Qualitative Stability of Convex Programs with Probabilistic Constraints

René Henrion
Weierstrass Institute Berlin

December 15, 1998

1991 Mathematics Subject Classification. 90C15, 90C31.

Keywords. stochastic programming, probabilistic constraints, qualitative stability, r-concave measures.

Abstract

We consider convex stochastic optimization problems with probabilistic constraints which are defined by so-called r-concave probability measures. Since the true measure is unknown in general, the problem is usually solved on the basis of estimated approximations, hence the issue of perturbation analysis arises in a natural way. For the solution set mapping and for the optimal value function, stability results are derived. In order to include the important class of empirical estimators, the perturbations are allowed to be arbitrary in the space of probability measures (in contrast to the convexity property of the original measure). All assumptions relate to the original problem. Examples show the necessity of the formulated conditions and illustrate the sharpness of results in the respective settings.

1 Introduction

Most constraint sets in optimization problems can be described by an inclusion $0 \in H(x)$, where H is some multifunction. In a large class of applied problems, the constraints are subject to uncertainty such that their description changes to $\xi \in H(x)$, where ξ is some random variable. Usually, the optimization of x- variables has to be carried out without or with partial knowledge only about the realizations of the random variable. Then, of course, the above formulation has to be replaced by some reasonable deterministic equivalent. One possible way is to define an admissible x as to satisfiy the inclusion $\xi \in H(x)$ with high probability: $\mu(\xi \mid \xi \in H(x)) \geq p$ or briefly $\mu(H(x)) \geq p$, where μ is the probability distribution of ξ and $p \in (0,1)$ is some specified probability level. We shall refer to such constraints as to probabilistic constraints. In the following, we shall consider optimization problems of the type

$$(P) \qquad \min\{g(x) \mid x \in X,\ \mu(H(x)) \geq p\}.$$

Here, g is a cost function on $\mathbb{R}^m$, $X \subset \mathbb{R}^m$ is a non-specified set of deterministic constraints, $H : \mathbb{R}^m \rightrightarrows \mathbb{R}^s$ is a multifunction and μ is the probability distribution of an s-dimensional random variable ξ, i.e. $\mu \in \mathcal{P}(\mathbb{R}^s)$, where $\mathcal{P}(\mathbb{R}^s)$ denotes the space of probability measures on $\mathbb{R}^s$. Throughout this paper, we shall make the following basic assumptions for problem (P):

$$g : \mathbb{R}^m \to \mathbb{R} \text{ is convex.} \tag{1}$$

$$X \subseteq \mathbb{R}^m \text{ is closed and convex.} \tag{2}$$

$$H : \mathbb{R}^m \rightrightarrows \mathbb{R}^s \text{ has closed and convex graph.} \tag{3}$$

$$\mu \in \mathcal{P}(\mathbb{R}^s) \text{ is } r\text{- concave for some } r < 0. \tag{4}$$

We note, that (3) is equivalent to a description

$$H(x) = \{z \in \mathbb{R}^s \mid h(x,z) \leq 0\},$$

where $h : \mathbb{R}^m \times \mathbb{R}^s \to \mathbb{R}^k$ is convex and lower semicontinuous (in both variables). Concerning assumption (4), we refer to Section 2.1.

A peculiarity of stochastic optimization problems of type (P) is that usually there is no or only partial information on the measure μ available. In solution procedures, μ is therefore replaced by sample-based estimators which for increasing sampling size are supposed to approximate μ. In this context, the question of stability arises in quite a natural way: When the sampling size tends to infinity, do the optimal solutions and their cost function values of the approximate problems converge towards an optimal solution and its cost function value, respectively, of the original problem? This issue is intimately related with the qualitative stability of the solution set mapping and of the optimal value function, both depending on perturbed probability measures in a neighbourhood of the original one. For a list of papers dealing with stability aspects in stochastic programming problems with probabilistic constraints, we refer to e.g. [1], [4],[6],[7],[9],[13],[15] and references therein.

