Convergence of measurable selections and measurable solutions in stochastic optimization

Werner Römisch

1. Introduction

Results about the convergence of measurable multifunctions (random sets) and of their measurable selections are of interest in different fields (see e.g. the volume [9]). For instance, such results have been used for the design and study of approximation schemes in stochastic analysis and stochastic optimization ([5] , [12]). The convergence of measurable selections was first studied in [11] (for recent additions see e.g. [1] , [8]).
In this note, we establish conditions under which sequences (of sets) of measurable selections (of multifunctions with measurable graph and values in Polish spaces) converge almost surely and in probability (referring essentially to [8]). These conditions are related to the respective modes of convergence of the underlying sequence of measurable multifunctions. The main aim of this note is to apply these results to obtain convergence of measurable optimal solutions of stochastic minimization problems. The results obtained are derived from the epi-convergence of the underlying sequence of normal integrands.

Throughout this paper, let $(\Omega,\mathcal{A},P)$ be a complete probability space and X be a Polish space (i.e., complete separable metric) with metric d. Let $\mathcal{P}(X)$ denote the power set of X, $\mathcal{B}(X)$ the σ-algebra of Borel subsets of X and $\mathcal{A}\times\mathcal{B}(X)$ the smallest σ-algebra on $\Omega\times X$ containing $\{A\times B \mid A\in\mathcal{A},\ B\in\mathcal{B}(X)\}$. We write $d(x,F)$ for $\inf\{d(x,y) \mid y\in F\}$, $x\in X$, $F\subseteq X$, and put $d(x,\emptyset) = +\infty$. We say that a property depending on $\omega\in\Omega$ holds almost surely (in abbreviation a.s.) if there is a set $A\in\mathcal{A}$ with $P(A) = 0$ such that the property holds for all $\omega\in\Omega\setminus A$.

For a multifunction C from Ω into X (i.e., a mapping from Ω into $\mathcal{P}(X)$), we define its domain by $\operatorname{dom} C = \{\omega\in\Omega \mid C(\omega)\neq\emptyset\}$, its graph by $\operatorname{Gr} C = \{(\omega,x)\in\Omega\times X \mid x\in C(\omega)\}$ and we put $C^{-1}(F) = \{\omega\in\Omega \mid C(\omega)\cap F\neq\emptyset\}$ for $F\subseteq X$. We say that a multifunction C is closed-valued (nonempty-valued) if $C(\omega)$ is closed for all $\omega\in\Omega$ ($\operatorname{dom} C = \Omega$). C is called Gr-measurable if $\operatorname{Gr} C\in\mathcal{A}\times\mathcal{B}(X)$. It is a well-known conclusion from the projection theorem ([4, Theorem III.23]) that $C^{-1}(B)\in\mathcal{A}$ holds for every $B\in\mathcal{B}(X)$ if C is Gr-measurable.
For a multifunction C from Ω into X we denote by S(C) the set of all measurable $x:\Omega\longrightarrow X$ (i.e., X-valued random variables defined on $(\Omega,\mathcal{A},P)$) that are a.s.-selections of C, i.e., $S(C) = \{x:\Omega\longrightarrow X \mid x \text{ is measurable and } x(\omega)\in C(\omega) \text{ a.s.}\}$.
Consistently, S(X) denotes the set of all X-valued random variables defined on $(\Omega,\mathcal{A},P)$. Note that $S(C)\neq\emptyset$ if C is nonempty-valued and Gr-measurable ([15, Theorem 5.10]).
Excellent monographs about multifunctions, selections, integrands, variational problems and their convergence are [6], [4], [7], [2] and [12] .

2. Convergence of measurable multifunctions and selections

In this section, we review some definitions and results contained essentially in [11] and [8]. Let C and C_n , $n\in\mathbb{N}$, be Gr-measurable multifunctions from Ω into X.

Definition 2.1:
The sequence (C_n) is said to converge to C

(i) almost surely if there is an $A\in\mathcal{A}$ with $P(A) = 0$ such that for all $\omega\in\Omega\setminus A$, $C(\omega)$ is the limit of the sequence $(C_n(\omega))$ in the sense of Kuratowski convergence of sets (see e.g. [11, p. 277] and [2, Sect. 1.4] for an introduction to set-convergence);

(ii) in probability if for all $\varepsilon>0$ and any compact subset K of X, $\lim\limits_{n\to\infty} P(\Delta^{-1}_{\varepsilon,n}(K)) = 0$,
where $\Delta_{\varepsilon,n}(\omega) = (C(\omega)\setminus \varepsilon C_n(\omega)) \cup (C_n(\omega)\setminus \varepsilon C(\omega))$, $\omega\in\Omega$,
and $\varepsilon F = \{x\in X \mid d(x,F)<\varepsilon\}$ denotes the ε-enlargement of

a subset F of X.

