A multivalued strong law of large numbers

Pedro Terán

Departamento de Estadística e I.O. y D.M.

Escuela Politécnica de Ingeniería

Universidad de Oviedo

E-33071 Gijón, Spain

e-mail: teranpedro@uniovi.es

Abstract

We prove a strong law of large numbers for random closed sets in a separable Banach space. It improves upon and unifies the laws of large numbers with convergence in the Wijsman, Mosco and slice topologies, without requiring extra assumptions on either the properties of the space or the kind of sets that can be taken on by the random set as values.

Keywords: Aumann expectation; Gap topology; Random closed set; Slice topology.

1 Introduction

In the last decades, a steady flow of research has been produced about the Strong Law of Large Numbers (SLLN), as well as other limit theorems, for random closed sets. The operations between sets are performed elementwise, taking the closure if needed to ensure that the result remains a closed set. Then the sample averages of a sequence of i.i.d. random sets converge to its Aumann expectation. Applications of this result go back to [2, 21], see also the recent papers [22, 9, 11, 24, 10].

Artstein and Vitale [3] proved the SLLN for random compact sets in $\mathbb{R}^d$, with convergence in the Hausdorff metric. It was extended in two directions: to random compact sets in a separable Banach space by Artstein and Hansen [1] (and later in certain metric spaces by Terán and Molchanov [25]), and to random closed sets in $\mathbb{R}^d$ by Artstein and Hart [2].

It is somewhat harder to deal with both non-compact values and an infinite-dimensional carrier space at the same time. The Hausdorff metric is too strong for the law of large numbers to hold and so must be replaced by an appropriate weaker topology among a plethora of candidates from the literature of hyperspace topology.

Ideally, a topology should be identified which is as fine as possible, has reasonable properties, and allows the SLLN to be proved without extra assumptions on the space. Hiai [19] extended the Artstein-Hart SLLN to separable Banach spaces by using the Mosco topology. It turned out later that this topology misbehaves badly in non-reflexive Banach spaces: it even fails to be Hausdorff. The slice topology defines a convergence for convex sets which coincides with the Mosco topology in reflexive spaces and behaves better in non-reflexive ones; Hess [17] proved an SLLN in that topology, provided that one of the following assumptions holds:

(i) The carrier space has a strongly separable dual.

(ii) Almost surely, every bounded subset of X is relatively weakly compact.

Condition (ii) is quite artificial and, although it is granted in reflexive spaces, that case is already covered by condition (i). In order to prove a law of large numbers for possibly unbounded random sets in a separable Banach space without additional assumptions, Hess [18] resorted to the Wijsman topology. This is formally weaker, but applying it in combination with renorming theorems still yields Mosco and slice convergence in certain conditions (see Section 4 for some more details). In particular, Hess was able to show that the SLLN for the slice topology follows from the SLLN for the Wijsman topology, provided the space has a strongly separable dual (see [18, Proposition 3.9]).

Slice convergence (of convex sets) can always be written as Wijsman convergence over the family of all equivalent norms. It seems, after Hess, that the best approach would be to try and show that a *countable* subfamily suffices, then apply the SLLN for the Wijsman topology in those norms to obtain convergence except on a null set.

But a surprising finding is that the cardinality of the family of norms is not of the essence, since we present a way of reproving the SLLN for the Wijsman topology that will make it clear that the neglected null set can be taken to be *the same for every equivalent norm.* In order to do so, each Wijsman convergence is reduced, using a result of Beer and Tamaki [8], to a family of Painlevé–Kuratowski convergences. This reduction technique is interesting in itself because Painlevé–Kuratowski convergence is the weakest one we have mentioned, making the new proof not only simultaneously valid for all equivalent norms but also arguably simpler than the one in [18].

