

WORKING PAPER

EPI-CONSISTENCY OF CONVEX STOCHASTIC PROGRAMS

*Alan J. King
Roger J-B Wets*

July 1988
WP-88-57

EPI-CONSISTENCY OF CONVEX STOCHASTIC PROGRAMS

Alan J. King
Roger J-B Wets

July 1988
WP-88-57

Working Papers are interim reports on work of the International Institute for Applied Systems Analysis and have received only limited review. Views or opinions expressed herein do not necessarily represent those of the Institute or of its National Member Organizations.

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

FOREWORD

This paper presents consistency results for sequences of optimal solutions to convex stochastic optimization problems constructed from empirical data. Very few additional assumptions are required because of the special properties of convexity and empirical measures; nevertheless the results are broadly applicable to many situations arising in stochastic programming.

Alexander B. Kurzhanski
Chairman
System and Decision Sciences Program

CONTENTS

| | | |
|---|--|----|
| 1 | Introduction | 1 |
| 2 | Lower Semicontinuity of Integrals | 2 |
| 3 | Epi-Consistency | 5 |
| 4 | Application to Stochastic Recourse Problems in Finite Dimensions | 11 |
| | References | 13 |

EPI-CONSISTENCY OF CONVEX STOCHASTIC PROGRAMS

Alan J. King[†] and Roger J-B Wets[‡]

Abstract. This paper presents consistency results for sequences of optimal solutions to convex stochastic optimization problems constructed from empirical data. Very few additional assumptions are required because of the special properties of convexity and empirical measures; nevertheless the results are broadly applicable to many situations arising in stochastic programming.

Keywords: random sets, epi-convergence, statistical consistency, empirical measures, law of large numbers

[†] International Institute for Applied Systems Analysis, Laxenburg, Austria.

[‡] Department of Mathematics, University of California, Davis, CA 95616.

Research partially supported by a grant of the National Science Foundation

1. Introduction

To solve the stochastic program

$$(1.1) \quad \text{minimize } \int f(x, \xi) P(d\xi) := Ef(x), \text{ over all } x \in X,$$

it is frequently necessary to solve instead an approximating problem,

$$(1.2) \quad \text{minimize } \frac{1}{\nu} \sum f(x, \xi_i) := E^\nu f(x), \text{ over all } x \in X,$$

where the probability measure P is replaced by an empirical measure derived from an independent series of random observations $\{\xi_1, \dots, \xi_\nu\}$ each with common distribution P . Generally speaking, this arises for one of two reasons: either the measure P itself is known only through the observations; or the numerical solution of (1.1) requires the discretization of P , and one very simple technique is to generate a set of “pseudo-random observations” from the distribution of P . Any solution to such a problem, $x^\nu = x^\nu(\xi_1, \dots, \xi_\nu)$, is a random variable that depends on the observations; indeed the objective function itself is random in a certain sense that will be made clear below. As the number ν of sample observations grows large, we demand that the approximations (1.2) approach the true problem in the sense that the functions $E^\nu f$ be *epi-consistent* with limit Ef — that is, $E^\nu f$ *epi-converges* to Ef almost surely. This implies the essential property that cluster points of the sequence of solutions $\{x^\nu\}$ are, with probability one, minimizers of the function Ef . In this paper, we present a systematic investigation of epi-consistency tailored for the special case where the function $f(\cdot, \xi)$ is a.s. convex and the approximating measures are empirical.

Epi-consistency has been explored in Dupačová and Wets [9], and Artstein and Wets [10]. These papers present sufficient conditions on P and f such that $P \mapsto Ef$ is continuous as a map from the space of probability measures topologized by convergence in distribution into the space of lower semi-continuous functions topologized by epi-convergence. Related approximation results delivering local epi-continuity of $P \mapsto Ef$ have been reported in Kall [11] and Robinson and Wets [12]. While it is true that the empirical measures converge in distribution to P , and therefore the present situation can conform to the topological setting of these papers, the empirical/convex case is special and much better results are to be expected.

In this paper, the space X is assumed to be a reflexive Banach space with separable dual X^* ; in particular, X could be a finite-dimensional Euclidean space. Our epi-consistency result applies the strong law of large numbers for sums of random closed sets, as proved by Hess [8], to the epigraphs of the conjugates of the $E^\nu f$. It is remarkable that

an approach similar to the one presented here could be carried out in finite dimensions, based on the strong law of Artstein and Hart [13], but would not yield a better result than that for Banach spaces.

The organization of the paper is as follows. In Section 2, we set the definitions and discuss lower semicontinuity of integrals for not-necessarily-convex integrands. The main result appears in Section 3, accompanied by a brief discussion of epi-convergence. In Section 4, we prove epi-consistency for stochastic linear programs with recourse, under standard assumptions.