At this point, it is emphasized that we allow for arbitrary perturbations of μ in the space $\mathcal{P}(\mathbb{R}^s)$ of probability measures on $\mathbb{R}^s$. In particular, the important class of empirical measures is included as approximation. Hence, although the original measure μ is supposed to have some nice convexity property (assumption 4), the perturbations are allowed even to be discontinuous. The purpose of this paper is twofold: first, it aims at a fairly complete characterization of qualitative stability in the settings introduced above by means of verifiable conditions for the unperturbed problem; second, a series of examples collected in section 4 shall illustrate the necessity of assumptions and the sharpness of results.

2 Preliminaries

In this section, we collect some basic definitions and facts which are necessary for the following analysis.

2.1 r-concave probability measures

Here we recall the notion of an r-concave probability measure for some $r \in [-\infty, \infty]$ which was imposed as a basic assumption to the problem we are going to analyze (see (4)). We start with the definition of the generalized mean function m_r on $\mathbb{R}_+ \times \mathbb{R}_+ \times [0,1]$:

$$m_r(a,b;\lambda) = \begin{cases} (\lambda a^r + (1-\lambda)b^r)^{1/r} & \text{if } r \in (0,\infty) \text{ or } r \in (-\infty,0), ab > 0 \\ 0 & \text{if } ab = 0, r \in (-\infty, 0) \\ a^\lambda b^{1-\lambda} & \text{if } r = 0 \\ \max\{a,b\} & \text{if } r = \infty \\ \min\{a,b\} & \text{if } r = -\infty \end{cases} \tag{5}$$

The measure $\mu \in \mathcal{P}(\mathbb{R}^s)$ is called *r-concave* ([3]) for some $r \in [-\infty, \infty]$, if the inequality

$$\mu(\lambda B_1 + (1-\lambda)B_2) \geq m_r(\mu(B_1), \mu(B_2); \lambda) \tag{6}$$

holds for all Borel measurable, convex subsets B_1, B_2 of $\mathbb{R}^s$ and all $\lambda \in [0,1]$ for which the convex combination $\lambda B_1 + (1-\lambda)B_2$ is Borel measurable as well (note that convex sets need not be Borel measurable, see [5]). For $r = 0$ and $r = -\infty$, μ is also called log-concave and quasi-concave, respectively. Since $m_r(a, b; \lambda)$ is increasing in r if all the other variables are fixed, the sets $\mathcal{M}_r(\mathbb{R}^s)$ of all r-concave probability measures are increasing if r is decreasing, i.e., we have for all $-\infty < r_1 \leq r_2 < \infty$ that

$$\mathcal{M}_{-\infty}(\mathbb{R}^s) \supseteq \mathcal{M}_{r_1}(\mathbb{R}^s) \supseteq \mathcal{M}_{r_2}(\mathbb{R}^s) \supseteq \mathcal{M}_{\infty}(\mathbb{R}^s). \tag{7}$$

Recall, that the distribution function F_μ corresponding to some $\mu \in \mathcal{P}(\mathbb{R}^s)$ is defined by

$$F_\mu(z) = \mu(\xi \mid \xi_i \leq z_i \ (i = 1, \ldots, s)) = \mu(z + \mathbb{R}^s_-).$$

For this particular case of cells $B = z + \mathbb{R}^s_-$, $z \in \mathbb{R}^s$, and for $r \in (-\infty, 0)$, the inequality (6) implies the distribution function F_μ to have the property that the extended-real-valued function F_μ^r is convex on $\mathbb{R}^s$. Moreover, (6) and (7) entail that F_μ is quasi-concave on $\mathbb{R}^s$.

As a consequence of a Theorem by Prékopa ([11], Th. 4.2.1.), the probability measure μ induced by a log-concave density f (i.e. a density the logarithm of which is concave) is log-concave as well, in particular it is r-concave for all $r < 0$ in view of (7). Examples of distributions having log-concave densities are the uniform distribution (on any bounded convex subset of $\mathbb{R}^s$ with non-zero Lebesgue measure), the (nondegenerate) multivariate normal distribution, the Dirichlet distribution, the multivariate Student and Pareto distributions. These examples qualify our basic assumption (4) as being not very restrictive. For more information on this issue, proofs and details we refer to Chapter 4 of [11].