Finally, we will show that the proof of the non-convex case (commonly called a 'deconvexification' step) is very easy once Wijsman convergence has been established. As a by-product, we can state the SLLN in a topology which, for non-convex sets, is even stronger than the slice topology. That topology is the gap topology generated by the family of all non-empty closed bounded convex sets, henceforth simply called 'the' *gap topology.* The *gap* between two sets A and C is

$$D(A, C) = \inf_{x \in A} \inf_{y \in C} \|x - y\|,$$

and gap topologies are the weak topologies generated by gap functionals. Namely, in our case a net $\{A_i\}_i$ converges to A provided

$$D(A_i, C) \to D(A, C)$$

for all non-empty closed bounded convex C.

Strengthening the convergence in the law of large numbers has practical implications. The failure of the Hausdorff metric shows that reasonable assessments of the behaviour of the sample mean can fail to indicate correctly that of the expectation. For instance, the Hausdorff distance between the sample mean and a set of interest may not converge to the distance between the expectation and that set. Convergence in the Wijsman topology guarantees (almost surely) that every singleton eventually covered by the sample average must be in the expectation as well. The gap topologies address the more realistic question whether it is possible that the expectation is at a positive distance from a set C and yet all sample averages hit C. In applications to stochastic optimization, this translates e.g. as C containing solutions of every member of a sequence of approximating problems while being away from the solutions of the true problem. What we will establish is that this cannot happen when the test set C is bounded and convex.

2   Notation and preliminaries

Let $\mathbb{E}$ denote a separable Banach space and $(\Omega, \mathcal{A}, P)$ a probability space. The dual space of $\mathbb{E}$ is denoted $\mathbb{E}^*$ and its unit sphere S^*. A *random element* of $\mathbb{E}$ is a Borel measurable mapping $\xi : \Omega \rightarrow \mathbb{E}$. Its expectation $E\xi$ is calculated using the Bochner integral.

For any subset A of a metric space with distance d, its *distance function* is given by

$$d(x, A) = \inf_{y \in A} d(x, y)$$

and, for any $r > 0$, we set

$$A^r = \{x \in \mathbb{E} \mid d(x, A) \leq r\}.$$

If $A \subset \mathbb{E}$, then $\overline{\text{co}}\, A, \text{co}\, A, \text{cl}\, A$ denote, respectively, the closed convex hull, convex hull, and closure of A. The *support function* of A is the mapping $s(\cdot, A) : S^* \rightarrow [-\infty, \infty]$ given by $s(f, A) = \sup f(A)$.

Denote by $\mathcal{F}$ the class of all non-empty closed subsets of $\mathbb{E}$. If the additional assumption of convexity is made, we write $\mathcal{F}_c$. A mapping $X : \Omega \rightarrow \mathcal{F}$ is a *random closed set* if it is Effros measurable, namely $\{X \cap G \neq \varnothing\}$ is $\mathcal{A}$-measurable for each open $G \subset \mathbb{E}$. A *selection* of X is a random element ξ such that $\xi \in X$ almost surely.

A random closed set is *integrable* if $Ed(0, X) < \infty$, equivalently if it is has integrable selections. Then its *Aumann expectation* [4] is

$$EX = \text{cl}\{E\xi \mid \xi \text{ is an integrable selection of } X\}.$$

The $\liminf$ of a sequence of closed sets $\{A_n\}_n$ is the set of all limits of sequences $x_n \in A_n$, and the $\limsup$ is the set of all limits of subsequences $x_{n_k} \in A_{n_k}$. The sequence is said to converge to the limit A in the *Painlevé–Kuratowski sense* if

$$\limsup_n A_n = A = \liminf_n A_n.$$

Painlevé–Kuratowski convergence is not necessarily topological, but it becomes so if strengthened by defining the $\limsup$ in the sense of the weak topology instead of the norm topology of $\mathbb{E}$; that yields the *Mosco topology*. The sequence converges in the *Wijsman topology* if $d(x, A_n) \rightarrow d(x, A)$ for all $x \in \mathbb{E}$, and it converges in the *slice topology* if $D(A_n, C) \rightarrow D(A, C)$ for all sets C being a non-empty slice of a ball (the intersection of a ball and a closed halfspace). The *gap topology* defined in the Introduction yields that convergence for all bounded convex C.