The situation considered here has many similarities with maximum likelihood estimation in statistics. In the terminology of that field, $x^\nu = x^\nu(\xi_1, \dots, \xi_\nu)$ is a *statistic* and of the many important properties of a statistic *consistency* stands first, i.e. there is a constant x^* with $x^\nu \rightarrow x^*$ a.s. This concept does not transfer very well to optimization where in many practical situations a unique minimizing x^* for (1.1) is unlikely. For a detailed discussion of these similarities and contrasts, and for a much more complete presentation of the definitions, notations, and motivations than is possible in the confines of the present work, we refer the reader to Dupačová and Wets [9].

2. Lower Semicontinuity of Integrals

The analysis will be based on the geometrical point of view that associates to each extended real-valued function $g : X \rightarrow \overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$ its *epigraph*

$$\text{epi } g = \{(x, \alpha) \in X \times \mathbb{R} \mid \alpha \geq f(x)\}.$$

The function g is called *lower semicontinuous* (l.s.c.) if $\text{epi } g$ is a closed subset of $X \times \mathbb{R}$, this corresponds to

$$(2.1) \quad \liminf_{x' \rightarrow x} g(x') \geq g(x), \quad \forall x \in X,$$

and it is *convex* if $\text{epi } g$ is a convex subset of $X \times \mathbb{R}$.

We next introduce some fundamental measurability concepts for which the standard references are Rockafellar [5], and Castaing and Valadier [6]. Let (Ξ, \mathcal{A}, P) be a probability space completed with respect to P . We say that a closed-valued multifunction $G : \Xi \rightrightarrows X \times \mathbb{R}$ is *measurable* if for all closed subsets $C \subset X$ one has

$$(2.2) \quad G^{-1}(C) := \{\xi \in \Xi \mid G(\xi) \cap C \neq \emptyset\} \in \mathcal{A}.$$

Following usual practice, we shall also call G a *random closed set*. The *domain* of G is the measurable set $\text{dom } G = \{\xi \in \Xi \mid G(\xi) \neq \emptyset\}$.

The epigraphical viewpoint leads to the following definition for the integrand in (1.1). We say that a function $f : X \times \Xi \rightarrow \overline{\mathbb{R}}$ is *random lower semicontinuous* (random l.sc.) if the epigraphical multifunction $\xi \mapsto \text{epi } f(\cdot, \xi)$, where

$$\text{epi } f(\cdot, \xi) = \{(x, \alpha) \in X \times \mathbb{R} \mid \alpha \geq f(x, \xi)\},$$

is a random closed set. Random l.sc. functions were introduced, under the name “normal integrands”, by Rockafellar [5] as a generalization of Caratheodory integrands — functions that are continuous on X and measurable on Ξ . If f is random l.sc. then the *infimal function*

$$(2.3) \quad \xi \mapsto \inf f(\cdot, \xi) := \inf \{f(x, \xi) \mid x \in X\}$$

is measurable, and the *optimal solution multifunction*

$$(2.4) \quad \xi \mapsto \text{argmin } f(\cdot, \xi) := \{x \in X \mid f(x, \xi) = \inf f(\cdot, \xi)\}$$

is a random closed set.

(A random l.sc. function can take the value $+\infty$, permitting the implicit representation of constraints. The set of points x with $f(x, \xi) = +\infty$ is obviously undesirable from the point of view of minimization; thus if in the event $\xi \in \Xi$ we wish to limit our possible decisions to $x \in M(\xi)$, then this can be accomplished by setting $f(x, \xi) = +\infty$ for $x \notin M(\xi)$. The resulting objective is called the *essential objective* and will be a random l.sc. function if M is closed-valued and measurable.)

A random l.sc. function $f : X \times \Xi \rightarrow \overline{\mathbb{R}}$ can be integrated over Ξ , for each $x \in X$, to form the function $Ef : X \rightarrow \overline{\mathbb{R}}$ in the usual way as the difference of the integrals of $f^+(x, \xi) := \max[0, f(x, \xi)]$ and $f^-(x, \xi) := \max[0, -f(x, \xi)]$, once we decide which is the proper value to assign when both the positive and negative parts turn out to be $+\infty$. The natural convention is $(+\infty) - (+\infty) = +\infty$, incorporating the principle that infeasibility $(+\infty)$ dominates. The integral so defined is order preserving and subadditive, i.e. $E(f + g) \leq Ef + Eg$. However, the implicit constraints (the *induced constraints*), defined as the set of points x for which $f(x, \xi) < +\infty$, can be satisfied only with probability one in this definition of the integral. The situation is spelled out in the following result, where, following the traditions of stochastic programming, we define the *weak feasibility set*

$$(2.5) \quad K_2 = \{x \in X \mid f(x, \xi) < +\infty \text{ with probability one}\}.$$

Proposition 2.1. *The weak feasibility set K_2 is closed and can be written in the form*

$$K_2 = \bigcap_{\xi \in \Sigma} \{x \in X \mid f(x, \xi) < +\infty\},$$

where Σ is the measurable set

$$\Sigma = \bigcap_{x \in K_2} \{\xi \in \Xi \mid f(x, \xi) < +\infty\};$$

and, moreover, $P(\Sigma) = 1$.