2.2 The parametric problem and $\mathcal{B}$-discrepancy between probability measures

In order to study qualitative stability, we imbed problem (P) into the parametric problem

$$(P_\nu) \qquad \min\{g(x) \mid x \in \Phi(\nu)\} \quad (\nu \in \mathcal{P}(\mathbb{R}^s)),$$

where the constraint set mapping $\Phi : \mathcal{P}(\mathbb{R}^s) \rightrightarrows \mathbb{R}^m$ is defined as $\Phi(\nu) = \{x \in X \mid \nu(H(x)) \geq p\}$. Clearly, $(P) = (P_\mu)$. We are interested in the behaviour of the solution set and value function corresponding to this parametric problem. For technical reasons, we introduce the slightly more general localized concepts for some open $V \subseteq \mathbb{R}^m$ and $\nu \in \mathcal{P}(\mathbb{R}^s)$:

$$\begin{aligned} \varphi_V(\nu) &= \inf\{g(x) \mid x \in X \cap \operatorname{cl} V,\ \nu(H(x)) \geq p\} \\ \Psi_V(\nu) &= \operatorname{argmin}\{g(x) \mid x \in X \cap \operatorname{cl} V,\ \nu(H(x)) \geq p\} \end{aligned}$$

We recall the following elementary fact:

$$\emptyset \neq \Psi(\nu) \subseteq V \Longrightarrow \Psi(\nu) = \Psi_V(\nu), \varphi(\nu) = \varphi_V(\nu) \tag{8}$$

By φ, Ψ without index, we refer to the usual optimal value function and set of optimal solutions respectively (i.e. $V = \mathbb{R}^m$).

Before stating any stability result for optimal solutions and values as functions of perturbed probability measures $\nu \in \mathcal{P}(\mathbb{R}^s)$ in a neighbourhood of the original measure μ, we want to specify a distance on $\mathcal{P}(\mathbb{R}^s)$ which is suitable for our purposes (cf. discussion in [14]):

$$\alpha_{\mathcal{B}}(\nu_1, \nu_2) = \sup_{B \in \mathcal{B}} |\nu_1(B) - \nu_2(B)|, \quad \nu_1, \nu_2 \in \mathcal{P}(\mathbb{R}^s) \tag{9}$$

Here, $\mathcal{B}$ is a system of closed subsets of $\mathbb{R}^s$ such that it contains all sets $\{H(x)|x \in X\}$ (with H and X as introduced in (2) and (3)) and that it forms a determining class, i.e. whenever $\nu_1|_{\mathcal{B}} = \nu_2|_{\mathcal{B}}$, then $\nu_1 = \nu_2$. This last condition implies $\alpha_{\mathcal{B}}$ to be a distance, which is also referred to as the $\mathcal{B}$-discrepancy. A useful choice in the setting of our problem (P) under the stated assumptions is $\mathcal{B} = \{H(x)|x \in X\} \cup \{z + \mathbb{R}^s_-|z \in \mathbb{R}^s\}$, where the second part of the union serves to turn $\mathcal{B}$ into a determining class, while the first part is essential to obtain the important observations of the following Proposition:

Proposition 2.1 *In problem (P_ν), it holds that*

1. *The multifunction $\Phi : (\mathcal{P}(\mathbb{R}^s), \alpha_{\mathcal{B}}) \rightrightarrows \mathbb{R}^m$ has closed graph.*
2. *For $\nu \in \mathcal{P}(\mathbb{R}^s)$, define $w_\nu(x) := \nu(H(x))$. Assume that there exists some subset $Q \subseteq X$, such that $w_\nu(x) \geq \rho > 0$ for all $x \in Q$. Then, for all $r < 0$, there exist constants $c, \delta > 0$ such that*

$$|w_\nu^r(x) - w_{\nu'}^r(x)| \leq c\alpha_{\mathcal{B}}(\nu, \nu') \quad \forall x \in Q\, \forall \nu' \in \mathcal{P}(\mathbb{R}^s),\ \alpha_{\mathcal{B}}(\nu, \nu') < \delta.$$