(Note that $\Delta_{\varepsilon,n}$ is a Gr-measurable multifunction.)

These modes of convergences of measurable multifunctions are studied extensively in [11] and [12]. Note that (C_n) converges in probability (to C) if it converges almost surely (to C).

Now, let ρ denote a.s.-convergence and convergence in probability ("P-convergence"), respectively, of sequences of X-valued random variables, and let us consider the following limits (with respect to ρ) of the sequence $(S(C_n))$ of sets of measurable a.s.-selections (cf. [5, Def. 2.1], [8]):

$$\rho\text{- Liminf } S(C_n) = \{ x \in S(X) \mid \text{there exist } x_n \in S(C_n),\ n \in \mathbb{N}, \text{ such that } (x_n) \text{ converges to } x \text{ w.r.t. } \rho \},$$

$$\rho\text{- Limsup } S(C_n) = \{ x \in S(X) \mid \text{there exist } n_1 < n_2 < n_3 < \dots \in \mathbb{N} \text{ and } x_k \in S(C_{n_k}),\ k \in \mathbb{N}, \text{ such that } (x_k) \text{ converges to } x \text{ w.r.t. } \rho \},$$

$$\rho\text{- Lim } S(C_n) = \rho\text{- Liminf } S(C_n) \quad \text{if} \quad \rho\text{- Limsup } S(C_n) \subseteq \rho\text{- Liminf } S(C_n).$$

(Note that the inclusion ρ- Liminf $S(C_n) \subseteq \rho$- Limsup $S(C_n)$ holds obviously.)

In the following, we use the notation a.s.- Lim $S(C_n)$, P- Limsup $S(C_n)$ etc. if ρ means a.s.-convergence and P-convergence, respectively.

In our first result, we state conditions that imply

$$S(C) = \rho\text{- Lim } S(C_n),$$

i.e., convergence of sets of measurable selections.

Theorem 2.2:

Let C and C_n, $n \in \mathbb{N}$, be nonempty-valued Gr-measurable multifunctions from Ω into X.

(a) Let (C_n) be almost surely convergent to C. Then we have

$$S(C) = \text{a.s.- Lim } S(C_n).$$

(b) Let C be closed-valued and (C_n) be convergent in probability to C. Then

$$S(C) = \text{P- Lim } S(C_n).$$

For the proof and further informations we refer to [8].

3. An application to the epi-convergence of normal integrands: convergence of optimal solutions

In the following, we make use of the concept of integrands and their epigraph multifunctions developed in [6]. Let $f:\Omega \times X \longrightarrow \overline{\mathbb{R}}$ be an integrand taking values in the extended reals $\overline{\mathbb{R}}$ and let $E_f:\Omega \longrightarrow \mathcal{P}(X \times \mathbb{R})$ be its epigraph multifunction defined by

$$E_f(\omega) = \text{epi } f(\omega,\cdot) = \{ (x,r) \in X \times \mathbb{R} \mid f(\omega,x) \leq r \},\quad \omega \in \Omega.$$

We shall say that f is a normal integrand if f is $\mathcal{A} \times \mathcal{B}(X)$-measurable (this modified definition of normality is suggested in [6, p. 174] for the case that X is not finite-dimensional). Note that the epigraph multifunction E_f is Gr-measurable if f is a normal integrand (see the proof of [6, Theorem 2A]). A normal integran. f will be called tight if for each $\delta > 0$ there is a compact subset K_δ of $X \times \mathbb{R}$ such that $P(E_f^{-1}(K_\delta)) > 1 - \delta$.

Let $m:\Omega \longrightarrow \overline{\mathbb{R}}$ denote the infimal function of f, i.e., $m(\omega) = \inf \{ f(\omega,x) \mid x \in X \}$, $\omega \in \Omega$, and let $M:\Omega \longrightarrow \mathcal{P}(X)$ denote the multifunction of optimal solutions of f, i.e., $M(\omega) = \text{argmin } f(\omega,\cdot) = \{ x \in X \mid f(\omega,x) = m(\omega) \}$, $\omega \in \Omega$.

Lemma 3.1:

(a) If f is a normal integrand, then m is measurable and M is Gr-measurable.