Proof. See Appendix A of Walkup and Wets [17]. □

From the definition of the integral, minimizing $Ef(x)$ over $x \in X$ is equivalent to minimizing $Ef(x)$ over $x \in K_2$. The above proposition shows that generally this formulation can take into account the implicit constraints $f(x, \xi) < +\infty$ only for the events $\xi \in \Sigma$. A study of the set Σ is beyond the scope of this paper; cf. [17] for more on this subject in the setting of stochastic linear programming.

To obtain lower semi-continuity of the integral Ef we must concern ourselves with the lower boundary of $\text{epi } f(\cdot, \xi)$, and, in particular, when $f(\cdot, \xi)$ approaches $-\infty$. The following proposition is essentially from [19; 5.13].

Proposition 2.2. *Let $f : X \times \Xi \rightarrow \overline{\mathbb{R}}$ be random l.sc. and suppose*

$$(2.6) \quad \int_{\Xi} \inf f(\cdot, \xi) P(d\xi) > -\infty.$$

Then $Ef : X \rightarrow \overline{\mathbb{R}}$ is lower semicontinuous.

Proof. It suffices to show for an arbitrary point $x \in X$ and sequence $\{x^\nu\}$ converging to x that

$$\liminf_{\nu \rightarrow \infty} Ef(x^\nu) \geq Ef(x).$$

We have

$$\liminf_{\nu \rightarrow \infty} Ef(x^\nu) \geq \int_{\Xi} \liminf_{\nu \rightarrow \infty} f(x^\nu, \xi) P(d\xi) \geq Ef(x);$$

the first inequality follows from Fatou's Lemma applied to the nonnegative functions $\xi \mapsto (f(x^\nu, \xi) - \inf f(\cdot, \xi))$ and the second follows from lower semicontinuity of $f(\cdot, \xi)$. □

Remark 2.3. It is generally difficult to weaken condition (2.6). For an example that illustrates the complexities involved, let $\Xi = [0, 1]$ with Lebesgue measure and let $f : [0, 1] \times [0, 1] \rightarrow \overline{\mathbb{R}}$ be defined by

$$f(x, \xi) = \begin{cases} \frac{1}{x}\xi - \frac{1}{\xi} & \text{if } x \neq 0 \text{ and } \xi \neq 0 \\ +\infty & \text{otherwise.} \end{cases}$$

Then for almost all ξ , $f(\cdot, \xi)$ is proper and even l.sc., but $Ef(0) = +\infty$ and $Ef(x) = -\infty$ for $x \neq 0$; thus Ef cannot be l.sc. This example can be modified so that $Ef(x) > -\infty$ for all x (take $\max[f(x, \xi), -\frac{1}{x}]$) without altering the conclusions. However, if X is finite dimensional and $f(\cdot, \xi)$ is *convex*, then condition (2.6) may be weakened, as pointed out in Remark 3.5

3. Epi-Consistency

We suppose that X is a reflexive Banach space with separable dual X^* . Our asymptotic study will be based on the concept of the epi-convergence of sequences of extended real-valued functions on X , and this in turn can be defined via a certain convergence of the sequence of epigraphs as subsets of $X \times \mathbb{R}$. In Proposition 3.1, we see why epi-convergence is important in approximation theory for optimization. We only give the bare outlines here; for more details, see [1], [2], and [3], for example.

Let τ be a topology on a metrizable space Y , and let $\{A_\nu\}$ be a sequence of subsets of Y . We define the following set limits:

$$(3.1) \quad \tau\text{-}\limsup_{\nu} A_\nu = \{y = \tau\text{-}\lim_{\nu} y_\nu \mid y_\nu \in A_\nu \text{ for infinitely many } \nu\},$$

$$(3.2) \quad \tau\text{-}\liminf_{\nu} A_\nu = \{y = \tau\text{-}\lim_{\nu} y_\nu \mid y_\nu \in A_\nu \text{ for all but finitely many } \nu\}.$$

Now let g^ν be a sequence of extended real-valued functions on X . We say that g^ν *Mosco-epi-converges* to g , $\text{Mosco-epi-}\lim_{\nu} g^\nu = g$, if

$$(3.3) \quad \text{epi } g = w\text{-}\limsup(\text{epi } g^\nu) = s\text{-}\liminf(\text{epi } g^\nu),$$

where the $w\text{-}\limsup$ is taken with respect to the product of the weak topology on X and the usual topology on \mathbb{R} , and the $s\text{-}\liminf$ with the product of the strong on X and the usual topology on \mathbb{R} . This type of convergence is neither implied by, nor does it imply, pointwise convergence; cf. [1]. Its superiority to pointwise convergence for applications in optimization theory is evident from the following proposition, where we see that epi-convergence implies that weak cluster points of sequences of points x^ν , each a minimizer of g^ν , must minimize g — an essential property in numerical approximation that is generally false for pointwise convergent sequences of functions.

