
STOCHASTIC PROGRAMS WITH CHANCE CONSTRAINTS: GENERALIZED CONVEXITY AND APPROXIMATION ISSUES

Roger J-B Wets

Department of Mathematics, University of California, Davis

ABSTRACT

Averaging has a smoothing and convexifying effect. So expectation functionals are ‘usually’ convex. However, for an important class of expectation functionals that arise in stochastic programs with chance constraints one can obtain no more than quasi-convexity. Approximation questions for this class of expectation functionals are also being considered.¹

1 INTRODUCTION

Stochastic programming deals with decision making in an uncertain environment. More specifically, stochastic programming models assume that some of the parameters of the problem are only known in a statistical sense. Rather than knowing the values to assign to these parameters, the available information is a probability distribution of their possible values. The uncertainty about these parameters may come from physical considerations such as uncertainty about atmospheric conditions, reliability of measurements, etc., but most commonly it comes from the uncertainty associated with future events such as future demand, future interest or exchange rates, future technological advances, and so on.

The typical stochastic programming model just like deterministic programming models involves (deterministic) constraints on the choice of the decision, say in the form $x \in X \subset \mathbb{R}^n$, ‘costs’ associated with such a decision, say $c(x)$, but

¹Research supported in part by a grant of the National Science Foundation

also a probabilistic description of the random parameters of the problem. Let's denote the random elements of the problem by ξ and with $\Xi \subset \mathbb{R}^N$ the set of possible values of ξ ; more precisely if μ is the probability measure induced by ξ on the Borel subsets of \mathbb{R}^N , Ξ denotes the support of μ , i.e., the smallest closed set of μ -measure one. A decision x and a realization ξ of the random elements ξ brings us to the 'state': (x, ξ) . The choice of a decision must take into account not just the associated costs but also the distribution of the possible 'states'.

For example, if one has to choose the design of a communication network, one should consider not only the cost of the construction of the network but also take into account variations in the demand/traffic as well as the reliability of the network. Or still, when choosing production levels for an assortment of products, one should consider not only per unit profits but also potential demands for these products.

Stochastic programming models either associate with each state a value, say $u(\xi, x)$, or restrict the states to an acceptable class A . In the first case the decision maker is guided by cost/utility considerations and usually, in the second case, by reliability/contractual considerations.

A general formulation for the first class of models would be:

$$\min_{x \in X} E\{f_0(\xi, x)\} \text{ where } f_0(\xi, x) = c(x) + u(\xi, x).$$

One usually refers to

$$\min_{x \in X} E f_0(x) \text{ with } E f_0(x) = \int_{\Xi} f_0(\xi, x) \mu(d\xi),$$

as the *deterministic equivalent problem*. Assuming that the constraint set X is convex, convexity questions related to this optimization then center around the properties of the *expectation functional*: $E f_0$.

When reliability considerations are of paramount importance such as when the decision affects the availability of basic utilities (water, electricity) or seeks to avoid catastrophic situations, the stochastic programming model is often formulated as

$$\min_{x \in X} c(x) \quad \text{subject to } \mu\{\xi \mid (\xi, x) \in A\} \geq \alpha,$$

where $A \subset \Xi \times \mathbb{R}^n$ denotes the 'acceptable' states and $\alpha \in (0, 1]$ indicates the level of reliability that should be achieved. One usually refers to a constraint of this type as a *chance constraint*. A model might actually involve a number

of such constraints but for our purposes it will suffice to consider just one such constraint. In §2 and §3, we are going to be concerned with conditions under which the set

$$S := \left\{ x \in \mathbb{R}^n \mid \mu\{\xi \mid (\xi, x) \in A\} \geq \alpha \right\}$$

is convex, and quasi-convexity plays here a key role.

In §4, we provide a brief introduction to how the convergence theory for expectation functionals can be exploited to obtain approximation results for stochastic programs with chance constraints, namely when the measure μ is replaced by an empirical measure obtained from sampling ξ .

