Zeitschrift für Analyiss
und ihre Anwendungen
Bd. 5 (3) 1986, S. 277 – 288

On the Convergence of Measurable Selections and an Application to Approximations in Stochastic Optimization

W. Römisch

Es werden Bedingungen angegeben, die garantieren, daß eine Folge (von Mengen) meßbarer Selektoren fast sicher, in Wahrscheinlichkeit und im Mittel konvergiert. Diese Bedingungen sind mit den entsprechenden Konvergenzeigenschaften der zugrunde liegenden Folge meßbarer Multifunktionen verwandt. Schließlich werden die Resultate zur Untersuchung von Approximationen für das sogenannte Verteilungsproblem der stochastischen Optimierung verwendet.

Формулируются условия, гарантирующие сходимость последовательности (множеств) измеримых селекторов почти наверное, по вероятности и в среднем. Такие условия связаны со соответствующей сходимостью последовательности измеримых многозначных отображений. Результаты применяются к аппроксимации определённой задачи стохастического программирования.

Conditions are given that guarantee that a sequence (of sets) of measurable selections converges almost surely, in probability and in mean. These conditions are related to the convergence of the underlying sequence of measurable multifunctions. The results are applied to approximations for the so-called "distribution problem" of stochastic optimization.

1. Introduction and preliminaries

The study of measurable multifunctions and measurable selections as well as of their convergence is motivated by several applications. These include probability theory [9], stochastic geometry [16], stochastic analysis (e.g. [81]), stochastic optimization [7, 20, 22, 23, 25], control theory, and mathematical economics, among other fields. Particularly, results about the convergence of measurable multifunctions and their measurable selections play an essential role for the design and study of approximation schemes in stochastic analysis and stochastic optimization (see e.g. [8] and [22]). The first results on the convergence of measurable selections seem to be given by Salinetti and Wets in [21] (for finite-dimensional spaces). Probably, [21] initiated the recent research in this field (see [1, 8: Sect. 4, 17, 22]).

In this paper we establish conditions under which sequences (of sets) of measurable selections (of multifunctions with measurable graph and values in Polish spaces) converge almost surely, in probability and in mean. (Convergence in distribution is not considered; this is done in [22] and [1]). These conditions are related to the respective modes of convergence of the underlying sequence of measurable multifunctions. Finally, we outline the use of the results about measurable selection convergence in the study of approximation schemes for the "distribution problem" of stochastic optimization.

Throughout this paper, let $(\Omega, \mathcal{A}, P)$ be a complete probability space (cf. Remark 1.2) and X be a *Polish* space (i.e., complete separable metrizable) with metric d. Let $\mathcal{P}(X)$ be the set of all non-empty subsets of X and $\mathcal{B}(X)$ be the σ-algebra of Borel sets of X. For any $F \subseteq X$ and $x \in X$ let $d(x, F) := \inf \{d(x, y) \mid y \in F\}$, $d(x, F) := +\infty$ if

$F = \emptyset$. The smallest σ-algebra on $\Omega \times X$ containing $\{A \times B \mid A \in \mathcal{A}, B \in \mathcal{B}(X)\}$ will be denoted by $\mathcal{A} \times \mathcal{B}(X)$. As usual, we say that a property depending on $\omega \in \Omega$ holds *almost surely* (a.s.) or *for* P-*almost all* $\omega \in \Omega$ if there is a set $A \in \mathcal{A}$ with $P(A) = 0$ such that the property holds for all $\omega \in \Omega \setminus A$.

A set-valued map from Ω into X is a mapping from Ω into the set of all subsets of X. For a set-valued map C let

$$\operatorname{dom} C := \{\omega \in \Omega \mid C(\omega) \neq \emptyset\}$$

be its domain,

$$\operatorname{Gr} C := \{(\omega, x) \in \Omega \times X \mid x \in C(\omega)\}$$

be its graph and for $B \subseteq X$ let

$$C^{-1}(B) := \{\omega \in \Omega \mid C(\omega) \cap B \neq \emptyset\}.$$

A set-valued map C is called *measurable* (*weakly measurable*) if $C^{-1}(B) \in \mathcal{A}$ for each closed (resp. open) subset B of X. C is called Gr-*measurable* if $\operatorname{Gr} C \in \mathcal{A} \times \mathcal{B}(X)$. If $\operatorname{dom} C = \Omega$, then C is called a *multifunction*. For a multifunction C from Ω into X we denote by $S(C)$ the set of all measurable $x : \Omega \to X$ (X-*valued random variable defined on* $(\Omega, \mathcal{A}, P)$) that are a.s.-*selections* of C, i.e.,

$$S(C) := \{x : \Omega \to X \mid x \text{ measurable and } x(\omega) \in C(\omega) \text{ a.s.}\}.$$

Consistently, $S(X)$ is the set of all X-calued random variables (defined on $(\Omega, \mathcal{A}, P)$).

Excellent sources for properties of measurable set-valued maps and measurable selection theorems (i.e., results stating when $S(C) \neq \emptyset$) are [5, 10, 15], [18] (for $X = \mathbf{R}^m$) and [24]. There the following facts can be found.

Proposition 1.1: *Let C be a set-valued map from Ω into X.*

a) *C is* Gr-*measurable implies that C is measurable, and this implies that C is weakly measurable. If C is closed-valued, then C is* Gr-*measurable iff C is weakly measurable.*

b) *C is weakly measurable iff for all $x \in X$ the map $d(x, C(\cdot))$ from Ω into the extended reals is measurable. If C is weakly measurable, then* $\operatorname{dom} C \in \mathcal{A}$.

c) *If C is* Gr-*measurable, then there exists a measurable map $x : \operatorname{dom} C \to X$ such that $x(\omega) \in C(\omega)$ a.s. If C is a* Gr-*measurable multifunction, then $S(C) \neq \emptyset$ (where "a.s." can be replaced by "for all $\omega \in \Omega$").*

Let $x \in S(X)$ and $x_n \in S(X)$ $(n \in \mathbf{N})$. The following modes of convergence of the sequence $(x_n)_{n \in \mathbf{N}}$ will be considered:

$(x_n)_{n \in \mathbf{N}}$ converges to x

(i) *almost surely* ("*a.s.*-convergence") if there is an $A \in \mathcal{A}$ with $P(A) = 0$ such that for all $\omega \in \Omega \setminus A$,

$$\lim_{n \to \infty} d(x_n(\omega), x(\omega)) = 0;$$

(ii) *in probability* ("P-convergence") if for every $\varepsilon > 0$

$$\lim_{n \to \infty} P(\{\omega \in \Omega \mid d(x_n(\omega), x(\omega)) \geqq \varepsilon\}) = 0;$$

(iii) *in mean* ("m-convergence") if

$$\lim_{n \to \infty} \int_{\Omega} d(x_n(\omega), x(\omega))\, dP = 0.$$

For the following, let ϱ denote any of these modes of convergence in $S(X)$. If $(x_n)_{n\in\mathbf{N}}$ converges to x in the sense of ϱ, we shortly write $x = \varrho - \lim_{n\to\infty} x_n$.

Now, let C and C_n $(n \in \mathbf{N})$ be multifunctions from Ω into X. Let us consider the following limits of the sequence $\big(S(C_n)\big)_{n\in\mathbf{N}}$ of sets of measurable selections (see also [8: Def. 2.1]):

$$\varrho - \operatorname*{Liminf}_{n\to\infty} S(C_n) := \{x \in S(X) \mid x = \varrho - \lim_{n\to\infty} x_n, \quad x_n \in S(C_n), \text{ for all } n \in \mathbf{N}\},$$

$$\varrho - \operatorname*{Limsup}_{n\to\infty} S(C_n) := \{x \in S(X) \mid x = \varrho - \lim_{k\to\infty} x_k, \quad x_k \in S(C_{n_k}), \text{ for all } k \in \mathbf{N} \text{ and for some infinite ordered subset } (n_k)_{k\in\mathbf{N}} \text{ of } \mathbf{N}\},$$

$$\varrho - \operatorname*{Lim}_{n\to\infty} S(C_n) := \varrho - \operatorname*{Liminf}_{n\to\infty} S(C_n) = \varrho - \operatorname*{Limsup}_{n\to\infty} S(C_n).$$

Now, we are in the position to state the aim of this paper as follows: Find conditions that guarantee that

$$S(C) \subseteq \varrho - \operatorname*{Liminf}_{n\to\infty} S(C_n) \tag{1.1}$$

and

$$S(C) = \varrho - \operatorname*{Lim}_{n\to\infty} S(C_n), \tag{1.2}$$

respectively. (We will write a.s. $-\operatorname*{Liminf}_{n\to\infty}$, P $-\operatorname*{Limsup}_{n\to\infty}$ and m $-\operatorname*{Lim}_{n\to\infty}$ etc. in the case of a.s. — convergence, P-convergence and m-convergence, respectively.)

Remark 1.2: We need that the underlying probability space is complete for establishing Prop. 1.1 and Prop. 4.1 in the general setting of this paper. The reader is referred to [18: p. 164/165] for a discussion of „completeness". But, note that Theorem 2.4 (Theorem 3.4) is also valid for complete measure spaces with σ-finite (finite) measure.

2. Almost sure convergence of measurable selections

The study of convergence of sequences of measurable selections has been initiated by Salinetti and Wets in [21]. There, the case of $X = \mathbf{R}^m$ and of closed-valued measurable set-valued maps is studied. In the following let C and $C_n (n \in \mathbf{N})$ be Gr-measurable multifunctions from Ω into a Polish space X.

Definition 2.1: $(C_n)_{n\in\mathbf{N}}$ is said to converge *almost surely* to C if there is an $A \in \mathcal{A}$ with $P(A) = 0$ such that for all $\omega \in \Omega \setminus A$,

$$C(\omega) = \operatorname*{Liminf}_{n\to\infty} C_n(\omega) = \operatorname*{Limsup}_{n\to\infty} C_n(\omega), \tag{2.1}$$

where

$$\operatorname*{Liminf}_{n\to\infty} C_n(\omega) := \{x \in X \mid x = \lim_{n\to\infty} x_n, x_n \in C_n(\omega) \quad \text{for all} \quad n \in \mathbf{N}\}, \tag{2.2}$$

$$\operatorname*{Limsup}_{n\to\infty} C_n(\omega) := \{x \in X \mid x = \lim_{k\to\infty} x_k, x_k \in C_{n_k}(\omega) \quad \text{for all} \quad k \in \mathbf{N}, \tag{2.3}$$

and for some infinite ordered subset $(n_k)_{k\in\mathbf{N}}$ of $\mathbf{N}\}$.

Remark 2.2: Almost sure convergence of (measurable) multifunctions was introduced in [21] (see also [25]). Note that the sets $\operatorname*{Liminf}_{n\to\infty} C_n(\omega)$ and $\operatorname*{Limsup}_{n\to\infty} C_n(\omega)$ $(\omega \in \Omega)$

are clearly closed. If X is locally compact and C is closed-valued, then it follows from [16: p. 10] (see also [21]) that $(C_n)_{n\in\mathbf{N}}$ converges almost surely to C if and only if there exists an $A \in \mathcal{A}$ with $P(A) = 0$ such that for all $x \in X$ and $\omega \in \Omega \setminus A$,

$$\lim_{n\to\infty} d\big(x, C_n(\omega)\big) = d\big(x, C(\omega)\big). \tag{2.4}$$

Now, we are interested in conditions that guarantee (1.1) and (1.2), respectively, for the case of almost sure convergence, and their relations to the notion in Def. 2.1. The next result turns out to be useful for the proof of convergence results for measurable selections.