Proposition 3.1. *If $g = \text{Mosco-epi-}\lim_{\nu} g^\nu$ then*

$$(3.4) \quad w\text{-}\limsup(\text{argmin } g^\nu) \subset \text{argmin } g,$$

and

$$(3.5) \quad \limsup(\inf g^\nu) \leq \inf g.$$

Proof. Cf. [4; Theorems 1 and 3] whose arguments easily generalize to the infinite dimensional setting. \square

Of concern in this study is that cluster points of minimizers to the approximates (1.2) should minimize (1.1). We shall show, in a moment, that the objective functions $E^\nu f$ are random l.sc. on a certain probability space. The epi-convergence, therefore, need only take place on a set of probability one. We formalize this in the following definition.

Definition 3.2. A sequence $\{h^\nu\}$ of random l.sc. functions is *epi-consistent* if there is a (necessarily) l.sc. function h such that

$$(3.6) \quad \text{Mosco-epi-lim } h^\nu = h$$

with probability one.

The main result of this paper provides conditions under which the functions $E^\nu f$ in (1.2) are epi-consistent with limit Ef . The proof employs conjugate duality arguments to arrange things so that a central limit theorem for sums of random closed sets can be applied to the epi-graphs of the conjugates of $E^\nu f$. For this reason the functions are required to be convex. We pause here to review some of the facts about convexity that will be used; these results are standard and may be found in Rockafellar [5], and Castaing and Valadier [6], for example. We shall continue to use the notation established in Section 2.

A random l.sc. function $f : X \times \Xi \rightarrow \overline{\mathbb{R}}$ is said to be *convex* if the epi-graphical multifunction $\xi \mapsto \text{epi } f(\cdot, \xi)$ is closed, convex-valued and measurable. It is *almost surely proper* if for almost all ξ the function $f(\cdot, \xi)$ is proper. The *conjugate* of a random l.sc. convex function f is the mapping $f^* : X^* \times \Xi \rightarrow \overline{\mathbb{R}}$ given by

$$(3.7) \quad f^*(x^*, \xi) = \sup\{\langle x^*, x \rangle - f(x, \xi) \mid x \in X\};$$

the conjugate f^* is a random l.sc. convex function that is a.s. proper if and only if f is a.s. proper. The *subgradient* of a random l.sc. convex function f at a given pair $x \in X$ and $\xi \in \Xi$ is the set

$$(3.8) \quad \partial f(x, \xi) = \{x^* \in X^* \mid f(x', \xi) \geq f(x, \xi) + \langle x^*, x' - x \rangle, \forall x' \in X\}.$$

If f is a.s. proper, then the multifunction $\xi \mapsto \text{gph } \partial f(\cdot, \xi)$ is non-empty, closed-valued and measurable. An important relationship between the subgradient and conjugate is the following: for fixed $\xi \in \Xi$ and $x \in X$, a point x^* belongs to $\partial f(x, \xi)$ if and only if

$$(3.9) \quad \langle x^*, x \rangle = f(x, \xi) + f^*(x^*, \xi).$$

The *continuity set* of a proper l.s.c. convex function $g : X \rightarrow \overline{\mathbb{R}}$ is the set of points $\text{cont}(g)$ where the function is finite and continuous; when $X = \mathbb{R}^n$, this set is equal to the relative interior of $\text{dom } g$ [7].

We will also need the following (epigraphical) operation: by $\overset{+}{\underset{e}{+}}$ we denote the *epi-addition* defined by the identity:

$$\begin{aligned} (f \overset{+}{\underset{e}{+}} g)(x) &= \inf \{ f(u) + g(v) \mid u + v = x \} \\ &= \inf \{ f(u) + g(x - u) \mid u \in X \}; \end{aligned}$$

with, as usual, $\infty - \infty = \infty$. The subscript “e” refers to the fact that the operation takes place on epigraphs. Indeed,

$$\text{epi}_s f + \text{epi}_s g =: \text{epi}_s (f \overset{+}{\underset{e}{+}} g)$$

where $\text{epi}_s h$ is the *strict* epigraph of h , i.e.,

$$\text{epi}_s h = \{(x, \alpha) \mid \alpha > h(x)\}.$$

In the literature one also finds the epi-sum $f \overset{+}{\underset{e}{+}} g$ denoted by $f \square g$ (or $f \nabla g$) and called the *inf-convolution* of f and g . The reference to “convolution” is formal, whereas the epigraphical terminology refers to the geometric interpretation of these operations. The use we make of this concept in the proof of the next lemma should be enough of an illustration.