2 CONVEXITY

Let

$$\kappa^{-1}(\xi) := \{x \in \mathbb{R}^n \mid (\xi, x) \in A\}$$

and

$$\kappa(x) := \{\xi \in \Xi \mid (\xi, x) \in A\},$$

and assume that for all x , the set $\kappa(x)$ is measurable, an insignificant assumption in practice. The question about the convexity of S can be reformulated as follows:

$$\text{Is } \text{lev}_{\geq \alpha} \mu \circ \kappa = \{x \in \mathbb{R}^n \mid \mu \circ \kappa(x) \geq \alpha\} \text{ convex?}$$

This will certainly be the case when $\mu \circ \kappa$ has convex level sets, i.e., when $\mu \circ \kappa : \mathbb{R}^n \rightarrow [0, 1]$ is *quasi-concave* which means that for all $x^0, x^1 \in \mathbb{R}^n$ and $\tau \in [0, 1]$,

$$\mu \circ \kappa(x^\tau) \geq \min \{ \mu \circ \kappa(x^0), \mu \circ \kappa(x^1) \} \text{ where } x^\tau = (1 - \tau)x^0 + \tau x^1.$$

Clearly if $\mu \circ \kappa$ is quasi-concave and $x^0, x^1 \in S$ then $\mu \circ \kappa(x^0) \geq \alpha$ and $\mu \circ \kappa(x^1) \geq \alpha$ which implies that for all $\tau \in [0, 1]$, $\mu \circ \kappa(x^\tau) \geq \alpha$ and hence $x^\tau \in S$.

That this isn't going to be the case in general, can be surmised from the following couple of simple examples.

Example 2.1. Let $A = \{(\xi, x) \mid \xi_1 x_1 + \xi_2 x_2 \leq 0\}$ with $\xi = (\xi_1, \xi_2)$ a random vector taking on the values $(1, 1)$, $(1, -1)$, $(-1, 0)$ with equal probability. With $x^0 = (1/2, 1)$, $x^1 = (1/2, -1)$, one has $\mu \circ \kappa(x^0) = \mu \circ \kappa(x^1) = 2/3$. But for all $\tau \in (1/4, 3/4)$, $1/3 = \mu \circ \kappa(x^\tau) < \min \{ \mu \circ \kappa(x^0), \mu \circ \kappa(x^1) \}$. \square

Example 2.2. Let ξ_1, ξ_2 be independent uniformly distributed random variables with support $[0, 1]$, and

$$A = \{(\xi, x) \mid \xi_1 x_1 \leq 1, \xi_2 x_2 \leq 1\} \cap [0, 1]^2 \times [1, \infty)^2.$$

Then for all $x \in [1, \infty) \times [1, \infty)$, $\mu \circ \kappa(x) = x_1^{-1} x_2^{-1}$. With $x^0 = (2, 4)$, $x^1 = (4, 2)$ and $\tau = 0.5$, one has $1/9 = \mu \circ \kappa(x^\tau) < \mu \circ \kappa(x^0) = \mu \circ \kappa(x^1) = 1/8$. Hence $\mu \circ \kappa$ is not quasi-concave. Thus, with $S := \{x \in \mathbb{R}^n \mid \mu\{\xi \mid (\xi, x) \in A\} \geq \alpha\}$, if $\alpha = 8/9$, the points x^0, x^1 belong to S , but for all $\tau \in (0, 1)$, $x^\tau \notin S$. \blacksquare

In the first example, μ is a discrete distribution, but it's easy to modify it so that μ is absolutely continuous. For example, instead of the probability mass being concentrated on three points, almost all the probability mass could be concentrated on sufficiently small neighborhoods of these three points.

The examples have been selected so that for each ξ the sets $\kappa^{-1}(\xi)$ are convex; in fact, these sets are defined by simple linear inequalities. In these examples, quasi-concavity fails because of two basic reasons:

- The sets $\kappa(x^\tau)$ are smaller than the convex combination of $\kappa(x^0)$ and $\kappa(x^1)$.
- The probability measure μ is not of the right 'shape'.

On the basis of these examples, one might conjecture that μ should have 'unimodality' properties.

Let us consider the relationship between the sets $\kappa(x)$ for $x = x^0, x^1$ and $x^\tau = (1 - \tau)x^0 + \tau x^1$ in some specific instances.