(b) Let $X = \mathbb{R}^m$ and f be a normal integrand such that E_f is nonempty-valued. Then f is tight.

Proof:

(a) The measurability of m follows from [4, Lemma III.39]. Since f is $\mathcal{A} \times \mathcal{B}(X)$-measurable and m is measurable, we have

$$\Omega \times X \setminus \text{Gr } M = \{ (\omega,x) \in \Omega \times X \mid f(\omega,x) > m(\omega) \} =$$
$$= \bigcup_{k \in \mathbb{N}} \{ (\omega,x) \in \Omega \times X \mid f(\omega,x) > r_k \} \cap \{ \omega \in \Omega \mid m(\omega) < r_k \} \times X \in \mathcal{A} \times \mathcal{B}(X),$$

where $\{ r_k \mid k \in \mathbb{N} \}$ is the set of all rational numbers. Hence M is Gr-measurable.

(b) Since $X \times \mathbb{R} = \mathbb{R}^{m+1}$ is σ-compact, there exists a sequence of increasing compact subsets K_n, $n \in \mathbb{N}$, of $\mathbb{R}^{m+1}$ with $\mathbb{R}^{m+1} = \bigcup_{n \in \mathbb{N}} K_n$. This implies $E_f^{-1}(X \times \mathbb{R}) = \bigcup_{n \in \mathbb{N}} E_f^{-1}(K_n)$.

Since E_f is nonempty-valued, we obtain

$$1 = P(E_f^{-1}(X \times \mathbb{R})) = \lim_{n \to \infty} P(E_f^{-1}(K_n)) .$$

This completes the proof. □

Now, let $f_n : \Omega \times X \longrightarrow \overline{\mathbb{R}}$, $n \in \mathbb{N}$, be a collection of normal integrands and let m_n resp. M_n, $n \in \mathbb{N}$, be their infimal functions resp. multifunctions of optimal solutions. Because of Lemma 3.1, m_n is measurable and M_n is Gr-measurable for every $n \in \mathbb{N}$.
In the following, we are interested in conditions guaranteeing convergence of (m_n) and of sequences of measurable optimal solutions, i.e., of measurable selections of M_n, $n \in \mathbb{N}$. As it is known for approximations of deterministic variational problems, epi-convergence implies essentially the convergence of approximate solutions (see [3], [2], [7]). The stochastic counterparts are the following modes of convergence of normal integrands introduced in [12].
A sequence (f_n) of normal integrands is said to epi-converge almost surely (in probability) to the normal integrand f if the sequence (E_{f_n}) of epigraph multifunctions converges almost surely (in probability) to E_f (cf. Def. 2.1).
Note that (f_n) epi-converges almost surely to f if and only if there exists $A \in \mathcal{A}$ with $P(A) = 0$ such that for all $\omega \in \Omega \setminus A$ the sequence $(f_n(\omega, \cdot))$ of mappings from X into $\overline{\mathbb{R}}$ epi-converges to $f(\omega, \cdot)$ ([12, Prop. 3.8]). This argument, together with well-known properties of epi-convergence (see e.g. [2]), yields the following result in a straightforward manner.

Proposition 3.2:
Let the sequence (f_n) of normal integrands epi-converge almost surely to the normal integrand f. Then the following relations hold:

(i) $\limsup_{n \to \infty} m_n(\omega) \leq m(\omega)$ a.s.,

(ii) a.s.- Limsup $S(M_n) \subseteq S(M)$.

Proof:
It suffices to prove (ii). Let $x \in$ a.s.- Limsup $S(M_n)$ and $x_k \in S(M_{n_k})$, $k \in \mathbb{N}$, be such that (x_k) converges almost surely to x, i.e., $\lim_{k \to \infty} x_k(\omega) = x(\omega)$ for P-almost all $\omega \in \Omega$. Since

there exists $A \in \mathcal{A}$ with $P(A) = 0$ such that $(f_n(\omega, \cdot))$ epi-converges to $f(\omega, \cdot)$ for every $\omega \in \Omega \setminus A$, it follows (see e.g. [2, Sect. 2]) that for all $\omega \in \Omega \setminus A$,

$$\text{Limsup } M_n(\omega) = \text{Limsup argmin } f_n(\omega, \cdot) \subseteq M(\omega)$$

(where Limsup is a Kuratowski limit (see [2, Sect. 1.4])).
Since for all $k \in \mathbb{N}$ $x_k(\omega) \in M_{n_k}(\omega)$ a.s., this implies $x(\omega) \in M(\omega)$ a.s. Hence $x \in S(M)$. □

[12, Theorem 8.11] contains a result on the a.s.-convergence of (m_n) to m. Our next result provides an analogue of Prop. 3.2 for the case of epi-convergence in probability (of (f_n)) and extends [8, Theorem 4.2].