First, we construct, via Kolmogorov’s method, the *sample probability space*, (Z, \mathcal{Z}) , whose elements are sequences $\zeta = \{\xi_1, \xi_2, \dots\}$, and equip it with a measure μ that is consistent in the sense that if $\pi_\nu : Z \rightarrow \Xi$ is the ν -th coordinate projection and if $A \in \mathcal{A}$, then $\mu \pi_\nu^{-1}(A) = P\{\xi_\nu \in A\}$.

Lemma 3.3. *Let X be a reflexive Banach space and $f : X \times \Xi \rightarrow \overline{\mathbb{R}}$ be a P -a.s. proper random l.s.c. convex function, and suppose that for μ -almost all sequences $\zeta = \{\xi_1, \xi_2, \dots\} \in Z$ one has*

$$(3.10) \quad \bigcap_{i=1}^{\nu} \text{cont}(f(\cdot, \xi_i)) \neq \emptyset, \quad \forall \nu = 1, 2, \dots$$

Then for all $\nu = 1, 2, \dots$ the conjugate functions $(E^\nu f)^*$ are μ -a.s. proper random l.sc. convex functions and, moreover,

$$(3.11) \quad \text{epi}(E^\nu f(\cdot)(\zeta))^* = \frac{1}{\nu} \sum_{i=1}^{\nu} \text{epi} f^*(\cdot, \xi_i).$$

Proof. We have

$$E^\nu f(x)(\zeta) = \frac{1}{\nu} \sum_{i=1}^{\nu} f(x, \pi_i(\zeta)).$$

Usually we write $E^\nu f(x)$, the ζ dependence being implicit. By assumption (3.10), for almost all ζ , the function $E^\nu f(\cdot)(\zeta)$ is proper, convex, and l.sc.; cf. [6; I-21]. The argument of [5; 2M] then applies to show that the $E^\nu f$ are random l.sc. convex functions. Applying the conjugacy formula (3.7) we find

$$(E^\nu f)^*(x^*) = \frac{1}{\nu} \sup \{ \langle \nu x^*, x \rangle - \sum_{i=1}^{\nu} f(x, \xi_i) \mid x \in X \},$$

and hence by [6; I-19] — since $f^{**} = f$ — we have

$$(E^\nu f)^*(x^*) = \frac{1}{\nu} (f^*(\cdot, \xi_1) \dot{+} \cdots \dot{+} f^*(\cdot, \xi_\nu))(\nu x^*).$$

Finally, by [6; I-18 and I-20] and assumption (3.10), we have

$$\text{epi}(f^*(\cdot, \xi_1) \dot{+} \cdots \dot{+} f^*(\cdot, \xi_\nu)) = \sum_i \text{epi}(f^*(\cdot, \xi_i)).$$

The formula (3.11) follows, since $(x^*, \alpha) \in \frac{1}{\nu} \sum \text{epi}(f^*(\cdot, \xi_i))$ if and only if

$$\nu \alpha \geq (f^*(\cdot, \xi_1) \dot{+} \cdots \dot{+} f^*(\cdot, \xi_\nu))(\nu x^*).$$

□

We are ready to state and prove the main result. For the convenience of the reader, we have collected all needed assumptions in the statement of the theorem. Denote by $\|\cdot\|_*$ the norm on the dual X^* .

Theorem 3.4. *Let X be a reflexive Banach space with separable dual X^* , and $f : X \times \Xi \rightarrow \overline{\mathbb{R}}$ be a random l.sc. convex function that satisfies the three conditions:*

(3.12) *there is a point \bar{x} such that $E f(\bar{x})$ is finite, and a measurable selection*

$$\bar{u}(\xi) \in \partial f(\bar{x}, \xi) \text{ such that } \int \|\bar{u}(\xi)\|_* P(d\xi) \text{ is finite;}$$

$$(3.13) \quad Ef(\cdot) \text{ is lower semi-continuous;}$$

$$(3.14) \quad \text{for } \mu\text{-almost all sequences } \zeta = \{\xi_1, \xi_2, \dots\} \text{ one has}$$

$$\bigcap_{i=1}^{\nu} \text{cont}(f(\cdot, \xi_i)) \neq \emptyset, \quad \forall \nu = 1, 2, \dots$$

Then $\{E^\nu f\}$ is epi-consistent with limit Ef , and, with probability one, any weak cluster point of any sequence of minimizers of the $E^\nu f$ is a minimizer of Ef .

Proof. The existence of a measurable selection $\bar{u}(\xi) \in \partial f(\bar{x}, \xi)$ is assured by the first part of (3.12), since $\xi \mapsto \partial f(\bar{x}, \xi)$ is an a.s. nonempty, closed-valued measurable multifunction. Thus the second part of condition (3.12) only requires that among those measurable selections there exist one that is integrable.