Lemma 2.3: *Let C_n $(n \in \mathbf{N})$ be* Gr*-measurable multifunctions from Ω into X. For all $x \in S(X)$ there is a sequence $x_n \in S(C_n)$, $n \in \mathbf{N}$, such that for all $n \in \mathbf{N}$ and $\omega \in \Omega$, we have*

$$d\big(x(\omega), x_n(\omega)\big) \leqq d\big(x(\omega), C_n(\omega)\big) + n^{-1}.$$

Proof: Let $x \in S(X)$ and $n \in \mathbf{N}$ be arbitrary, but fixed. We define

$$B(x_0, r) := \{z \in X \mid d(z, x_0) \leqq r\}, \quad \text{for} \quad x_0 \in X, r > 0;$$

$$D_n : \Omega \to \mathcal{P}(X),$$

$$D_n(\omega) := \big\{z \in C_n(\omega) \mid d\big(z, x(\omega)\big) \leqq d\big(x(\omega), C_n(\omega)\big) + n^{-1}\big\}$$

$$= C_n(\omega) \cap B\big(x(\omega), r_n(\omega)\big) \quad \text{for} \quad \omega \in \Omega,$$

where

$$r_n(\omega) := d\big(x(\omega), C_n(\omega)\big) + n^{-1}, \ \omega \in \Omega.$$

Since the map $(\omega, z) \mapsto d\big(z, C_n(\omega)\big)$ is a Caratheodory function from $\Omega \times X$ into $\mathbf{R}$, $d\big(x(\cdot), C_n(\cdot)\big)$ is a real random variable. Because of [5: p. 88] and Prop. 1.1, $B\big(x(\cdot), r_n(\cdot)\big)\colon \Omega \to \mathcal{P}(X)$ is a Gr-measurable closed-valued multifunction. This implies

$$\mathrm{Gr}\, D_n = \mathrm{Gr}\, C_n \cap \mathrm{Gr}\, B\big(x(\cdot), r_n(\cdot)\big) \in \mathcal{A} \times \mathcal{B}(X),$$

i.e., D_n is Gr-measurable. Again using Prop. 1.1 we obtain a measurable map $x_n : \Omega \to X$ such that $x_n(\omega) \in D_n(\omega)$, for all $\omega \in \Omega$ ∎

Theorem 2.4: *Let C and $C_n (n \in \mathbf{N})$ be* Gr*-measurable multifunctions from Ω into X.*

a) *$S(C) \subseteqq$ a.s. $-$ $\underset{n\to\infty}{\mathrm{Liminf}}\, S(C_n)$ if and only if there is an $A \in \mathcal{A}$ with $P(A) = 0$ such that for all $\omega \in \Omega \setminus A$, $C(\omega) \subseteqq \underset{n\to\infty}{\mathrm{Liminf}}\, C_n(\omega)$ (equivalently: $\lim_{n\to\infty} d\big(x, C_n(\omega)\big) = 0$, for all $x \in C(\omega)$).*

b) *Let $(C_n)_{n\in\mathbf{N}}$ be almost surely convergent to C. Then*

$$S(C) = \text{a.s.} - \lim_{n\to\infty} S(C_n).$$

Proof: a) Let $x \in S(C)$ be arbitrary, but fixed. Because of Lemma 2.3 there is a sequence $x_n \in S(C_n)$, $n \in \mathbf{N}$, such that for all $n \in \mathbf{N}$ and $\omega \in \Omega$,

$$d\big(x(\omega), x_n(\omega)\big) \leqq d\big(x(\omega), C_n(\omega)\big) + n^{-1}.$$

This implies

$$\lim_{n\to\infty} d\big(x(\omega), x_n(\omega)\big) = 0 \quad \text{a.s., i.e.} \quad x \in \text{a.s.} - \underset{n\to\infty}{\mathrm{Liminf}}\, S(C_n).$$

Thus, the if-part of assertion a) is proved. For the converse the reader is referred to [8: pp. 271–273].

b) Because of part a), it remains to show that a.s. $-\operatorname{Limsup}_{n\to\infty} S(C_n) \subseteqq S(C)$. Let

$$x \in \text{a.s.} - \operatorname{Limsup}_{n\to\infty} S(C_n), \quad \text{i.e.,} \quad \lim_{k\to\infty} d\big(x(\omega), x_k(\omega)\big) = 0 \quad \text{a.s.},$$

where $x_k \in S(C_{n_k})$ for all $k \in \mathbf{N}$ and $(n_k)_{k\in\mathbf{N}}$ is a subsequence of $\mathbf{N}$. Thus, we have $x(\omega) \in \operatorname{Limsup}_{n\to\infty} C_n(\omega) = C(\omega)$ a.s., i.e., $x \in S(C)$ ∎

Remark 2.5: Theorem 2.4a) is also stated and proved as part a) of Theorem 4.1 in [8]. But, note that the "if-part" is proved using Lemma 2.3. In [8], the selections x_n $(n \in \mathbf{N})$ are constructed in a direct way from Castaing representations of C_n $(n \in \mathbf{N})$. Theorem 2.4b) generalizes Theorem 4.1c) in [8], since almost sure convergence of $(C_n)_{n\in\mathbf{N}}$ to C does not imply (2.4) in arbitrary Polish spaces. It is not clear whether the converse holds in this part of the Theorem! For a discussion of this subject and of other aspects, the reader is referred to [8: Remark 4.3]. Theorem 2.4 can be viewed as a generalization of [21: Theorem 4.3] from $\mathbf{R}^m$ to a Polish space X. Another such generalization is stated as Theorem 1.1 in [17]. There the author assumes the following: $(\Omega, \mathcal{A})$ is a measurable space, X a separable metric space and C, $C_n (n \in \mathbf{N})$ are complete-valued weakly measurable multifunctions. Then he proves a result similar to Theorem 2.4a), but does not permit exceptional sets of measure zero for the selections. Note that in our concept of $S(C)$ the exceptional sets may depend on the selections.

Example 2.6: Let Y be a metric space, $f : \Omega \times X \to Y$ be $\mathcal{A} \times \mathcal{B}(X)$-measurable [8: Def. 1.2] and $B \in \mathcal{B}(Y)$. Let us consider the following set-valued map C from Ω into X

$$\omega \mapsto C(\omega) := \{x \in X \mid f(\omega, x) \in B\}.$$

Clearly we have

$$\operatorname{Gr} C := \{(\omega, x) \in \Omega \times X \mid f(\omega, x) \in B\} = f^{-1}(B) \in \mathcal{A} \times \mathcal{B}(X),$$

i.e., C is Gr-measurable.