3 Main result

In this section, we state and prove our main result.

Theorem 1. *Let $\mathbb{E}$ be a separable Banach space, and let $\{X_n\}_n$ be a sequence of pairwise independent and identically distributed integrable random closed sets in $\mathbb{E}$. Then,*

$$n^{-1} \sum_{i=1}^{n} X_i \to \overline{\mathrm{co}}\, EX$$

almost surely, in the gap topology.

Observe that, if the initial space is non-atomic, then EX is convex and so the limit is EX.

In order to prove this SLLN, we will rely on the SLLN for random *elements* of a Banach space due to Etemadi [15].

Lemma 2. (Etemadi's Strong Law of Large Numbers) *Let $\mathbb{E}$ be a separable Banach space, and let $\{\xi_n\}_n$ be a sequence of pairwise independent and identically distributed random elements of $\mathbb{E}$. Then,*

$$n^{-1} \sum_{i=1}^{n} \xi_i \to E\xi$$

almost surely.

We will also use the fact that random closed sets can be regarded as Borel measurable random variables with values in a metric space [16].

Lemma 3. (Hess's measurability theorem) *Let $\mathbb{E}$ be a separable Banach space. A mapping X is a random closed set if and only if it is measurable with respect to the Borel σ-algebra of the Wijsman topology.*

Let B be the closed unit ball of a fixed arbitrary equivalent norm of $\mathbb{E}$, and $r > 0$. The first step in the proof is to show that

$$n^{-1} \sum_{i=1}^{n} X_i + rB \to \overline{\mathrm{co}}\, EX + rB \tag{1}$$

in the Painlevé-Kuratowski sense, except on a null set independent of both r and the chosen norm.

Proof of $\overline{\mathrm{co}}\, EX + rB \subset \liminf_n(n^{-1}\sum_{i=1}^n X_i + rB)$.

Recall that the expectation of X may depend on the atomic structure of the underlying probability space $(\Omega, \mathcal{A}, P)$, and consider the *reduced expectation* $E^{\sigma(X)}X$, namely the Aumann expectation of X when its distribution is defined on the minimal σ-algebra generated by X. Although it may be the case that $EX \neq E^{\sigma(X)}X$, we have

$$\overline{\mathrm{co}}\, E^{\sigma(X)}X = \overline{\mathrm{co}}\, EX$$

[23, Theorem 1.32(i)].

By the separability of $\mathbb{E}$, there exists a countable dense subset

$$D \subset \{E\xi \mid \xi \text{ integrable } \sigma(X)\text{-measurable selection of } X\}.$$

Since $E^{\sigma(X)}X$ is the closure of the set in the right-hand side, D is dense in $E^{\sigma(X)}X$ too.

Take an arbitrary $x \in D$. By definition, there is a $\sigma(X)$-measurable selection ξ^x such that x is the Bochner expectation $E\xi^x$. By Lemma 3, X is Borel measurable with respect to the Wijsman topology, which is Polish [5]. By the Doob–Dynkin Lemma (e.g. [20, Lemma 1.13]), there exists a Borel function $\phi^x : \mathcal{F} \to \mathbb{E}$ such that $\xi^x(\omega) = \phi^x(X(\omega))$ for each $\omega \in \Omega$.

We are thus able to find a sequence $\{\phi^x \circ X_n\}_n$ which is pairwise independent and identically distributed as ξ^x. Lemma 2, applied to that sequence, yields

$$n^{-1}\sum_{i=1}^n \phi^x(X_n) \to E\xi^x = x$$

except on a null set. Since D is countable, we can dispose of a null subset $N_1 \subset \Omega$ so that, for all $x \in D$,

$$x = \lim_n n^{-1}\sum_{i=1}^n \phi^x(X_n(\omega))$$

for all $\omega \in \Omega \backslash N_1$.