The function Ef is l.sc. (assumption (3.13)) and convex (by subadditivity of the integral; cf. [19]) and so are the functions $E^\nu f$ (cf. Lemma 3.3). Thus it suffices to show

$$(3.15) \quad \text{Mosco-epi-} \lim_{\nu \rightarrow \infty} (E^\nu f)^* = (Ef)^*$$

with probability one, since $(Ef)^{**} = Ef$ and $(E^\nu f)^{**} = E^\nu f$ and hence by Mosco's theorem [2], which states that Mosco-epi-convergence of functions implies Mosco-epi-convergence of the conjugates, (3.15) would imply that the $E^\nu f$ are epi-consistent with limit Ef .

From Lemma 3.3 again, we have

$$\text{epi}(E^\nu f(\cdot)(\zeta))^* = \frac{1}{\nu} \sum_{i=1}^{\nu} \text{epi} f^*(\cdot, \xi_i).$$

The random closed sets $\text{epi} f^*(\cdot, \xi_i)$ are independent and identically distributed subsets of the separable reflexive Banach space X^* . We seek to apply Hess's [8; p. 12-34] strong law of large numbers for unbounded random closed sets, which states that:

$$(3.16) \quad \lim_{\nu \rightarrow \infty} \frac{1}{\nu} \sum_{i=1}^{\nu} \text{epi} f^*(\cdot, \xi_i) = \overline{\text{co}} F^*, \quad \mu\text{-a.s.},$$

where $\overline{\text{co}} F^*$ is the closed convex hull of the set

$$F^* = \left\{ \int_{\Xi} (u(\xi), \alpha(\xi)) P(d\xi) \mid (u(\xi), \alpha(\xi)) \text{ is an integrable selection of } \text{epi} f(\cdot, \xi) \right\},$$

provided only that the distance function

$$\xi \mapsto d(0, \text{epi } f^*(\cdot, \xi)) := \inf \{ \|x^*\|_* + |\alpha| : (x^*, \alpha) \in \text{epi } f^*(\cdot, \xi) \}$$

is integrable. This last proviso is implied by our assumption (3.12), since we have

$$d(0, \text{epi } f^*(\cdot, \xi)) \leq \|\bar{u}(\xi)\|_* + |f^*(\bar{u}(\xi), \xi)|;$$

the first term is integrable by the second part of (3.12), and by (3.9) we have

$$f^*(\bar{u}(\xi), \bar{x}) = \langle \bar{u}(\xi), \bar{x} \rangle - f(\bar{x}, \xi)$$

which is integrable by both parts of (3.12). Hence (3.16) is indeed valid. It remains only to show that $\overline{\text{co}} F^* = \text{epi}(Ef)^*$ or, equivalently, that $\overline{\text{co}} F^*$ is the epi-graph of an l.sc. convex function and

$$(3.17) \quad Ef(x) = \sup \{ \langle x^*, x \rangle - \alpha \mid (x^*, \alpha) \in \overline{\text{co}} F^* \}.$$

Evidently $\overline{\text{co}} F^*$ is a closed convex subset of $X^* \times \mathbb{R}$; that it is an epigraph is also clear. Hence $\overline{\text{co}} F^*$ is the epigraph of some l.sc. convex function. In (3.17), note that the supremum is unaffected if we replace $\overline{\text{co}} F^*$ by F^* , and it is thus equal to

$$(3.18) \quad \sup_{(u(\cdot), \alpha(\cdot)) \in \mathcal{L}^1} \left\{ \int_{\Xi} [\langle u(\xi), x \rangle - \alpha(\xi) - \psi(u(\xi), \alpha(\xi), \xi)] P(d\xi) \right\},$$

where ψ is the a.s. proper random l.sc. convex function

$$\psi(u, \alpha, \xi) = \begin{cases} 0 & \text{if } (u, \alpha) \in \text{epi } f^*(\cdot, \xi) \\ +\infty & \text{otherwise,} \end{cases}$$

and \mathcal{L}^1 is the space of P -integrable functions from Ξ into $X^* \times \mathbb{R}$. Since \mathcal{L}^1 is decomposable, we may exchange supremum and integration in (3.18), cf. [6; VII-14] or [5; 3A] for example, and obtain

$$\sup \{ \langle x^*, x \rangle - \alpha \mid (x^*, \alpha) \in \overline{\text{co}} F^* \} = \int_{\Xi} \sup_{x^* \in X^*} \{ \langle x^*, x \rangle - f^*(x^*, \xi) \} P(d\xi).$$

The right hand side is evidently equal to $Ef(x)$; hence (3.17) is proved. This verifies epi-consistency. The final conclusion, concerning cluster points of minimizers, is evident from epi-consistency and Proposition 3.1; see also [9]. \square

