{co}}) \geq \delta/2)\,. \end{aligned}$$

Applying lemma 2 and the large deviation principle for $(S_n^{\mathrm{co}})_{n\in\mathbb{N}^*}$ we get

$$\liminf_{n\to\infty} \frac{1}{n} \ln P(S_n \in \mathcal{U}) \,\geq\, -\Lambda^*(U)\,.$$

Taking the supremum over all sets U in interior$(\mathcal{U})$ yields the desired lower bound.

We finally prove the upper bound. Let $\mathcal{U}$ be a subset of $\mathcal{K}(F)$. For any $\delta > 0$ we set $\mathcal{U}^\delta = \{\, A \in \mathcal{K}(F) : d(A,\mathcal{U}) \leq \delta \,\}$. We write then

$$P(S_n \in \mathcal{U}) \,\leq\, P(S_n^{\mathrm{co}} \in \mathcal{U}^\delta) \,+\, P(d(S_n,S_n^{\mathrm{co}}) > \delta)\,.$$

Applying lemma 2 and the large deviation principle for $(S_n^{\mathrm{co}})_{n\in\mathbb{N}^*}$ we get

$$\limsup_{n\to\infty} \frac{1}{n} \ln P(S_n \in \mathcal{U}^\delta) \,\leq\, -\inf\{\, \Lambda^*(U) : U \in \mathrm{closure}_{\mathrm{co}}\,(\mathcal{U}^\delta) \,\}\,.$$

However $\mathrm{closure}_{\mathrm{co}}\,(\mathcal{U}^\delta) = \mathcal{U}^\delta \cap \mathrm{co}\,\mathcal{K}(F)$ and in addition $\bigcap_{\delta>0} \mathrm{closure}_{\mathrm{co}}\,(\mathcal{U}^\delta) = \mathrm{closure}(\mathcal{U}) \cap \mathrm{co}\,\mathcal{K}(F)$. Since Λ^* is a "good" rate function we have that

$$\lim_{\delta\to 0} \inf\{\, \Lambda^*(U) : U \in \mathrm{closure}_{\mathrm{co}}\,(\mathcal{U}^\delta) \,\} \,=\, \inf\{\, \Lambda^*(U) : U \in \mathrm{closure}(\mathcal{U}) \cap \mathrm{co}\,\mathcal{K}(F) \,\}\,.$$

The righthandside is clearly larger than the lefthand side; let $(U_n)_{n\in\mathbb{N}^*}$ be a sequence such that U_n belongs to $\mathrm{closure}_{\mathrm{co}}\,(\mathcal{U}^{1/n})$ for all n and $\Lambda^*(U_n)$ converges to the lefthandside. The level sets of Λ^* being compact, we can extract from $(U_n)_{n\in\mathbb{N}^*}$ a subsequence converging to a set U which necessarily belongs to $\mathrm{closure}(\mathcal{U}) \cap \mathrm{co}\,\mathcal{K}(F)$. By the lower semicontinuity of Λ^*, $\Lambda^*(U)$ is smaller than the lefthand side.

Thus letting δ go to zero in the previous inequality gives the desired upper bound. $\square$

References

1. Z. Artstein and R. A. Vitale, *A strong law of large numbers for random compact sets*, Ann. Prob. **3** (1975), 879–882.
2. J. W. S. Cassels, *Measures of the non–convexity of sets and the Shapley–Folkman–Starr theorem*, Math. Proc. Camb. Phil. Soc. **78** (1975), 433–436.
3. N. Cressie, *A central limit theorem for random sets*, Z. Wahrscheinlichkeitstheor. Verw. Geb. **49** (1979), 37–47.
4. J.D. Deuschel and D.W. Stroock, *Large deviations*, Academic Press, 1989.
5. N. Dunford and J. T. Schwartz, *Linear operators. Part I: General theory*, John Wiley & Sons, 1958.
6. A. Dembo and O. Zeitouni, *Large deviations techniques and applications*, Jones and Bartlett publishers, 1993.
7. E. Giné, M. G. Hahn and J. Zinn, *Limit theorems for random sets: an application of probability in Banach space results*, Proc. Fourth Int. Conf. on Prob. Banach spaces, Oberwolfach (1983), Springer Lecture Notes in Math. 990, 112–135.
8. C. Hess, *Multivalued strong laws of large numbers in the slice topology. Applications to integrands.*, Set–Valued Anal. **2** (1994), 183–205.
9. L. Hörmander, *Sur la fonction d'appui des ensembles convexes dans un espace localement convexe*, Arkiv Matematik **3** (1954), 181–186.
10. N. N. Lyashenko, *On limit theorems for sums of independent compact random subsets of Euclidean space*, Zap. Nauchn. Semin. Leningr. Otd. Mat. Inst. Steklova **85** (1979), 113–128.
11. M. L. Puri and D. A. Ralescu, *Limit theorems for random compact sets in Banach space*, Math. Proc. Camb. Phil. Soc. **97** (1985), 151–158.
12. H. Rådström, *An embedding theorem for spaces of convex sets*, Proc. Amer. Math. Soc. **3** (1952), 165–169.
13. W. Rudin, *Functional Analysis*, McGraw–Hill, 1973.
14. W. Rudin, *Real and complex analysis*, McGraw–Hill, 1966.
15. R. Starr, *Quasi–equilibria in markets with non–convex preferences*, Econometrica **37** (1969), 25–38.
16. W. Weil, *An application of the central limit theorem for Banach–space–valued random variables to the theory of random sets*, Z. Wahrscheinlichkeitstheor. Verw. Geb. **60** (1982), 203–208.

R. Cerf, Université Paris Sud, Mathématique, Bâtiment 425, 91405 Orsay Cedex–France
E-mail address: Raphael.Cerf@math.u-psud.fr