Proof. 1. is shown in [13] (Prop. 3.1). For the second assertion, note that

$$|u^r - v^r| \leq |r| \max\{u^{r-1}, v^{r-1}\}\, |u - v| \quad \forall u, v > 0.$$

Then, choosing $\delta := \rho/2$, one has $w_{\nu'}(x) \geq \rho/2 > 0 \quad \forall x \in Q\, \forall \nu' \in \mathcal{P}(\mathbb{R}^s),\ \alpha_{\mathcal{B}}(\nu, \nu') < \delta$. Fix c as $|r|\,(\rho/2)^{r-1}$. ■

As a consequence of assertion 1. in the last proposition, all constraint sets are closed and, hence, so are all solution sets $\Psi(\nu)$.

3 Qualitative Stability

In this section we study qualitative stability in terms of upper and lower semicontinuity of the solution set mapping and (upper Lipschitz) continuity of the optimal value function in the parametric version (P_ν) of the problem (P). The following theorem gives the main result in this direction.

Theorem 3.1 *Consider the parametric problem (P_ν) under the basic assumptions (1-4). Let additionally the following assumptions be satisfied at μ:*

1. *$\Psi(\mu)$ is nonempty and bounded.*

2. There exists some $\hat{x} \in X$ such that $\mu(H(\hat{x})) > p$ (Slater condition).

Then, the multifunction $\Psi : (\mathcal{P}(\mathbb{R}^s), \alpha_{\mathcal{B}}) \rightrightarrows \mathbb{R}^m$ is upper semicontinuous at μ, and there exist constants $L, \delta > 0$, such that

$$\Psi(\nu) \neq \emptyset \text{ and } |\varphi(\nu) - \varphi(\mu)| \leq L\alpha_{\mathcal{B}}(\nu, \mu) \text{ for all } \nu \in \mathcal{P}(\mathbb{R}^s) \text{ with } \alpha_{\mathcal{B}}(\nu, \mu) < \delta.$$

Proof. We define $f(x) := \mu^r(H(x)) - p^r$ with the improper value ∞ allowed in case of $\mu(H(x)) = 0$. Then, in view of $r < 0$, the unperturbed constraint set may be written as $\Phi(\mu) = \{x \in X \mid f(x) \leq 0\}$, where f is convex due to (3) and (4). Furthermore, $f(\hat{x}) = \mu^r(H(\hat{x})) - p^r < 0$ by assumption 2., i.e. $\hat{x}$ is a Slater point of f w.r.t. X. Using a result by Klatte [10], it was shown in [13] (Cor. 3.7.) that under the assumptions made here, the desired continuity properties at μ hold in the <u>localized</u> case. More precisely, with V being some bounded, open neighbourhood of $\Psi(\mu)$ (see assumption 1.), one has that Ψ_V is upper semicontinuous at μ and there exist constants $L_1, \delta_1 > 0$, such that $|\varphi_V(\nu) - \varphi(\mu)| \leq L_1\alpha_{\mathcal{B}}(\nu, \mu)$ for all $\nu \in \mathcal{P}(\mathbb{R}^s)$ with $\alpha_{\mathcal{B}}(\nu, \mu) < \delta_1$. Note that, by definition, one has $\varphi_V(\mu) = \varphi(\mu)$ and $\Psi_V(\mu) = \Psi(\mu)$ since $\emptyset \neq \Psi(\mu) \subseteq V$ (see (8)).