Additionally, let $f_n : \Omega \times X \to Y$ $(n \in \mathbf{N})$ be $\mathcal{A} \times \mathcal{B}(X)$-measurable mappings and C_n $(n \in \mathbf{N})$ be defined by

$$C_n(\omega) := \{x \in X \mid f_n(\omega, x) \in B\}, \quad \text{for} \quad \omega \in \Omega.$$

Assume that C and C_n $(n \in \mathbf{N})$ are multifunctions. The following proposition gives sufficient conditions for the a.s. — convergence of $(C_n)_{n\in\mathbf{N}}$ to C.

Proposition 2.7: *Let C and C_n $(n \in \mathbf{N})$ be as in Example 2.6 and assume that there is an $A \in \mathcal{A}$ with $P(A) = 0$ such that for all $\omega \in \Omega \setminus A$,*

(i) *if $x, x_n \in X$ $(n \in \mathbf{N})$ are such that $x = \lim_{n\to\infty} x_n$, then we have*

$$\lim_{n\to\infty} f_n(\omega, x_n) = f(\omega, x),$$

(ii) $C(\omega) = \operatorname{cl}\{x \in X \mid f(\omega, x) \in \operatorname{int} B\}$, *where* "cl" *denotes the closure and* "int" *the interior of a set,*

(iii) *B is closed.*

Then $(C_n)_{n\in\mathbf{N}}$ converges to C almost surely.

The proof is a consequence of [14: Sätze 4.1 and 4.2] applied to $C(\omega)$ and $C_n(\omega)$ $(n \in \mathbf{N})$ for each $\omega \in \Omega \setminus A$. Especially, (i) and (ii) imply

$$C(\omega) \subseteqq \operatorname{Liminf}_{n\to\infty} C_n(\omega), \quad \text{for} \quad \omega \in \Omega \setminus A,$$

and (i) and (iii) imply

$$C(\omega) \subseteq \operatorname*{Limsup}_{n\to\infty} C_n(\omega), \quad \text{for} \quad \omega \in \Omega \setminus A \ \blacksquare$$

Remark 2.8: Note that (2.1) is closely related to the notions "open" ("lower semi-continuous according to Berge") and "closed" in [11, 14, 2] in the context of parametric optimization. Example 2.6 and Prop. 2.7 only serve as an illustration of the general result. Of course, more general Gr-measurable multifunctions (as in [7, 18: p. 171/172]) could be considered. It is well-known that the "constraint qualification" (ii) is rather restrictive. This fact is discussed in [2: Sect. 3.1] and in [13] and it is shown how to overcome this obstacle (in particular, by imposing certain convexity or regularity conditions).

3. Convergence of measurable selections in probability and in mean

As in Section 2, let C and $C_n (n \in \mathbf{N})$ be Gr-measurable multifunctions from Ω into a Polish space X. For later use we define the following set-valued maps from Ω into X for all $n \in \mathbf{N}$ and all $\varepsilon > 0$,

$$\left.\begin{aligned} \triangle^+_{\varepsilon,n}(\omega) &:= C(\omega) \setminus \varepsilon C_n(\omega) \\ \triangle_{\varepsilon,n}(\omega) &:= \triangle^+_{\varepsilon,n}(\omega) \cup \big(C_n(\omega) \setminus \varepsilon C(\omega)\big) \end{aligned}\right\} \quad \text{for all} \quad \omega \in \Omega,$$

where $\varepsilon F := \{x \in X \mid d(x, F) < \varepsilon\}$ for any subset $F \subseteq X$. Note that the set-valued maps $\triangle^+_{\varepsilon,n}$ and $\triangle_{\varepsilon,n} (n \in \mathbf{N}, \varepsilon > 0)$ are clearly Gr-measurable.

Definition 3.1: $(C_n)_{n\in\mathbf{N}}$ is said to converge *in probability* to C if for all $\varepsilon > 0$ and any compact subset $K \subseteq X$, $\lim_{n\to\infty} P\big(\triangle^{-1}_{\varepsilon,n}(K)\big) = 0$.

Convergence in probability of measurable multifunctions was introduced in [21: Sect. 5] and [22: Sect. 1]. In [21] and [22] the usual relations between convergence in probability and almost sure convergence (and convergence in distribution, respectively) are proved for the case $X = \mathbf{R}^m$. The following result establishes one of these relations in a more general situation.

Proposition 3.2: *Let X be a non-compact Polish space. $(C_n)_{n\in\mathbf{N}}$ converges in probability to C if it converges almost surely to C.*

Proof: Let $(C_n)_{n\in\mathbf{N}}$ converge almost surely to C and let $\varepsilon > 0$ and $K \subset X$ compact be arbitrary. Because of Prop. 1.1, there exist measurable mappings $\bar{x}_n : \operatorname{dom}\big(\triangle_{\varepsilon,n}(\cdot) \cap K\big) \to X \ (n \in \mathbf{N})$ such that

$$\bar{x}_n(\omega) \in \triangle_{\varepsilon,n}(\omega) \cap K \quad \text{for all} \quad \omega \in \operatorname{dom}\big(\triangle_{\varepsilon,n}(\cdot) \cap K\big) \setminus N_n,$$

where $N_n \in \mathcal{A}$ with $P(N_n) = 0$. We define

$$N := \bigcup_{n\in\mathbf{N}} N_n, \; A_n := \operatorname{dom}\big(\triangle_{\varepsilon,n}(\cdot) \cup K\big) \setminus N, \quad \text{for all } n,$$

and we note that $P(N) = 0$. For a fixed $y \in X \setminus K$ we define for each $n \in \mathbf{N}$ measurable mappings $x_n : \Omega \to X$,

$$x_n(\omega) := \begin{cases} \bar{x}_n(\omega), & \omega \in A_n \\ y, & \omega \in \Omega \setminus A_n. \end{cases}$$