Example 2.3. With D a $m \times n$ -matrix, let

$$\begin{aligned} \kappa(x) &= \{\xi \in \Xi \subset \mathbb{R}^m \mid Dx \leq \xi\} \\ &= (Dx) + \mathbb{R}_+^m. \end{aligned}$$

Then, for all x , $\kappa(x)$ is convex and

$$\kappa(x^\tau) = (1 - \tau)\kappa(x^0) + \tau\kappa(x^1).$$

This is the case when the chance constraints are defined by a linear system of inequalities and some, or all, of the right hand sides entries are random. \blacksquare

Example 2.4. With $G : \mathbb{R}^n \rightarrow \mathbb{R}^m$, let

$$\begin{aligned}\kappa(x) &= \{\xi \in \Xi \subset \mathbb{R}^m \mid G(x) \leq \xi\} \\ &= G(x) + \mathbb{R}_+^m.\end{aligned}$$

For all x , $\kappa(x)$ is convex. Now, assuming also that the mapping G is ‘convex’, i.e., for each $i = 1, \dots, m$, the coordinate function $G_i : \mathbb{R}^n \rightarrow \mathbb{R}$ is convex, one has $G(x^\tau) \leq (1 - \tau)G(x^0) + \tau G(x^1)$. And, consequently,

$$\kappa(x^\tau) \supset (1 - \tau)\kappa(x^0) + \tau\kappa(x^1).$$

This is just a generalization of 2.3; it’s the case of joint chance constraints with random right hand sides. Because G isn’t necessarily linear one only obtains an inclusion. \square

Examples 2.5. With $D : \Xi \rightarrow \mathbb{R}^{m \times n}$ be a random matrix, let

$$\kappa(x) = \{D(\xi) \in \Xi \mid D(\xi)x \leq b\}.$$

Again, for all x , the set $\kappa(x)$ is convex, in fact, it’s a convex polyhedron. As a special case, one has

$$\kappa(x) = \{\xi \in \Xi \subset \mathbb{R}^n \mid \langle \xi, x \rangle \leq \beta\}$$

which defines the intersection of Ξ with a closed half-space. Even in this instance, one can’t assert in general that $\kappa(x^\tau) \supset (1 - \tau)\kappa(x^0) + \tau\kappa(x^1)$. Indeed, $\xi^0 \in \kappa(x^0)$, i.e., $\langle \xi^0, x^0 \rangle \leq \beta$, and $\xi^1 \in \kappa(x^1)$, i.e., $\langle \xi^1, x^1 \rangle \leq \beta$, does not imply $\langle \xi^\tau, x^\tau \rangle \leq \beta$. \square

But even if the inclusion $\kappa(x^\tau) \supset (1 - \tau)\kappa(x^0) + \tau\kappa(x^1)$ holds, there is no guarantee that

$$\mu \circ \kappa(x^\tau) \geq (1 - \tau)\mu \circ \kappa(x^0) + \tau\mu \circ \kappa(x^1),$$

which would yield the concavity of $\mu \circ \kappa$, or even that the weaker inequality

$$\mu \circ \kappa(x^\tau) \geq \min[\mu \circ \kappa(x^0), \mu \circ \kappa(x^1)]$$

holds, which yields the quasi-concavity of $\mu \circ \kappa$. However, this will be the case whenever μ is a *logconcave*, or just a *quasi-concave measure*, i.e., if for any pair A_0, A_1 of nonempty, convex, measurable subsets of Ξ , one has

$$\mu(A_\tau) \geq \mu(A_0)^{1-\tau} \mu(A_1)^\tau \text{ where } A_\tau = (1 - \tau)A_0 + \tau A_1,$$

in the logconcave case, or

$$\mu(A_\tau) \geq \min[\mu(A_0), \mu(A_1)], \quad \forall \tau \in [0, 1]$$

in the quasi-concave case. Note that logconcavity of μ implies quasi-concavity since for $\tau \in [0, 1]$ and $\alpha_0, \alpha_1 \geq 0$, the minimum of the expression $\alpha_0^{1-\tau} \alpha_1^\tau$ is attained at $\tau = 0$ or $\tau = 1$ depending on α_0 being smaller than or larger than α_1 ; in the case of equality the minimum is attained at both $\tau = 0$ and $\tau = 1$.