Suppose now that Ψ was not upper semicontinuous at μ. Then, by the compactness of $\Psi(\mu)$ (see assumption 1. and recall the closedness of $\Psi(\mu)$), there exists some $\varepsilon > 0$ as well as sequences ν_n, x_n such that $\alpha_{\mathcal{B}}(\nu_n, \mu) \to 0$, $x_n \in \Psi(\nu_n)$ and $d(x_n, \Psi(\mu)) \geq \varepsilon$. On the other hand, in case that local nonemptiness of Ψ is violated, $\Psi(\nu_n) = \emptyset$ would hold for a sequence ν_n with $\alpha_{\mathcal{B}}(\nu_n, \mu) \to 0$. Since $\Psi(\mu) \neq \emptyset$ by assumption 1., there is some $x^* \in \Psi(\mu)$, hence $x^* \in X \cap V$ and $f(x^*) \leq 0$. With the Slater point $\hat{x}$ from assumption 2., select $\lambda \in (0, 1]$ such that $\tilde{x} := \lambda\hat{x} + (1 - \lambda)x^* \in X \cap V$ (by convexity of X). Since f is convex, it follows that $f(\tilde{x}) \leq \lambda f(\hat{x}) + (1 - \lambda)f(x^*) < 0$. Hence, $\mu(H(\tilde{x})) > p$ and $\nu(H(\tilde{x})) \geq p$ for $\nu \in \mathcal{P}(\mathbb{R}^s)$ with $\alpha_{\mathcal{B}}(\nu, \mu) < \delta := 1/2(\mu(H(\tilde{x})) - p)$. In particular, the localized perturbed constraint sets $\Phi(\nu_n) \cap \operatorname{cl} V$ are non-empty (they contain $\tilde{x}$) and compact for n large enough. Consequently, $\Psi_V(\nu_n) \neq \emptyset$ (since g is continuous as a convex function which is finite-valued on $\mathbb{R}^m$, see (1)). But then, $\Psi(\nu_n) = \emptyset$ means the existence of a sequence $x_n \in \Phi(\nu_n) \setminus \operatorname{cl} V$ with $g(x_n) \leq \varphi_V(\nu_n)$.

Summarizing, if the upper semicontinuity or the local non-emptiness of Ψ is violated at μ, then there are sequences ν_n, x_n such that (with some $\varepsilon > 0$)

$$\alpha_{\mathcal{B}}(\nu_n, \mu) \to 0,\ x_n \in \Phi(\nu_n),\ g(x_n) \leq \varphi_V(\nu_n) \text{ and } d(x_n, \Psi(\mu)) \geq \varepsilon. \tag{10}$$

In the following, we lead these relations to a contradiction. We define the set

$$A := \begin{cases} \mathbb{R}^m & \text{if } \Psi(\mu) = \Phi(\mu) \\ g^{-1}(-\infty, g(x')] & \text{if there is some } x' \in \Phi(\mu) \setminus \Psi(\mu),\ f(x') < 0. \end{cases} \tag{11}$$

Note first, that the case distinction above is complete. Indeed, assume that $f(x) = 0$ for all $x \in \Phi(\mu) \setminus \Psi(\mu)$ and choose an arbitrary such x. Then, $f(\lambda\hat{x} + (1 - \lambda)x) \leq \lambda f(\hat{x}) + (1 - \lambda)f(x) < 0$ for $\lambda \in (0, 1]$, hence, due to $\hat{x}, x \in \Phi(\mu)$ and to convexity of $\Phi(\mu)$, one gets $\lambda\hat{x} + (1 - \lambda)x \in \Psi(\mu)$. This, however, entails $\Psi(\mu) = \Phi(\mu)$.

We note that $A \cap \Phi(\mu)$ is convex and compact. The convexity being evident from the convexity of g, the compactness follows in the first case above from the compactness of

$\Psi(\mu) = A \cap \Phi(\mu)$ due to assumption 1. In the second case, one may use the fact that the level set $g^{-1}(-\infty, \varphi(\mu)]$ of the convex function g intersected with the convex set $\Phi(\mu)$ equals $\Psi(\mu)$ which is compact by assumption 1. Then, according to [12] (Cor. 8.7.1), the intersection $\Phi(\mu) \cap g^{-1}(-\infty, \alpha)]$ has to be compact for all levels α, hence the compactness of $A \cap \Phi(\mu)$ follows with $\alpha := g(x')$.