For all $n \in \mathbf{N}$ we have

$$P\left(\triangle_{\varepsilon,n}^{-1}(K)\right) = P(A_n) = P\left(\{\omega \in \Omega \mid d(x_n(\omega), y) > 0\}\right)$$
$$= P\left(\{\omega \in \Omega \mid d(x_n(\omega), y) \geqq d(y, K)\}\right).$$

Because of the almost sure convergence of $(C_n)_{n\in\mathbf{N}}$ to C, the sequence $(\triangle_{\varepsilon,n})_{n\in\mathbf{N}}$ converges almost surely to $\emptyset$. This implies

$$\lim_{n\to\infty} d\left(x_n(\omega), y\right) = 0 \text{ a.s., and thus, } \lim_{n\to\infty} P\left(\{\omega \in \Omega \mid d(x_n(\omega), y) \geqq d(y, K)\}\right) = 0 \quad \blacksquare$$

Remark 3.3: The following result can be proved analogously to Prop. 3.2: For all $\varepsilon > 0$ and any compact subset $K \subset X$,

$$\lim_{n\to\infty} P\left((\triangle_{\varepsilon,n}^{+})^{-1}(K)\right) = 0$$

if there is an $A \in \mathcal{A}$ with $P(A) = 0$ such that for all $\omega \in \Omega \setminus A$,

$$C(\omega) \subseteqq \operatorname*{Liminf}_{n\to\infty} C_n(\omega).$$

Theorem 3.4: *Let C and $C_n (n \in \mathbf{N})$ be* Gr*-measurable multifunctions from Ω into X.*
a) *$S(C) \subseteqq P - \operatorname{Liminf}_{n\to\infty} S(C_n)$ if for all $\varepsilon > 0$ and any compact subset $K \subseteqq X$, $\lim_{n\to\infty} P\left((\triangle_{\varepsilon,n}^{+})^{-1}(K)\right) = 0$.*
b) *Let C be closed-valued and $(C_n)_{n\in\mathbf{N}}$ converge in probability to C. Then*

$$S(C) = P - \operatorname*{Lim}_{n\to\infty} S(C_n).$$

Proof: a) Let $x \in S(C)$ and let $\varepsilon > 0$ and $\delta > 0$ be arbitrary, but fixed. Because of [4: Theorem 1.4], there is a compact $K \subseteqq X$ such that

$$P\left(x^{-1}(X \setminus K)\right) \leqq \frac{\delta}{2}.$$

By assumption, there is an $n_0 \in \mathbf{N}$ such that for all $n \geqq n_0$,

$$P\left(\{\omega \in \Omega \mid \triangle_{(\varepsilon/2),n}^{+}(\omega) \cap K \neq \emptyset\}\right) \leqq \frac{\delta}{2}.$$

Because of Lemma 2.3, there is a sequence $x_n \in S(C_n)$, $n \in \mathbf{N}$, such that for all $n \in \mathbf{N}$ and $\omega \in \Omega$,

$$d\left(x(\omega), x_n(\omega)\right) \leqq d\left(x(\omega), C_n(\omega)\right) + n^{-1}.$$

Then, there is an $n_1 \in \mathbf{N}$, $n_1 \geqq n_0$, such that for all $n \geqq n_1$,

$$P\left(\{\omega \in \Omega \mid d(x(\omega), x_n(\omega)) \geqq \varepsilon\}\right)$$
$$\leqq P\left(\left\{\omega \in \Omega \mid d\left(x(\omega), C_n(\omega)\right) \geqq \frac{\varepsilon}{2}\right\}\right) = P\left(\{\omega \in \Omega \mid x(\omega) \in \triangle_{(\varepsilon/2),n}^{+}(\omega)\}\right)$$
$$\leqq P\left(\{\omega \in \Omega \mid x(\omega) \in \triangle_{(\varepsilon/2),n}^{+}(\omega) \cap K\}\right) + \frac{\delta}{2} \leqq \delta.$$

Thus, $x \in P - \operatorname{Liminf}_{n\to\infty} S(C_n)$.

b) Because of part a), It remains to show that

$$S(C) \supseteqq P - \operatorname*{Limsup}_{n\to\infty} S(C_n).$$

Let $x \in \operatorname{Limsup}_{n\to\infty} S(C_n)$, i.e., a sequence $(x_k)_{k\in\mathbf{N}}$ converges to x in probability, where $x_k \in S(C_{n_k})$, for all $k \in \mathbf{N}$, and $(n_k)_{k\in\mathbf{N}}$ is a subsequence of $\mathbf{N}$.

Let $\varepsilon > 0$ and $\delta > 0$ be arbitrary, but fixed. Because of Prohorov's Theorem [4: Theorem 6.2], there is a compact subset $K_\delta \subseteqq X$ such that

$$\sup_{k\in\mathbf{N}} P\big(\{\omega \in \Omega \mid x_k(\omega) \in X \setminus K_\delta\}\big) \leqq \frac{\delta}{2}.$$

By assumption, there is a $k_0 \in \mathbf{N}$ such that for all $k \geqq k_0$,

$$P\big(\triangle^{-1}_{\varepsilon,n_k}(K_\delta)\big) \leqq \frac{\delta}{2}.$$

Then we have for all $k \geqq k_0$,

$$\begin{aligned} &P\big(\{\omega \in \Omega \mid d(x_k(\omega), C(\omega)) \geqq \varepsilon\}\big) \\ &\leqq P\big(\{\omega \in \Omega \mid x_k(\omega) \in K_\delta, d(x_k(\omega), C(\omega)) \geqq \varepsilon\}\big) + \frac{\delta}{2} \\ &\leqq P\big(\{\omega \in \Omega \mid (C_{n_k}(\omega) \setminus \varepsilon C(\omega)) \cap K_\delta \neq \emptyset\}\big) + \frac{\delta}{2} \\ &\leqq P\big(\triangle^{-1}_{\varepsilon,n_k}(K_\delta)\big) + \frac{\delta}{2} \leqq \delta. \end{aligned}$$