Remark 3.5. Assumption (3.13), of the lower semi-continuity of Ef , can be proved in several ways. Proposition 2.2 gives one possibility. If X is finite dimensional and f is a random l.sc. convex function, it is shown in [18] that one can get by with a weaker condition, namely

$$Ef(x) > -\infty, \quad \forall x \in X.$$

4. Application to Stochastic Recourse Problems in Finite Dimensions

In this section, we show how the conditions of the epi-consistency theorem may be satisfied in the important class of *two-stage stochastic linear programs with recourse*:

$$(4.1) \quad \begin{aligned} & \text{minimize} && c'x + \int Q(x, \xi)P(d\xi) && \text{over all } x \in \mathbb{R}^n \\ & \text{subject to} && Ax = b, \\ & && x \geq 0, \end{aligned}$$

where the function $Q : \mathbb{R}^n \times \Xi \rightarrow \overline{\mathbb{R}}$ is the minimum value in the *second stage linear program*

$$(4.2) \quad Q(x, \xi) = \inf\{q'y \mid Wy = Tx - h, y \in \mathbb{R}^m, y \geq 0\},$$

and where c' denotes the transpose of c . We regard ξ as the random vector consisting of the vectors and matrices in the second stage program, i.e. $\xi = (q, W, T, h)$. This class of problems models decisions that must take into account future costs $Q(x, \xi)$, represented as linear programs, responding to presently uncertain events $\xi \in \Xi$, distributed according to P ; see, for example, [14], [15] and [16]. As in the introduction, we suppose that (4.1) cannot be solved as stated, because either P is not known or must be made discrete. Instead, one solves the problems

$$(4.3) \quad \begin{aligned} & \text{minimize} && c'x + \frac{1}{\nu} \sum_{i=1}^{\nu} Q(x, \xi_i) && \text{over all } x \in \mathbb{R}^n \\ & \text{subject to} && Ax = b \\ & && x \geq 0, \end{aligned}$$

where the ξ_i are independent random variables with common distribution P . We shall show that the essential objectives of (4.3) are epi-consistent with limit equal to the essential objective (4.1), under assumptions that are standard in the stochastic programming literature.

A comprehensive study of the properties of $Q(x, \xi)$ appears in the papers of Walkup and Wets [17] and [18], and Wets [20]. Most of what follows is drawn from these papers. For convenience, let us denote by K_1 the set of x satisfying the constraints of (4.1), i.e.

$$(4.4) \quad K_1 = \{x \in \mathbb{R}^n \mid Ax = b, x \geq 0\}.$$

We make here the blanket assumptions that the matrix W is *fixed*, i.e. nonrandom, and that for every $x \in K_1$ the second stage problem is feasible a.s., i.e., $K_1 \subset K_2$ where K_2 is the w.p.1 feasibility set

$$(4.5) \quad K_2 = \{x \in \mathbb{R}^n \mid Q(x, \xi) < +\infty \text{ with probability one}\}.$$

When these two assumptions are satisfied, the problem (4.1) is called a stochastic linear program with *fixed, relatively complete recourse*. Finally, we shall also assume that the random vector ξ satisfies the *weak covariance condition*:

$$(4.6) \quad \text{For all } i, j, k \text{ the random variables } q_i h_j \text{ and } q_i T_{jk} \text{ have finite first moments.}$$

This assumption is obviously satisfied if ξ is square integrable.

Let us now define the *essential integrand* as follows:

$$f(x, \xi) = c'x + Q(x, \xi) + \delta_{K_1}(x)$$

where $\delta_{K_1}(x) = +\infty$ if x is not in K_1 and zero otherwise. Clearly, the essential objective of the problem (4.1) is Ef , and that of the estimated problem (4.3) is $E^\nu f$. The essential objectives for the estimated problems are therefore epi-consistent with limit equal to the essential objective of the original problem, by Theorem 3.4, if f is a random l.sc. convex function that satisfies (3.12–14). We present the results of our investigations in a single theorem with a single set of assumptions. Stronger partial results are obviously possible; these may be found in the citations.

Theorem 4.1. *Suppose that the stochastic linear program (4.1) has fixed, relatively complete recourse and that the random elements satisfy the weak covariance condition (4.6). If there exists a single point $\bar{x} \in K_1$ with $EQ(\bar{x}) > -\infty$, then the $E^\nu f$ are epi-consistent with limit Ef , and, with probability one, all cluster points of sequences of minimizers to the problems (4.3) are minimizers of the original problem (4.1).*