We summarize our analysis in the following statement:

Theorem 2.6. If the probability measure μ defined on Ξ is quasi-concave and for any x^0, x^1 and $\tau \in [0, 1]$

$$\kappa(x^\tau) \supset (1 - \tau)\kappa(x^0) + \tau\kappa(x^1),$$

the function $\mu \circ \kappa : \mathbb{R}^n \rightarrow [0, 1]$ is quasi-concave, and thus $S = \text{lev}_{\geq \alpha} \mu \circ \kappa$ is convex.

Checking the inclusion involving κ must be done on a case by case basis but one can usually not go much beyond Example 2.4. The question thus boils down to identifying the probability measures that are quasi-concave.

3 QUASI-CONCAVE MEASURES

The basic results about quasi-concave probability measures are due to Prékopa [14, 15] and Leindler [11], with extensions and refinements by Borell [7], Brascamp and Lieb [8], Barndorff-Nielsen [6], Rinott [17] and Norkin and Roenko [13]; for applications in Mathematical Statistics one can consult the books of Barndorff-Nielsen [6], and Dharmadhikari and Joag-Dev [9], and for applications in Stochastic Programming the book of Prékopa [16].

We are going to limit ourselves to showing that if the probability measure μ has a density function that is logconcave then μ is a concave measure. A related result identifying density functions generating quasi-concave measures can be obtained by a similar argument. The presentation that follows is a much streamlined version of that found in the literature cited above.

Theorem 3.1. Let $f : \mathbb{R}^n \rightarrow \mathbb{R}_+$ be a probability density function and μ the probability measure induced by f . If f is logconcave, i.e., $f(x) = e^{q(x)}$ with $q : \mathbb{R} \rightarrow [-\infty, \infty)$ concave, then μ is a concave measure.

Proof. Let $C_0, C_1 \subset \mathbb{R}^n$ be nonempty and measurable, and for $\tau \in (0, 1)$, set $C_\tau = (1 - \tau)C_0 + \tau C_1$. The logconcavity of f implies that for all x^0, x^1 and $x^\tau = (1 - \tau)x^0 + \tau x^1$:

$$(f^{1-\tau}(x^0)\mathbb{1}_{C_0}(x^0)) (f^\tau(x^1)\mathbb{1}_{C_1}(x^1)) \leq (f(x^\tau)\mathbb{1}_{C_\tau}(x^\tau)),$$

and thus,

$$f^{1-\tau}(x^0)\mathbb{1}_{C_0}(x^0) \geq \zeta^{1-\tau}, \quad f^\tau(x^1)\mathbb{1}_{C_1}(x^1) \geq \zeta^\tau \implies f(x^\tau)\mathbb{1}_{C_\tau}(x^\tau) \geq \zeta.$$

Hence, for all $\zeta \in \mathbb{R}_+$,

$$(1 - \tau)\{x \in C_0 \mid f(x) \geq \zeta\} + \tau\{x \in C_1 \mid f(x) \geq \zeta\} \subset \{x \in C_\tau \mid f(x) \geq \zeta\},$$

from which it follows that

$$(1 - \tau)\mu(C_0) + \tau\mu(C_1) \leq \mu(C_\tau),$$

since, with λ the Lebesgue measure, for any measurable set C ,

$$\mu(C) = \int_C f(x) dx = \int_0^\infty F_C(\zeta) d\zeta \text{ where } F_C(\zeta) = \lambda(\{x \in C \mid f(x) \geq \zeta\}),$$

and the (inner) Lebesgue measure of the sum of two nonempty sets exceeds or equals the sum of the Lebesgue measure of these sets, cf. [12]. \square

The preceding result is somewhat sharper than the one found in the literature cited earlier where it is shown that μ is a logconcave measure by which one means that for C_0, C_1 (measurable) convex sets and $\tau \in (0, 1)$

$$\mu(C_0)^{1-\tau}\mu(C_1)^\tau \leq \mu(C_\tau).$$

But every concave measure is also a logconcave measure. The arithmetic mean/geometric mean inequality implying $\mu(C_0)^{1-\tau}\mu(C_1)^\tau \leq (1 - \tau)\mu(C_0) + \tau\mu(C_1)$. Note also that we didn't have to restrict the argument to convex sets C_0, C_1 .