This means that the sequence $\big(d(x_k(\cdot), C(\cdot))\big)_{k\in\mathbf{N}}$ converges in probability to zero. There exists a subsequence $(x_{k_j})_{j\in\mathbf{N}}$ such that

$$\lim_{j\to\infty} d\big(x_{k_j}(\omega), x(\omega)\big) = 0 \text{ a.s.} \quad \text{and} \quad \lim_{j\to\infty} d\big(x_{k_j}(\omega), C(\omega)\big) = 0 \text{ a.s.}$$

This implies $d\big(x(\omega), C(\omega)\big) = 0$ a.s. Since C is closed-valued by assumption, we obtain $x \in S(C)$ ∎

Let us introduce the following notations:

$$D^+(E, F) := \sup_{x\in E} d(x, F),$$

$$D(E, F) := \max\left\{\sup_{x\in E} d(x, F), \sup_{x\in F} d(x, E)\right\}. \qquad \big(E, F \in \mathcal{P}(X)\big)$$

("Hausdorff-distance")

Corollary 3.5: *Let C and C_n $(n \in \mathbf{N})$ be as in Theorem* 3.4.

a) *$S(C) \subseteqq P - \operatorname{Liminf}_{n\to\infty} S(C_n)$ if for all $\varepsilon > 0$,*

$$\lim_{n\to\infty} P\big(\{\omega \in \Omega \mid D^+(C(\omega), C_n(\omega)) \geqq \varepsilon\}\big) = 0.$$

b) *Let C be closed-valued and assume that for all $\varepsilon > 0$,*

$$\lim_{n\to\infty} P\big(\{\omega \in \Omega \mid D(C(\omega), C_n(\omega)) \geqq \varepsilon\}\big) = 0.$$

Then $S(C) = P - \operatorname{Lim}_{n\to\infty} S(C_n)$.

Proof: Let $n \in \mathbf{N}$, $\varepsilon > 0$ and $K \subseteqq X$ compact be arbitrary. By definition follows

$$\begin{aligned} (\triangle^+_{\varepsilon,n})^{-1}(K) &= \big\{\omega \in \Omega \mid \{x \in C(\omega) \cap K \mid d(x, C_n(\omega)) \geqq \varepsilon\} \neq \emptyset\big\} \\ &\subseteqq \big\{\omega \in \Omega \mid D^+\big(C(\omega), C_n(\omega)\big) \geqq \varepsilon\big\}. \end{aligned}$$

Thus, a) follows from part a) of Theorem 3.4. Analogously,

$$\triangle_{\varepsilon,n}^{-1}(K) \subseteq \{\omega \in \Omega \mid D(C(\omega), C_n(\omega)) \geqq \varepsilon\},$$

and assertion b) also follows from the Theorem ∎

Remark 3.6: It can be seen from the corollary that Theorem 3.4 generalizes [8: Theorem 4.2]. It follows from Prop. 3.2 and Remark 3.3 that the conditions of Theorem 2.4 are stronger than those of Theorem 3.4. Because of [8: Remark 2.2] it is clear that

$$S(C) = \text{a.s.} - \operatorname*{Lim}_{n\to\infty} S(C_n) \quad \text{implies} \quad S(C) = P - \operatorname*{Lim}_{n\to\infty} S(C_n).$$

Definition 3.7: $(C_n)_{n\in\mathbf{N}}$ is said to converge in *mean* to C if

$$\lim_{n\to\infty} \int_{\Omega} D(C_n(\omega), C(\omega))\, dP = 0.$$

This mode of convergence of measurable multifunctions was considered in [23, 9]. Clearly, convergence in mean implies convergence in probability.

Corollary 3.8: *Let C and $C_n (n \in \mathbf{N})$ be as in Theorem 3.4.*

a) $S(C) \subseteq \mathrm{m} = \operatorname*{Liminf}_{n\to\infty} S(C_n)$ *if* $\lim_{n\to\infty} \int_{\Omega} D^+(C_n(\omega), C(\omega))\, dP = 0$.

b) *Let C be closed-valued and assume that $(C_n)_{n\in\mathbf{N}}$ converges in mean to C. Then*

$$S(C) = \mathrm{m} - \operatorname*{Lim}_{n\to\infty} S(C_n).$$

Proof: It is an immediate consequence of a Chebyshev-type inequality and Corollary 3.5 ∎

Theorem 3.4 and its corollaries represent the main results of this paper. Applications to concrete measurable multifunctions seem to be possible if e.g. the results of [2] are used. A simple example can be found in [8: p. 278/279].

4. An application to approximations in stochastic optimization

In this section, we outline the use of measurable selection convergence for the study of approximations for the so-called "distribution problem" of stochastic optimization. This application of our results is only meant to be illustrative.

Let $f : \Omega \times X \to \mathbf{R}$ be $\mathcal{A} \times \mathcal{B}(X)$ — measurable and $C : \Omega \to \mathcal{P}(X)$ be a Gr-measurable multifunction from Ω into X. We introduce the following notations:

$$\overline{\mathbf{R}} := \mathbf{R} \cup \{\pm\infty\},$$

$$\varphi : \Omega \to \overline{\mathbf{R}},\ \varphi(\omega) := \inf\{f(\omega, x) \mid x \in C(\omega)\} \quad \text{for} \quad \omega \in \Omega,$$

$$\psi : \Omega \to X, \qquad \psi(\omega) := \{x \in C(\omega) \mid f(\omega, x) = \varphi(\omega)\} \quad \text{for} \quad \omega \in \Omega.$$

Proposition 4.1: *Let f and C be as above. Then, φ is measurable and ψ is* Gr*-measurable.*

Proof: The measurability of φ follows from [5: Lemma III.39]. The Gr-measurability of ψ follows from

$$\mathrm{Gr}\,\psi = \mathrm{Gr}\,C \cap \{(\omega, x) \in \Omega \times X \mid f(\omega, x) = \varphi(\omega)\} \in \mathcal{A} \times \mathcal{B}(X)$$

by standard arguments and by assumption ∎

The question for the distribution or some of its characteristic values of the "optimal value" φ and (or) of an "optimal solution" (i.e., a selection of ψ) is usually called the *distribution problem* of stochastic optimization (see [6, 12]). Note that this problem makes sense by Prop. 4.1. The reader is referred to numerous results about this subject (e.g. [3, 6, 7, 20, 22, 23]).