But

$$n^{-1}\sum_{i=1}^n \phi^x(X_n(\omega)) \in n^{-1}\sum_{i=1}^n X_n(\omega),$$

therefore $x \in \liminf_n n^{-1}\sum_{i=1}^n X_i(\omega)$ for $\omega \in \Omega \backslash N_1$.

6

Fix any such ω. We check now that a convex combination $\sum_{j=1}^{m} \lambda_j x_j$ of points $x_j \in D$ is also in the liminf. But this is not hard to show: it reduces to choosing selections from X_n alternately so that the sample proportion $p_{n,j}$ of selections whose expectation is x_j tends to λ_j (e.g. for the nth step take the selection $\xi_n = \phi^{x_{j^*}}(X_n)$ if $|p_{n-1,j^*} - \lambda_{j^*}|$ is the largest distance $|p_{n-1,j} - \lambda_j|$). Thus, as a consequence of the reasoning we made for each x_j, we obtain

$$|n^{-1} \sum_{i=1}^{n} \xi_n(\omega) - x| \leq \max_{1 \leq j \leq m} |p_{n,j} - \lambda_j| \cdot \max_{1 \leq j \leq m} \|x_j\| \to 0.$$

Therefore, for any $y \in rB$,

$$n^{-1} \sum_{i=1}^{n} \xi_n(\omega) + y \to x + y$$

so that

$$\overline{co} \, EX + rB = \text{cl}(co \, D + rB) \subset \liminf_{n} (n^{-1} \sum_{i=1}^{n} X_i + rB)$$

since the right-hand side is closed.

Proof of $\limsup_{n} (n^{-1} \sum_{i=1}^{n} X_i + rB) \subset \overline{co} \, EX + rB.$

By the separation theorem,

$$\overline{co} \, EX = \bigcap_{x^* \in S^*} \{x \mid x^*(x) \leq s(x^*, EX)\}.$$

Letting D' be a countable dense subset of $\{x^* \in S^* \mid s(x^*, EX) < \infty\}$, one still has

$$\overline{co} \, EX = \bigcap_{x^* \in D'} \{x \mid x^*(x) \leq s(x^*, EX)\}$$

(for the case $D' = \emptyset$ we adopt the convention that the intersection over $x^* \in \emptyset$ denotes the whole $\mathbb{E}$.) Therefore,

$$\overline{co} \, EX + rB = \bigcap_{x^* \in D'} \{x \mid x^*(x) \leq s(x^*, EX) + s(x^*, rB)\}.$$

By the linearity of the support function, $s(x^*, EX) = Es(x^*, X)$ for every $x^* \in D'$, whence $s(x^*, X)$ is an integrable random variable for $x^* \in D'$. Let

N_2 be the union of the countably many null sets arising from the application of Etemadi's SLLN to those countably many random variables.

Fix any $\omega \notin N_2$. We must show that every convergent subsequence $n_k^{-1} \sum_{i=1}^{n_k} a_{i,k} +$ $y_k \to x$ such that $a_{i,k} \in X_i(\omega)$, $y_k \in rB$ and $x \in \mathbb{E}$, has $x \in \overline{\mathrm{co}}\, EX + rB$. Reasoning by contradiction, assume the contrary. Then, for some $\varepsilon > 0$,

$$x \notin \overline{\mathrm{co}}\, EX + (r + 2\varepsilon)B.$$

Accordingly, for some $x^* \in D'$, $x^*(n_k^{-1} \sum_{i=1}^{n_k} a_{i,k} + y_k)$ is at least $s(x^*, EX) + s(x^*, (r+\varepsilon)B)$. But

$$x^*(n_k^{-1} \sum_{i=1}^{n_k} a_{i,k} + y_k) \le s(x^*, n_k^{-1} \sum_{i=1}^{n_k} X_i(\omega)) + x^*(y_k)$$

$$= \sum_{i=1}^{n_k} s(x^*, n_k^{-1} X_i(\omega)) + x^*(y_k) \le \sum_{i=1}^{n_k} s(x^*, n_k^{-1} X_i(\omega)) + s(x^*, rB).$$

By Lemma 2, the right-hand side converges to $Es(x^*, X) + s(x^*, rB)$, but then we have

$$s(x^*, EX) + s(x^*, rB) + s(x^*, \varepsilon B) \le \limsup_k x^*(n_k^{-1} \sum_{i=1}^{n_k} a_{i,k} + y_k) \le s(x^*, EX) + s(x^*, rB),$$

a contradiction.