Here are a few examples of probability density functions that are logconcave.

Example 3.2. Let $D \subset \mathbb{R}^n$ be convex and suppose that $\lambda(D) < \infty$. The function

$$q(x) = \begin{cases} \ln(\lambda(D)) & \text{if } x \in D, \\ -\infty & \text{otherwise,} \end{cases}$$

is concave. The density $f(x) = e^{q(x)}$, the uniform distribution on D , is logconcave. \square

Example 3.3. If we let

$$q(x) = \gamma - \frac{1}{2}\langle (x - m), \Sigma^{-1}(x - m) \rangle \text{ with } \gamma = \ln [(2\pi)^n \det(\Sigma)]^{-1/2},$$

where Σ is symmetric, positive definite, then $f(x) = e^{q(x)}$ is the density function of a multivariate normally distributed random variable with mean m and covariance Σ . Again q is clearly concave and thus f is a logconcave density. \square

Other examples include the Dirichlet density (which specializes to the Beta density in dimension 1), the Wishart distribution, etc. One can refer to [6] and [16] for more complete dictionaries.

4 APPROXIMATIONS

Let \mathcal{A} the σ -field induced on $\Xi \subset \mathbb{R}^N$ by the random vector $\boldsymbol{\xi}$; let (Ξ, \mathcal{A}, μ) be a probability space. We refer to a function $f : \Xi \times \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ as a *random lsc function* if

- i) for all $\xi \in \Xi$, $x \mapsto f(\xi, x)$ is lsc,
- ii) $(\xi, x) \mapsto f(\xi, x)$ is (jointly) $\mathcal{A} \otimes \mathcal{B}$ -measurable; \mathcal{B} is the Borel field on \mathbb{R}^n .

The theoretical foundations of stochastic programming are centered around the properties of *expectation functionals* of the type:

$$x \mapsto Ef(x) = E\{f(\boldsymbol{\xi}, x)\} = \int_{\Xi} f(\xi, x) \mu(d\xi),$$

where f is a random lsc function. The recourse models for stochastic programming, i.e., models where the decision maker is guided by cost/utility considerations, can always be formulated in terms of minimizing an expectation functional, cf. §1 and for more details [19]. But also stochastic programming models with chance constraints, i.e., based on reliability considerations, involve

expectation functionals, albeit of a special type. Indeed, again with $A \subset \Xi \times \mathbb{R}^n$, the set of ‘acceptable’ states and $\alpha \in (0, 1]$, the chance constraint

$$\mu\left\{\xi \in \Xi \mid (\xi, x) \in A\right\} \geq \alpha$$

can be written as

$$E\{f(\boldsymbol{\xi}, x)\} \leq 0, \text{ where } f(\xi, x) = \alpha - \mathbb{1}_A(\xi, x).$$

If A is an $\mathcal{A} \otimes \mathcal{B}$ -measurable subset of $\Xi \times \mathbb{R}^n$ and for all $\xi \in \Xi$, $\kappa^{-1}(\xi) = \{x \in \mathbb{R}^n \mid (\xi, x) \in A\}$ is closed, then f is a random lsc function; for example, if $A = \{(\xi, x) \mid G(\xi, x) \leq 0\}$ where $G : \Xi \times \mathbb{R}^n \rightarrow \mathbb{R}^d$ is measurable with respect to ξ and continuous in x . A fortiori, f will be a random lsc function if A is a closed subset of $\Xi \times \mathbb{R}^n$ which, when $A = \{(\xi, x) \mid G(\xi, x) \leq 0\}$, would certainly be the case if G is (jointly) continuous.

For random lsc functions there is a fundamental law of large numbers that provides the theoretical justification for the use of sampling techniques to deal with the minimization of expectation functionals. Here we are going to exploit this law of large numbers to justify sampling when dealing with chance constraints; further elaboration can be found in [10].