Next, we verify that the sequence x_n in (10) satisfies $x_n \in A$ for n large enough. While this is trivial in the first case of (11), assuming the contrary in the second case would yield the existence of subsequences ν_{n_k}, x_{n_k} such that $\alpha_{\mathcal{B}}(\nu_{n_k}, \mu) \to_k 0$ and

$$\varphi_V(\mu) = \varphi(\mu) < g(x') < g(x_{n_k}) \leq \varphi_V(\nu_{n_k}) \xrightarrow[k \to \infty]{} \varphi_V(\mu)$$

which contradicts the already stated continuity of φ_V at μ.

Now, we claim that $f(x_n) > 0$ holds for the sequence x_n in (10) with n large enough. Indeed, otherwise $f(x_{n_k}) \leq 0$ holds for some subsequence. Since then $x_{n_k} \in A \cap \Phi(\mu)$ by definition of f and by the statement proven just before (also recall that $x_{n_k} \in \Phi(\nu_{n_k}) \subseteq X$) and, since $A \cap \Phi(\mu)$ was shown above to be compact, one has $x_{n_{k_l}} \to x^* \in A \cap \Phi(\mu)$ for another subsequence. Now, because of $g(x_{n_{k_l}}) \leq \varphi_V(\nu_{n_{k_l}})$ (see (10)), the continuity of φ_V at μ and that of g as a convex function provide $g(x^*) \leq \varphi_V(\mu) = \varphi(\mu)$ which entails the contradiction $x^* \in \Psi(\mu)$ to $d(x_{n_{k_l}}, \Psi(\mu)) \geq \varepsilon$ in (10).

Since $x_n \in \Phi(\nu_n)$, one has $\nu_n(H(x_n)) \geq p$, hence $\mu(H(x_n)) \geq p - \alpha_{\mathcal{B}}(\mu, \nu_n) \geq p/2 > 0$ for n large enough. Therefore, statement 2. in Proposition 2.1 yields the existence of some $c > 0$, such that $\mu^r(H(x_n)) - \nu_n^r(H(x_n)) \leq c\alpha_{\mathcal{B}}(\mu, \nu_n)$ for n large enough. Hence, $f(x_n) \leq c\alpha_{\mathcal{B}}(\mu, \nu_n)$. Next, define

$$\bar{x} := \begin{cases} \hat{x} & \text{in the first case of } (11) \\ x' & \text{in the second case of } (11). \end{cases}$$

Set $y_n := \tau_n \bar{x} + (1 - \tau_n)x_n$, where $\tau_n \in [0, 1]$ is chosen such that $f(y_n) = 0$ (recall that $f(\bar{x}) < 0$ and $0 < f(x_n) \leq c\alpha_{\mathcal{B}}(\mu, \nu_n) < \infty$). Then, $0 \leq \tau_n f(\bar{x}) + (1 - \tau_n)c\alpha_{\mathcal{B}}(\mu, \nu_n)$ by convexity of f. Since $(1 - \tau_n)c\alpha_{\mathcal{B}}(\mu, \nu_n) \to 0$ and $f(\bar{x}) < 0$, one derives that $\tau_n \to 0$. Furthermore, one has $\|y_n - \bar{x}\| = (1 - \tau_n)\|x_n - \bar{x}\|$. Now, $y_n \in \Phi(\mu)$ (since $f(y_n) = 0$ and $y_n \in X$ due to $\bar{x}, x_n \in X$ and to convexity of X). Finally, one has $y_n \in A$ which is trivial in the first case of (11) and which follows in the second case of (11) from $x', x_n \in A$ since A is convex. Knowing that $A \cap \Phi(\mu)$ is compact, the sequence y_n must be bounded, which, by the relations above, entails the sequence x_n to be bounded as well. Observing that $\|y_n - x_n\| = \tau_n\|\bar{x} - x_n\|$, we conclude $\|y_n - x_n\| \to 0$. It follows that $d(x_n, \Phi(\mu) \cap A) \to 0$ and, hence, $x_{n_k} \to x^* \in \Phi(\mu) \cap A$ for some subsequence. Now, similarly to an argumentation above, the relation $g(x_{n_k}) \leq \varphi_V(\nu_{n_k})$ along with continuity of φ_V at μ yields $g(x^*) \leq \varphi(\mu)$ and, hence, the contradiction $x^* \in \Psi(\mu)$ to $d(x_{n_k}, \Psi(\mu)) \geq \varepsilon$ in (10).