Theorem 3.3:
Let the sequence (f_n) of normal integrands epi-converge in probability to the normal integrand f. Assume that E_f and E_{f_n}, $n \in \mathbb{N}$, are nonempty-valued and that f is tight and E_f is closed-valued. Then the following relations hold:

(i) $\lim_{n \to \infty} P(\{\omega \mid m_n(\omega) - m^{\alpha}(\omega) > \varepsilon\}) = 0$ for all $\alpha > 0$ and $\varepsilon > 0$, where $m^{\alpha}(\omega) = \max\{-\alpha^{-1}, m(\omega) + \alpha\}$, $\omega \in \Omega$,

(ii) P- Limsup $S(M_n) \subseteq S(M)$.

Proof:
To prove (i) let $\alpha > 0$, $\varepsilon > 0$ and $\delta > 0$ be arbitrary, but fixed. Let $\omega \in \Omega$ be such that $m_n(\omega) - m^{\alpha}(\omega) > \varepsilon$. Then there exists $(x,r) \in X \times \mathbb{R}$ with $f(\omega,x) \leq m^{\alpha}(\omega) \leq r < m_n(\omega) - \varepsilon$. Hence

$$E_f(\omega) \cap \{(x,r) \in X \times \mathbb{R} \mid r + \varepsilon < f_n(\omega,x)\} \neq \emptyset .$$

The set $\{(x,r) \in X \times \mathbb{R} \mid r + \varepsilon < f_n(\omega,x)\}$ is contained in $X \times \mathbb{R} \setminus \varepsilon E_{f_n}(\omega) = \{(x,r) \in X \times \mathbb{R} \mid d'((x,r), E_{f_n}(\omega)) \geq \varepsilon\}$, where d' is the metric on $X \times \mathbb{R}$ defined by $d'((x,r),(y,s)) = \max\{d(x,y), |r - s|\}$. Thus $E_f(\omega) \setminus \varepsilon E_{f_n}(\omega) \neq \emptyset$. This implies

$$P(\{\omega \mid m_n(\omega) - m^{\alpha}(\omega) > \varepsilon\}) \leq P(\{\omega \mid E_f(\omega) \setminus \varepsilon E_{f_n}(\omega) \neq \emptyset\}) .$$

Since f is tight, there exists a compact subset K_δ of $X \times \mathbb{R}$ such that $P(\{\omega \mid E_f(\omega) \cap K_\delta \neq \emptyset\}) \geq 1 - \frac{\delta}{2}$. This implies

$$P(\{\omega \mid m_n(\omega) - m^{\alpha}(\omega) > \varepsilon\}) \leq P(\{\omega \mid E_f(\omega) \cap K \setminus \varepsilon E_{f_n}(\omega) \neq \emptyset\}) + \frac{\delta}{2}.$$

Since the epi-convergence in probability of (f_n) to f implies

$$\lim_{n\to\infty} P(\{\omega \mid E_f(\omega)\cap K_\delta \setminus \varepsilon E_{f_n}(\omega) \neq \emptyset\}) = 0 ,$$

it follows that $P(\{\omega \mid m_n(\omega) - m^\alpha(\omega) > \varepsilon\}) \leq \delta$ for large n.

(ii) Let $x \in P\text{-}\operatorname{Limsup} S(M_n)$, $n_1 < n_2 < n_3 < \ldots \in \mathbb{N}$ and $x_k \in S(M_{n_k})$, $k \in \mathbb{N}$, be such that (x_k) converges in probability to x. Let $\alpha > 0$ be arbitrary, but fixed. We define $m_n^\alpha(\omega) = \max\{m_n(\omega), m^\alpha(\omega)\}$, for all $\omega \in \Omega$ and $n \in \mathbb{N}$, and note that (i) yields convergence in probability of (m_n^α) to m^α. Hence, the sequence $((x_{n_k}, m_{n_k}^\alpha))$ converges in probability to (x, m^α) and it holds that $(x_{n_k}(\omega), m_{n_k}(\omega)) \in \operatorname{epi} f_{n_k}(\omega,\cdot)$ a.s. Thus $(x_{n_k}, m_{n_k}^\alpha) \in S(E_{f_{n_k}})$ for all $k \in \mathbb{N}$.