Theorem 4.1. Let $\boldsymbol{\xi}^1, \boldsymbol{\xi}^2, \dots$ be independent random vectors, all with the same distribution as $\boldsymbol{\xi}$, and let $f : \Xi \times \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be a random lsc function. Suppose that for all $x \in \mathbb{R}^n$ there is a neighborhood V of x and a measurable function $\alpha_- : \Xi \rightarrow \mathbb{R}_-$ with $\int_{\Xi} \alpha_-(\xi) \mu(d\xi) > -\infty$ such that μ -almost surely $f(\cdot, x') \geq \alpha$ for all $x' \in V$. Let $\boldsymbol{\mu}^\nu$ be the (random) empirical measure induced on (Ξ, \mathcal{A}) by the random variables $\boldsymbol{\xi}^1, \dots, \boldsymbol{\xi}^\nu$, and let

$$\mathbf{E}^\nu \mathbf{f} := \int_{\Xi} f(\xi, \cdot) \boldsymbol{\mu}^\nu(d\xi)$$

be the (random) expectation of f with respect to $\boldsymbol{\mu}^\nu$. Then, with $Ef(x) = \int_{\Xi} f(\xi, x) \mu(d\xi)$,

$$e\text{-}\lim_{\nu \rightarrow \infty} \mathbf{E}^\nu \mathbf{f} = Ef, \quad \mu^\infty\text{-almost surely.}$$

A somewhat weaker version of this law of large numbers, requiring a uniform lower bound on the functions $\{f(\xi, \cdot), \xi \in \Xi\}$ was originally obtained by Attouch and Wets [4]. The version stated here is essentially that found in Artstein and Wets [1]; further refinements can be found in Korf and Wets [10] where the independent, identically distributed assumptions are replaced by stationarity and ergodicity properties.

The assertion in Theorem 4.1 is in terms of an epi-limit. Recall that a function $g : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ is an epi-limit of a sequence of functions $g^\nu : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, i.e., $g = \text{e-lim}_\nu g^\nu$, if for all $x \in \mathbb{R}^n$:

- i) $\liminf_\nu g^\nu(x^\nu) \geq g(x)$, for all sequences $x^\nu \rightarrow x$,
- ii) $\limsup_\nu g^\nu(x^\nu) \leq g(x)$ for some sequence $x^\nu \rightarrow x$.

More generally one can define,

$$\begin{aligned} \text{e-lim inf}_\nu g^\nu(x) &= \inf_{\{x^\nu \rightarrow x\}} \liminf_\nu g^\nu(x^\nu) \\ \text{e-lim sup}_\nu g^\nu(x) &= \inf_{\{x^\nu \rightarrow x\}} \limsup_\nu g^\nu(x^\nu) \end{aligned}$$

with $\text{e-lim sup}_\nu g^\nu = \text{e-lim}_\nu g^\nu = \text{e-lim inf}_\nu g^\nu$ if the limit exists; the infimum in the preceding expression is with respect to all sequences $x^\nu \rightarrow x$. For a sequence of sets $\{C^\nu \subset \mathbb{R}^n, \nu \in \mathbb{N}\}$ and δ_D the indicator function of the set D , taking on the values 0 on D and ∞ outside D , one has

$$\begin{aligned} C = \limsup_\nu C^\nu &\Leftrightarrow \delta_C = \text{e-lim inf}_\nu \delta_{C^\nu}, \\ C = \liminf_\nu C^\nu &\Leftrightarrow \delta_C = \text{e-lim sup}_\nu \delta_{C^\nu}, \end{aligned}$$

and $C = \lim_\nu C^\nu$ if and only if $\delta_C = \text{e-lim}_\nu \delta_{C^\nu}$.

The main implications of epi-convergence have to do with the convergence of the optimal values and solutions of the minimization problems, consult [3,18,2] for example. For further use let's record the following immediate consequences of the definition.