Following the approach of [23, 20, 22] we now study approximations for the problem

$$f(\omega, x) \to \text{Min!}\ \text{s.t.}\ x \in C(\omega)\ (\omega \in \Omega); \tag{4.1}$$

Let, additionally, $f_n : \Omega \times X \to \mathbf{R}$ $(n \in \mathbf{N})$ be $\mathcal{A} \times \mathcal{B}(X)$ — measurable and $C_n : \Omega \to \mathcal{P}(X)$ $(n \in \mathbf{N})$ be Gr-measurable multifunctions. We consider a sequence of stochastic optimization problems

$$f_n(\omega, x) \to \text{Min!} \quad \text{s.t.} \quad x \in C_n(\omega)\ (\omega \in \Omega, n \in \mathbf{N}), \tag{4.2}$$

and we define for all $n \in \mathbf{N}$ and $\omega \in \Omega$,

$$\varphi_n(\omega) := \inf \{f_n(\omega, x) \mid x \in C_n(\omega)\},$$

$$\psi_n(\omega) := \{x \in C_n(\omega) \mid f_n(\omega, x) = \varphi_n(\omega)\}.$$

[3] and [23] contain results on the convergence of $(\varphi_n)_{n\in\mathbf{N}}$ and $(\psi_n)_{n\in\mathbf{N}}$ in the case of stochastic linear programming. In [20] and [22: Sect. 8] the convergence of "stochastic infima" (i.e., of $(\varphi_n)_{n\in\mathbf{N}}$) is studied (for the case $X = \mathbf{R}^m$) using the theory of convergence of measurable multifunctions. In the following we will present a result on convergence of the sequence of "optimal solution sets" $\big(S(\psi_n)\big)_{n\in\mathbf{N}}$.

Theorem 4.2: *Let f, $f_n (n \in \mathbf{N})$ be $\mathcal{A} \times \mathcal{B}(X)$ — measurable mappings from $\Omega \times X$ into $\mathbf{R}$ and C, $C_n (n \in \mathbf{N})$ be* Gr*-measurable multifunctions from Ω into X. Assume that*

(i) *C is closed-valued and $(C_n)_{n\in\mathbf{N}}$ converges in probability to C,*

(ii) *for all $x \in S(C)$ and $x_n \in S(C_n)$, $n \in \mathbf{N}$, such that $x = \mathrm{P}\text{-}\lim_{n\to\infty} x_n$, we have*

$$f\big(\cdot, x(\cdot)\big) = \mathrm{P}\text{-}\lim_{n\to\infty} f_n\big(\cdot, x_n(\cdot)\big),$$

(iii) *ψ_n is defined as in (4.2) and is a multifunction from Ω into X, for all $n \in \mathbf{N}$; ψ is as above.*

Then $\mathrm{P}\text{-}\operatorname{Limsup}_{n\to\infty} S(\psi_n) \subseteq S(\psi)$.

Proof: Because of Theorem 3.4, (i) implies

$$S(C) = P\text{-}\operatorname{Lim}_{n\to\infty} S(C_n). \tag{4.3}$$

Let $x \in \mathrm{P}\text{-}\operatorname{Limsup}_{n\to\infty} S(\psi_n)$, i.e., $x = \mathrm{P}\text{-}\lim_{k\to\infty} x_k$, $x_k \in S(\psi_{n_k})$, for all $k \in \mathbf{N}$. Especially, we have $x_k \in S(C_{n_k})$, $k \in \mathbf{N}$, and it follows from (4.3) that $x \in S(C)$. By (4.3) there exists a sequence $\bar{x}_n \in S(C_n)$, $n \in \mathbf{N}$, such that $x = \mathrm{P}\text{-}\lim_{n\to\infty} \bar{x}_n$. For all $n \in \mathbf{N}$, let

$$\tilde{x}_n := \begin{cases} x_k & \text{if} \quad n = n_k \quad k \in \mathbf{N} \\ \bar{x}_n & \text{otherwise.} \end{cases}$$

Then $x = \mathrm{P}\text{-}\lim_{n\to\infty} \tilde{x}_n$. (ii) implies

$$\mathrm{P}\text{-}\lim_{n\to\infty} f_n\big(\cdot, \tilde{x}_n(\cdot)\big) = P\text{-}\lim_{k\to\infty} f_{n_k}\big(\cdot, x_k(\cdot)\big) = \mathrm{P}\text{-}\lim_{k\to\infty} \varphi_{n_k}(\cdot) = f\big(\cdot, x(\cdot)\big).$$

It remains to show that $f\big(\omega, x(\omega)\big) \leqq \varphi(\omega)$ a.s. Let $\varepsilon > 0$ be arbitrary, but fixed. We consider the following set-valued map D : from Ω into X,

$$D(\omega) := \{x \in C(\omega) \mid f(\omega, x) \leqq r_\varepsilon(\omega)\}, \quad \text{for all} \quad \omega \in \Omega,$$

where

$$r_\varepsilon(\omega) := \begin{cases} \varphi(\omega) + \varepsilon & \text{if } \varphi(\omega) > -\infty \\ -\varepsilon^{-1} & \text{otherwise.} \end{cases}$$

It is clear that D is a Gr-measurable multifunction (see also Example 2.6). Let $\hat{x} \in S(D) \subseteq S(C)$ (see Prop. 1.1). By (4.3) there exists a sequence $\hat{x}_n \in S(C_n)$, $n \in \mathbb{N}$, with

$$\mathrm{P} - \lim_{n\to\infty} f_n\big(\cdot, \hat{x}_n(\cdot)\big) = f\big(\cdot, \hat{x}(\cdot)\big).$$