Proof. The essential integrand f can be written as the sum of Q and the convex lower semi-continuous function $c'x + \delta_{K_1}(x)$, that does not depend on ξ . Hence, f is random l.sc. convex if and only if Q is, and the function Q is random l.sc. by a standard result in measurability of multifunctions, e.g. [5; 2R], and $Q(\cdot, \xi)$ is convex by [17; 4.3]. We next show (3.13). The assumptions of fixed recourse and weak covariance imply that EQ is either identically $-\infty$ or finite and Lipschitz on K_1 , by [20; 7.6]. Our assumption of the existence of $\bar{x} \in K_1$ with $EQ(\bar{x}) > -\infty$ implies that the latter is true. Hence EQ is in particular lower semicontinuous, and therefore so is Ef . It remains only to prove (3.12) and (3.14). The functions $f(\cdot, \xi)$, $\xi \in \Xi$, are continuous on the relative interiors of their domains [7], and the assumption of relatively complete recourse implies that $\text{ri}(\text{dom } f(\cdot, \xi)) = \text{ri } K_1$ for almost all $\xi \in \Xi$ (cf. Proposition 2.1); hence (3.14) is satisfied. To establish (3.12), let $\bar{u}(\xi)$ be a selection from $\partial f(\bar{x}, \xi)$. By convex analysis [7; 29.1, 30.5] we have $\bar{u}(\xi) = \bar{y}(\xi)'T + \bar{a}$, where \bar{a} equals c' plus a fixed element from the

normal cone to the constraint set K_1 at \bar{x} , and where $\bar{y}(\xi)$ is the solution to the dual of (4.2):

$$(4.7) \quad \begin{aligned} & \text{maximize} && y'(T\bar{x} - h) \\ & \text{subject to} && y'W \leq q. \end{aligned}$$

It follows that $\bar{y}(\xi) = B^{-1}q'$, where B is some invertible square submatrix of W' ; hence

$$\bar{u}(\xi) = q(B^{-1})'T + \bar{a}$$

and this is integrable by the weak covariance assumption. □

References.

- [1] G. Salinetti and R.J-B Wets, "On the relation between two types of convergence for convex functions", *J. Mathematical Analysis and its Applications* **60**(1977), 211–226.
- [2] U. Mosco, "Convergence of convex sets and of solutions to variational inequalities", *Advances in Mathematics* **3**(1969), 510–585.
- [3] S. Dolecki, G. Salinetti, and R.J-B Wets, "Convergence of functions: equi-semicontinuity", *Transactions of the American Mathematical Society* **276**(1983), 409–429.
- [4] H. Attouch and R.J-B Wets, "Approximation and convergence in nonlinear optimization", in *Nonlinear Programming 4*, Academic Press, 1981, pp. 367–394.
- [5] R.T. Rockafellar, "Integral functionals, normal integrands and measurable selections", in *Nonlinear Operators and the Calculus of Variations*, Springer-Verlag Lecture Notes in Math. No. 543, 1976, pp. 157–207.
- [6] C. Castaing and M. Valadier, *Convex Analysis and Measurable Multifunctions*, Springer-Verlag Lecture Notes in Math. No. 580, 1977.
- [7] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, 1970.
- [8] C. Hess, *Contribution a l'Etude de la Mesaurabilité, de la Loi de Probabilité et de la Convergence de Multifunctions*, Thèse d'Etat, Université des Sciences et Techniques du Languedoc, 1986.
- [9] J. Dupačová and R.J-B Wets, "Asymptotic behavior of statistical estimators and of optimal solutions of stochastic optimization problems", *Annals of Mathematical Statistics* (to appear).
- [10] Z. Artstein and R.J-B Wets, "Approximating the integral of multifunction", *J. Multivariate Analysis* **24**(1988), 285–308.
- [11] P. Kall, "On approximations and stability in stochastic programming", in *Parametric Optimization and Related Topics*, J. Guddat, et al, eds., Akademie-Verlag, Berlin, 1987, pp. 387–407.
- [12] S.M. Robinson and R.J-B Wets, "Stability in two-stage stochastic programming", *SIAM J. Control and Optimization* **25**(1987), 1409–1416.

- [13] Z. Artstein and S. Hart, "Law of large numbers for random sets and allocation processes", *Mathematics of Operations Research* **6**(1981), 485–492.
- [14] M.A.H. Dempster, "Introduction to stochastic programming", in *Stochastic Programming*, M.A.H. Dempster, ed., Oxford Univ. Press, 1980.
- [15] P. Kall, *Stochastic Linear Programming*, Springer-Verlag, 1972.
- [16] R.J-B Wets, "Stochastic programming: solution techniques and approximation schemes", in *Mathematical Programming: The State of the Art 1982*, A.Bachem, M.Groetschel, and B.Korte, eds., Springer-Verlag, 1983, pp. 566–603.
- [17] D. Walkup and R.J-B Wets, "Stochastic programs with recourse", *SIAM J. Applied Mathematics* **15**(1967), 1299–1314.
- [18] D. Walkup and R.J-B Wets, "Stochastic programs with recourse II: on the continuity of the objective", *SIAM J. Applied Mathematics* **17**(1969), 98-103.
- [19] R.M. Van Slyke and R.J-B Wets. "Stochastic programs in abstract spaces", in *Stochastic Optimization and Control*, H.F. Karreman, ed., Wiley, 1968, pp. 25–45.
- [20] R. J-B Wets, "Stochastic programs with fixed recourse: the equivalent deterministic program", *SIAM Review* **16**(1974), 309–339.