Proposition 4.2. Let $\{g, g^\nu : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}\}$ be such that $g = \text{e-lim}_\nu g^\nu$ and $h : \mathbb{R}^n \rightarrow \mathbb{R}$ a continuous function. Then $g + h = \text{e-lim}_\nu (g^\nu + h)$.

Proposition 4.3. Suppose $f = \text{e-lim}_\nu f^\nu$. Then

$$\limsup_\nu \text{argmin} f^\nu \subset \text{argmin} f.$$

We are also going to rely on the following property.

Theorem 4.4 [5]. For functions g^ν and g on \mathbb{R}^n , one has:

- i) $g \leq \text{e-lim inf}_\nu g^\nu$ if and only if $\limsup_\nu (\text{lev}_{\leq \alpha^\nu} g^\nu) \subset \text{lev}_{\leq \alpha} g$ for all sequences $\alpha^\nu \rightarrow \alpha$;

- ii) $g \geq \text{e-lim sup}_\nu g^\nu$ if and only if $\liminf_\nu (\text{lev}_{\leq \alpha^\nu} g^\nu) \supset \text{lev}_{\leq \alpha} g$ for some sequence $\alpha^\nu \rightarrow \alpha$, in which case such a sequence can be chosen with $\alpha^\nu \searrow \alpha$.
- iii) $g = \text{e-lim}_\nu g^\nu$ if and only if both conditions hold.

These basic results allow us to almost immediately draw the following implications for stochastic programs with chance constraints.

Theorem 4.5. Let's consider the following stochastic program with chance constraints:

$$\min c(x) \text{ such that } \mu(\{\xi \mid (\xi, x) \in A\}) \geq \alpha. \quad (P)$$

where $c : \mathbb{R}^n \rightarrow \mathbb{R}$ is continuous, A is a closed subset of $\Xi \times \mathbb{R}^n$, $\alpha \in (0, 1]$. Let ξ^1, ξ^2, \dots be random samples of ξ and consider the following stochastic programs with chance constraints:

$$\min c(x) \text{ such that } \mu^\nu(\{\xi \mid (\xi, x) \in A\}) \geq \alpha^\nu$$

where μ^ν is the empirical measure generated from the samples ξ^1, \dots, ξ^ν and $\alpha^\nu \rightarrow \alpha$. Suppose that for all ν ,

$$x^\nu \in S^\nu = \left\{ x \mid \mu^\nu(\{\xi \mid (\xi, x) \in A\}) \geq \alpha^\nu \right\}, \quad (P^\nu)$$

then μ^∞ -almost surely, every cluster of the sequence $\{x^\nu\}_{\nu \in \mathbb{N}}$ is a feasible solution of the stochastic program with chance constraints (P) .

Moreover there exists a sequence $\alpha^\nu \nearrow \alpha$ such that if x^ν is an optimal solution of (P^ν) , then every cluster point of the sequence $\{x^\nu\}_{\nu \in \mathbb{N}}$ is actually an optimal solution of (P) .

Proof. From our earlier discussion it follows that

$$f(\xi, x) := -\mathbb{1}_A(\xi, x)$$

is a random lsc function and

$$Ef(x) = -\mu(\{\xi \mid (\xi, x) \in A\}), \quad E^\nu f(x) = -\mu^\nu(\{\xi \mid (\xi, x) \in A\}).$$

Theorem 4.1 and Proposition 4.2 imply $Ef = \text{e-lim}_\nu E^\nu f$ P^∞ -almost surely. In turn, this yields via Theorem 4.4, that

$$\limsup_\nu (\text{lev}_{\leq -\alpha^\nu} E^\nu f) \subset \text{lev}_{\leq -\alpha} Ef$$

which means that whenever \bar{x} is a cluster point of a sequence of points $\{x^\nu\}_{\nu \in N}$ with $x^\nu \in \text{lev}_{\leq -\alpha^\nu} E^\nu f$, then $\bar{x} \in \text{lev}_{\leq -\alpha} Ef$.