So far, we have proved

$$n^{-1} \sum_{i=1}^{n} X_i + rB \to \overline{\mathrm{co}}\, EX + rB$$

in the Painlevé–Kuratowski sense, except on $N_1 \cup N_2$, for all $r > 0$ and for B being the unit ball of *any* equivalent norm: a key fact is that the choice of N_1 and N_2 has not depended on B.

In order to transform this Painlevé-Kuratowski convergence into a stronger convergence we use the following result, which is a part of [8, Theorem 4.5]. For the reader's benefit, and to emphasize that the extra assumption in the original statement is not used in its proof, we rewrite said proof in the language of this paper.

Lemma 4. (Beer–Tamaki) *Let $\mathbb{E}$ be a metric space. Let $A \in \mathcal{F}$ and $\{A_n\}_n$ be a sequence in $\mathcal{F}$. If $(A_n)^r \to A^r$ in the sense of Painlevé-Kuratowski for every $r > 0$, then $A_n \to A$ in the Wijsman topology.*

Proof. Fix $x \in \mathbb{E}$. For any r, r' with $r > r' > d(x, A)$, we have

$$x \in A^{r'} \subset \liminf_n (A_n)^{r'}.$$

Thus $\{x\}^{r-r'}$ intersects $(A_n)^{r'}$ eventually, whence $d(x, A_n) < r$ eventually.
   On the other hand, if $d(x, A_{n_k}) \leq r$ for a subsequence $\{n_k\}_k$, then

$$x \in \limsup_n (A_n)^r \subset A^r$$

and so $d(x, A) \leq r$. Thus $d(x, A) > r$ ensures that $d(x, A_n) > r$ eventually.
   The combination of both parts yields $d(x, A_n) \to d(x, A)$. $\qquad\square$

   In a Banach space, $A^r = A + rB$, whence the lemma yields the following: the sequence $n^{-1} \sum_{i=1}^n X_i$ converges to $\overline{\mathrm{co}}\, EX$, in all the Wijsman topologies generated by equivalent norms, except on the null set $N_1 \cup N_2$.
   Note that this has been proven for X possibly non-convex, but in order to strengthen this convergence we need to consider the convex case separately.

Lemma 5. (Beer) *The slice topology, the gap topology and the supremum of all Wijsman topologies generated by equivalent norms coincide on $\mathcal{F}_c$.*

Proof. The part 'slice=gap' is [6, Theorem 5.2], whereas 'slice=supremum of Wijsman' is [7, Theorem 3.1]. $\qquad\square$

   Therefore, in the convex case, we achieve convergence in the slice or gap topology. It remains to prove that this convergence extends to the non-convex case. Let A be any non-empty closed bounded convex set. We must prove that

$$D(n^{-1} \sum_{i=1}^n X_i, A) \to D(\overline{\mathrm{co}}\, EX, A)$$

except on a null set simultaneously valid for all such A. Although it might be thought necessary to find a countable family of sets A determining gap convergence, again a simple device shows that the null set $N_1 \cup N_2$ works for an arbitrary A. Since

$$D(\overline{\mathrm{co}}\, EX, A) = \inf_{x \in \overline{\mathrm{co}}\, EX, y \in A} \|x - y\|,$$

for a fixed $\varepsilon > 0$ we find $y \in A$ such that

$$|D(\overline{\mathrm{co}}\,EX, A) - d(y, \overline{\mathrm{co}}\,EX)| \leq \varepsilon.$$