This means $(x, m^\alpha) \in P\text{-}\operatorname{Limsup} S(E_{f_n})$. Since (f_n) epi-converges in probability to f and E_f is closed-valued, Theorem 2.2(b) implies $(x, m^\alpha) \in S(E_f)$. Hence we obtain that

$$f(\omega,x(\omega)) \leq m^\alpha(\omega) = \max\{-\alpha^{-1}, m(\omega) + \alpha\} \quad \text{a.s.}$$

Since $\alpha > 0$ was arbitrary, this implies $x \in S(M)$ and the proof is complete. □

Remark 3.4:

Prop. 3.2 and Theorem 3.3 represent stochastic versions of well-known results in the deterministic case (see e.g. [3, Theorem 1], [2, Prop. 2.9]). The results remain true if $M_n(\omega)$ is replaced by ε_n - argmin $f_n(\omega,\cdot) = \{x \in X \mid f_n(\omega,x) \leq \max\{-\varepsilon_n^{-1}, m_n(\omega) + \varepsilon_n\}\}$ for $n \in \mathbb{N}$, $\omega \in \Omega$, for every sequence (ε_n), $\varepsilon_n > 0$, converging to zero. If $X = \mathbb{R}^m$, the assumption that f is tight is superfluous (Lemma 3.1(b)).

Convergence (almost surely and in probability) of multifunctions of optimal solutions in stochastic linear programming was first studied in [14] . For results concerning convergence (in distribution) of the stochastic infima (m_n) to m we refer to [12] and the recent papers [10] and [13] . In [10] and [13] the authors assume that X is locally compact and (f_n) epi-converges in distribution to f (for an introduction to this concept see [12]). Then they give minimal conditions that imply convergence in distribution of (m_n) to m.

References:

[1] Z. Artstein, On dense univalued representations of multi-valued maps, preprint, The Weizmann Institute, 1983.

[2] H. Attouch, Variational Convergence for Functions and Operators, Pitman Adv. Publ. Progr., Boston, 1984.

[3] H. Attouch, R. Wets, Convergence and approximation in non-linear optimization, in: Nonlinear Programming 4, eds. O. Mangasarian, R. Meyer and S. Robinson, Academic Press, New York, 1981, 367 - 394.

[4] C. Castaing, M. Valadier, Convex Analysis and Measurable Multifunctions, Lecture Notes in Mathematics, vol. 580, Springer-Verlag, Berlin, 1977.

[5] H.W. Engl, W. Römisch, Convergence of approximate solutions of non-linear random operator equations with non-unique solutions, Stochastic Anal. and Appl. 1 (1983), 239 - 298.

[6] R.T. Rockafellar, Integral functionals, normal integrands and measurable selections, in: Nonlinear Operators and Calculus of Variations, Lecture Notes in Mathematics, vol. 543, Springer-Verlag, Berlin, 1976, 157 - 207.

[7] R.T. Rockafellar, R. Wets, Variational systems, an introduction, in [9] , 1984, 1 - 54.

[8] W. Römisch, On the convergence of measurable selections and an application to approximations in stochastic optimization, Preprint Nr. 71, Humboldt-Univ. Berlin, Sekt. Mathematik, 1984 and Zeitschr. Anal. Anw. (to appear).

[9] G. Salinetti (ed.), Multifunctions and Integrands, Lecture Notes in Mathematics, vol. 1091, Springer-Verlag, 1984.

[10] G. Salinetti, Convergence of stochastic infima: Equi-semi-continuity, Tech. Report, Universita di Roma, 1985.

[11] G. Salinetti, R. Wets, On the convergence of closed-valued measurable multifunctions, Trans. Amer. Math. Soc. 266 (1981), 275 - 289.

[12] G. Salinetti, R. Wets, On the convergence in distribution of measurable multifunctions, normal integrands, stochastic processes and stochastic infima, Collab. paper, IIASA, Laxenburg, 1982 and Math. Operations Research (to appear).

[13] G. Salinetti, R. Wets, Convergence of infima, especially stochastic infima, manuscript, 1985.

[14] B. Van Cutsem, Problems of convergence in stochastic linear programming, in: Techniques of Optimization, ed. A. Balakrishnan, Academic Press, New York, 1972, 445 - 454.

[15] D.H. Wagner, Survey of measurable selection theorems, SIAM J. Control Optim. 15 (1977), 859 - 903.

Author's address: Humboldt-Universität Berlin
Sektion Mathematik
DDR - 1086 Berlin, PSF 1297