Theorem 4.4 also guarantees the existence of a sequence $\alpha^\nu \nearrow \alpha$ such that

$$\text{lev}_{\leq -\alpha} Ef =: C = \lim_\nu C^\nu \text{ with } C^\nu := \text{lev}_{\leq -\alpha^\nu} E^\nu f.$$

And thus $\delta_C = \text{e-lim}_\nu \delta_{C^\nu}$. Proposition 4.2 allows us to conclude that

$$c + \delta_C = \text{e-lim}_\nu (c + \delta_{C^\nu}).$$

Proposition 4.3 tells us: if $x^\nu \in \text{argmin}(c + \delta_{C^\nu})$ and the x^ν cluster at some point \bar{x} , then this cluster point $\bar{x} \in \text{argmin}(c + \delta_C)$, i.e., solves (P). \blacksquare

Acknowledgement. I am grateful for Ms L. Korf careful reading of the manuscript. Her comments have been very helpful.

REFERENCES

- [1] Z. Artstein and R.J-B Wets, “Consistency of minimizers and the SLLN for stochastic programs”, *J. of Convex Analysis* 2, 1995, 1-17.
- [2] J.-P. Aubin and H. Frankowska, *Set-Valued Analysis*, 1990, Birkhäuser Inc., Cambridge, Mass.
- [3] Hedy Attouch, “Variational Convergence for Functions and Operators”, 1984, *Applicable Mathematics Series*, Pitman, London.
- [4] H. Attouch and R.J-B Wets, “Epigraphical processes: laws of large numbers for random lsc functions”, *Séminaire d’Analyse Convexe*, Montpellier, 1990, 13.1–13.29.
- [5] G. Beer, R.T. Rockafellar and R. J-B Wets, “A characterization of epi-convergence in terms of convergence of level sets”, *Proceedings of the American Mathematical Society* 116, 1992, 753–761.
- [6] O. Barndorff-Nielsen, “Information and Exponential Families in Statistical Theory”, 1978, *J. Wiley*, New York.

- [7] C. Borell, “Convex set functions in d -space”, *Periodica Mathematica Hungarica* 6, 1975, 111–136.
- [8] H.J. Brascamp and E.H. Lieb, “On extensions of the Brunn-Minkowski and Prékopa-Leindler theorems, including inequalities for log-concave functions, and with applications to the diffusion equations”, *Journal of Functional Analysis* 22, 1976, 366–389.
- [9] S. Dharmadhikari and K. Joag-Dev, “Unimodality, Convexity and Applications”, 1988, Academic Press, New York.
- [10] L. Korf and R.J-B Wets, “Ergodic limit laws for stochastic optimization problems”, Manuscript, University of California, Davis.
- [11] L. Leindler, “On a certain converse of Hölder’s inequality II”, *Acta Scientiarum Mathematicarum (Szeged)* 33, 1972, 217–223.
- [12] L. Lusternik, “The Brunn-Minkowski inequality for Lebesgue measurable functions”, *Doklady Akademii Nauk S.S.S.R.* 3, 1935, 55–58.
- [13] V.I. Norkin and N.V. Roenko, “ α -concave functions and measures and their applications”, *Cybernetics and Systems Analysis*, 1991, 77–88.
- [14] A. Prékopa, “Logarithmic concave measures with application to stochastic programming”, *Acta Scientiarum Mathematicarum (Szeged)* 32, 1971, 301–316.
- [15] A. Prékopa, “On logarithmic concave measures and functions”, *Acta Scientiarum Mathematicarum (Szeged)* 34, 1973, 335–343.
- [16] A. Prékopa, *Stochastic Programming*, 1995, Kluwer Publishers, Dordrecht.
- [17] Y. Rinott, “On the convexity of measures”, *Annals of Probability* 4, 1976, 1020–1026.
- [18] R.T. Rockafellar and Roger J-B Wets, “Variational systems, an introduction”, *Multifunctions and Integrands: Stochastic Analysis, Approximation and Optimization*, G. Salinetti Editor, *Lecture Notes in Mathematics* 1091, 1984, Springer Verlag, Berlin, 1–54.
- [19] R.J-B Wets, “Stochastic programming”, *Handbook for Operations Research and Management Sciences*, G. Nemhauser, A. Rinnooy Kan and M. Todd Editors, 1989, Elsevier Science Publishers, Amsterdam, 573–629.