Now we sandwich $D(n^{-1}\sum_{i=1}^{n} X_i, A)$ as follows,

$$D(n^{-1}\sum_{i=1}^{n} \overline{\mathrm{co}}\,X_i(\omega), A) \leq D(n^{-1}\sum_{i=1}^{n} X_i, A) \leq d(y, n^{-1}\sum_{i=1}^{n} X_i(\omega)).$$

For any $\omega \notin N_1 \cup N_2$, by virtue of our intermediate results for the gap and Wijsman topologies, the lower bound converges to $D(E\,\overline{\mathrm{co}}\,X, A)$ and the upper bound converges to $d(y, \overline{\mathrm{co}}\,EX)$. But, since $E\,\overline{\mathrm{co}}\,X = \overline{\mathrm{co}}\,EX$ (see [23, Theorem 1.17(iii), p.154]), both limits differ by at most ε. We conclude that

$$\limsup_{n} |D(n^{-1}\sum_{i=1}^{n} X_i, A) - D(\overline{\mathrm{co}}\,EX, A)| \leq \varepsilon,$$

whence the arbitrariness of ε yields the desired result irrespective of the set A.

4   Additional remarks

1. Hess [18] observed that in reflexive spaces, where the Mosco topology is Hausdorff, the SLLN can be recovered from that with Wijsman convergence by applying it to an equivalent Fréchet differentiable norm. Let us show the stronger result that the SLLN with slice convergence in [17] (under separability of the dual space) is recovered too after a suitable renorming.

A Banach space with separable dual is weakly compactly generated (because it is separable) and Asplund. Thus it admits an equivalent norm whose dual norm has the weak* Kadec–Klee property (see e.g. [14]). And then, by [12, Theorem 3.2.(c)], Wijsman and slice convergence coincide for that norm.

That means that, even if the slice topology is never coarser than the Wijsman and Mosco topologies, in fact its SLLN had only been proved in situations when it was not more powerful: under condition (i) in the Introduction, it is the same as Wijsman convergence for a suitable renorm; and under condition (ii) Hess himself had shown that it is equivalent to Mosco convergence [17].

2. Using the Beer–Tamaki device and the comment above, one can effortlessly prove new versions of some existing almost sure convergence theorems

but with Wijsman convergence (or slice convergence if the dual space is separable) instead of Painlevé–Kuratowski or Mosco convergences. The technique is to apply countably many times a theorem stating $X_i \rightarrow_i X$ to $X_i + rB$ with r ranging over a countable dense subset of $(0, \infty)$.

In some cases, that leads to a weakening of the assumptions of the original theorem. For instance, in the Lévy Martingale Convergence Theorem of Couvreux and Hess [13, Theorem 1], a condition called countable supportedness is used to obtain conclusion (4.3), strong countable supportedness for (4.4), and very strong countable supportedness for (4.5). The conclusions state convergence in the Mosco, Wijsman and slice convergence, respectively. In view of the above, and since $X + rB$ satisfies the assumptions of [13, Theorem 1] as soon as X itself does, (4.4) already holds under countable supportedness, and (4.5) does too if the space has a separable dual. Moreover, in [13, Corollary 2.(b)] (4.3) and (4.4) are obtained, and so (4.5) holds too if $\mathbb{E}^*$ is separable.

However, in some situations (e.g. when the original theorem assumes that the random sets have weakly compact values) the random sets $X_i + rB$ fail the assumptions of the theorem and so this 'on the fly' technique could not be used.

Obtaining slice or gap convergence without the assumption of separable dual is not immediate. To avoid that assumption, it becomes necessary to replicate the steps in the proof of the SLLN leading to the conclusion that the null set involved can be taken the same for every equivalent norm. If that is possible, the way to do it depends on reworking the specific arguments used to prove each theorem.

References

[1] Z. Artstein, J. C. Hansen (1985). Convexification in limit laws of random sets in Banach spaces. *Ann. Probab.* **13**, 307–309.