Consequently, there is a further subsequence of $(n_k)_{k\in\mathbb{N}}$, say $(n_{k_j})_{j\in\mathbb{N}}$, such that $(\varphi_{n_{k_j}})_{j\in\mathbb{N}}$ and $\big(f_{n_{k_j}}(\cdot, \hat{x}_{n_{k_j}}(\cdot))\big)_{j\in\mathbb{N}}$ converge almost surely. Thus, we obtain

$$f\big(\omega, x(\omega)\big) = \lim_{j\to\infty} \varphi_{n_{k_j}}(\omega) \leqq f\big(\omega, \hat{x}(\omega)\big) \leqq r_\varepsilon(\omega) \text{ a.s.}$$

Since $\varepsilon > 0$ was arbitrary, this means $f\big(\omega, x(\omega)\big) \leqq \varphi(\omega)$ a.s. and thus, $x \in S(\psi)$ ∎

Remark 4.3: Note that it seems not to be easy to check whether $\mathrm{P} - \operatorname{Limsup}_{n\to\infty} S(\psi_n) \neq \emptyset$. One possibility for doing this is to show that $\big(S(\psi_n)\big)_{n\in\mathbb{N}}$ is contained in a set of X-valued random variables which is compact with respect to convergence in probability. Note that [8: Theorem 4.9] states a criterion for compactness with respect to this mode of convergence.

The study of approximation schemes (4.2) for the original stochastic optimization problem (4.1) is motivated by an approach to solve (4.1) via "discretizing" the random variables involved in (4.1) (see [20], and [19] in a somewhat different context; see also [26] for a recent survey on approximations in stochastic programming).

Acknowledgement: The author wishes to thank Prof. Heinz Engl (Universität Linz, Austria) for valuable and stimulating discussions about the subject of this paper.

REFERENCES

[1] Artstein, Z.: Distributions of random sets and random selections. Israel J. Math. **46** (1983), 313–324.

[2] Bank, B., Guddat, J., Klatte, D., Kummer, B., and K. Tammer: Non-Linear Parametric Optimization. Berlin: Akademie-Verlag 1982.

[3] Bereanu, B.: Stable stochastic linear programs and applications. Math. Operationsforsch. u. Statist. **6** (1975), 593–607.

[4] Billingsley, P.: Convergence of Probability Measures. New York: Wiley 1968.

[5] Castaing, C., and M. Valadier: Convex Analysis and Measurable Multifunctions (Lecture Notes in Mathematics: Vol. 580). Berlin–Heidelberg–New York: Springer-Verlag 1977.

[6] Dempster, M. A. H. (Ed.): Stochastic Programming. London: Academic Press 1980.

[7] Engl, H. W.: Existence of measurable optima in stochastic nonlinear programming and control. Appl. Math. Optim. **5** (1979), 271–281.

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

[9] Hiai, F., and H. Umegaki: Integrals, Conditional Expectations, and Martingales of Multivalued Functions. J. Multivariate Anal. **7** (1977), 149–182.

[10] Himmelberg, C. J.: Measurable relations. Fund. Math. **87** (1975), 53–72.

[11] Hogan, W. W.: Point-to-set maps in mathematical programming. SIAM Review **15** (1973), 591–603.

[12] KALL, P.: Stochastic programming. European J. Operational Research **10** (1982), 125 to 130.
[13] KIRSCH, A.: Continuous perturbations of infinite optimization problems. J. Optimization Theory Appl. **32** (1980), 171—182.
[14] KRABS, W.: Stetige Abänderung der Daten bei nichtlinearer Optimierung und ihre Konsequenzen. Meth. Operations Res. **25** (1976), 93—113.
[15] LEESE, S. J.: Multifunctions of Souslin type. Bull. Austral. Math. Soc. **11** (1974), 395—411.
[16] MATHERON, G.: Random Sets and Integral Geometry. New York: Wiley 1975.
[17] RICCERI, B.: Sur l'approximation des sélections measurables. C.R. Acad. Sc. Paris (Sér. I) **295** (1982), 527—530.
[18] ROCKAFELLAR, R. T.: Integral functionals, normal integrands and measurable selections. In: Nonlinear Operators and Calculus of Variations (Lecture Notes in Mathematics: Vol. 543). Berlin—Heidelberg—New York: Springer-Verlag 1976, 157—207.
[19] RÖMISCH, W.: An approximation method in stochastic optimization and control. Preprint Nr. 14 (1981). Berlin: Sekt. Math. der Humboldt-Univ. In: Mathematical Control Theory. Banach Center Publications (to appear).
[20] SALINETTI, G.: Convergence of measurable multifunctions: an application to stochastic optimization problems. In: Mathematical Programming and its Economic Applications (Eds.: G. Castellani and P. Mazzoleni). Milano: Angeli 1981, 757—774.
[21] SALINETTI, G., and R. WETS: On the convergence of closed-valued measurable multifunctions. Trans. Amer. Math. Soc. **266** (1981), 275—289.
[22] SALINETTI, G., and R. WETS: On the convergence in distribution of measurable multifunctions, normal integrands, stochastic processes and stochastic infima. Mathematics Operations Res. (to appear).
[23] VAN CUTSEM, B.: Problems of convergence in stochastic linear programming. In: Techniques of Optimization (Ed.: A. Balakrishnan). New York: Academic Press 1972, 445—454.
[24] WAGNER, D. H.: Survey of measurable selection theorems. SIAM J. Control Optim. **15** (1977), 859—903.
[25] WETS, R.: On the convergence of random convex sets. In: Convex Analysis and its Applications (Ed.: A. Auslender; Lecture Notes in Economics and Mathematical Systems: Vol. 144). Berlin—Heidelberg—New York: Springer-Verlag 1977, 191—206.
[26] WETS, R.: Stochastic programming: solution techniques and approximation schemes. In: Mathematical Programming: The State-of-the-Art 1982 (Eds.: A. Bachem, M. Grötschel and B. Korte). Berlin: Springer-Verlag 1983, 566—603.

Manuskripteingang: 06. 07. 1984

VERFASSER:

Dr. sc. WERNER RÖMISCH
Sektion Mathematik der Humboldt-Universität
Bereich Numerische Mathematik
DDR-1086 Berlin, PSF 1297