[2] Z. Artstein, S. Hart (1981). Law of large numbers for random sets and allocation processes. *Math. Oper. Res.* **6**, 485–492.

[3] Z. Artstein, R. Vitale (1975). A strong law of large numbers for random compact sets. *Ann. Probab.* **3**, 879-882.

[4] R. J. Aumann (1965). Integrals of set-valued functions. *J. Math. Anal. Appl.* **12**, 1–12.

[5] G. Beer (1991). A Polish topology for the closed subsets of a Polish space. *Proc. Amer. Math. Soc.* **113**, 1123–1133.

[6] G. Beer (1991). Topologies on closed and closed convex sets and the Effros measurability of set valued functions. Séminaire d'Analyse Convexe de l'Université de Montpellier **21**, Exposé **2**.

[7] G. Beer (1994). Wijsman convergence of convex sets under renorming. *Nonlinear Anal.* **22**, 207–216.

[8] G. Beer, R. K. Tamaki (1993). On hit-and-miss hyperspace topologies. *Comment. Math. Univ. Carolinae* **34**, 717–728.

[9] A. Beresteanu, F. Molinari (2008). Asymptotic properties for a class of partially identified models. *Econometrica* **76**, 763–814.

[10] S. Bobkov, M. Madiman, L. Wang (2011). Fractional generalizations of Young and Brunn-Minkowski inequalities. In: *Concentration, functional inequalities and isoperimetry* (C. Houdré, M. Ledoux, E. Milman, M. Milman, eds.), 35–53. Contemporary Mathematics **545**. American Mathematical Society, Providence.

[11] C. Borell (2009). Zonoids induced by Gauss measure with an application to risk aversion. *Alea* **6**, 133-147.

[12] J. Borwein, J. Vanderwerff (1994). Dual Kadec-Klee norms and the relationships between Wijsman, slice and Mosco convergence. *Michigan Math. J.* **41**, 371–387.

[13] J. Couvreux, C. Hess (1999). A Lévy type martingale convergence theorem for random sets with unbounded values. *J. Theoret. Probab.* **12**, 933–969.

[14] R. Deville, G. Godefroy, V. Zizler (1993). *Smoothness and renormings in Banach spaces.* Longman, Harlow.

[15] N. Etemadi (1981). An elementary proof of the strong law of large numbers. *Z. Wahrsch. Verw. Gebiete* **55**, 119–122.

[16] C. Hess (1991). On multivalued martingales whose values may be unbounded: martingale selectors and Mosco convergence. *J. Multivar. Anal.* **39**, 175–201.

[17] C. Hess (1994). Multivalued strong laws of large numbers in the slice topology. Application to integrands. *Set-Valued Analysis* **2**, 183–205.

[18] C. Hess (1999). The distribution of unbounded random sets and the multivalued strong law of large numbers in nonreflexive Banach spaces. *J. Convex Anal.* **6**, 163-182.

[19] F. Hiai (1985). Convergence of conditional expectations and strong laws of large numbers for multivalued random variables. *Trans. Amer. Math. Soc.* **291**, 613-627.

[20] O. Kallenberg (1997). *Foundations of modern probability.* Springer, New York.

[21] A. J. King, R. J.-B. Wets (1991). Epi-consistency of convex stochastic programs. *Stochastics Stochastics Rep.* **34**, 83-92.

[22] F. Maccheroni, M. Marinacci (2005). A strong law of large numbers for capacities. *Ann. Probab.* **33**, 1171-1178.

[23] I. Molchanov (2005). *Theory of random sets.* Springer, London.

[24] P. Terán (2010). On consistency of stationary points of stochastic optimization problems in a Banach space. *J. Math. Anal. Appl.* **363**, 569–578.

[25] P. Terán, I. Molchanov (2006). The law of large numbers in a metric space with a convex combination operation. *J. Theoret. Probab.* **19**, 875–898.