It remains to verify the statement on φ. Up to now, we have shown that Ψ is upper semicontinuous at μ and nonempty-valued close to μ. Accordingly, there is some $\delta > 0$ such that $\emptyset \neq \Psi(\nu) \subseteq V$ for all $\nu \in \mathcal{P}(I\!\!R^s)$ with $\alpha_{\mathcal{B}}(\nu, \mu) < \delta$. But then $\varphi(\nu) = \varphi_V(\nu)$ for these ν (see (8)) and the formulated continuity property of φ results from the same property of φ_V already stated in the beginning of this proof. ∎

As already stated in the proof of the preceding Theorem, a corresponding version with localized mappings Ψ_V and φ_V was shown in an earlier work. A restriction to localizations seemed to be necessary due to allowing for non-convex perturbations of the probabilistic constraint. However, the uniformity property of the $\mathcal{B}$-discrepancy introduced in Section 2.2 permits to obtain results even for the non-localized mappings.

A series of Examples in section 4 shows the necessity of the assumptions made in Theorem 3.1 in order to arrive at the assertions stated there. All examples relate to the problem (P') introduced below, which is a special case of problem (P). For instance, Examples 4.1 and 4.2 show that none of the three properties asserted in the theorem (upper semicontinuity and local non-emptiness of Ψ as well as continuity of φ) can be guaranteed when the Slater condition is dispensed with while all other assumptions are kept. In Example 4.3, all assumptions of the Theorem are met with the exception that $\Psi(\mu) = \emptyset$. As a consequence, the upper semicontinuity and, trivially, the local non-emptiness of Ψ fail to hold. Finally, the r-concavity of μ cannot be dropped in the theorem as it is shown in Example 4.4, where of course assumption 2. of the theorem can no longer be interpreted as a Slater condition due to absence of convexity. Nevertheless, it is possible to arrive at continuity results in the non-convex case. More precisely, exactly the same results as in Theorem 3.1 were shown in [7] (Th. 1) to hold for the localized mappings Ψ_V and φ_V in a more general setting of problem (P) than under assumptions (1-4), namely for locally Lipschitzian g, closed X, H with closed graph and $\mu \in \mathcal{P}(\mathbb{R}^s)$ arbitrary. Then, however, assumption 2. of Theorem 3.1 has to be replaced by the so-called metric regularity w.r.t. X of the probabilistic constraint, which in the setting of Theorem 3.1 is equivalent with the Slater condition, but which is a stronger requirement in the non-convex context.

Example 4.5 demonstrates that the assumptions of Theorem 3.1 do not suffice to derive the lower semicontinuity of Ψ. Inspecting Fig. 2 reveals that a lack of curvature in the level set of the chance constraint is responsible for the solution set to collapse after arbitrary small perturbations of the measure. Therefore it seems natural to require some strict convexity property of the probability measure. Before stating a corresponding result, some auxiliary facts are needed.

In the following, the multifunction H is specified as $H(x) = \{z \in \mathbb{R}^s \mid \xi \leq h(x)\}$, where $h : \mathbb{R}^m \to \mathbb{R}^s$ is supposed to have concave components h_i in order to satisfy the basic assumption (3). This specific system of inequalities, where the realizations z of the random vector ξ occur explicitly on the left-hand side, typically reflects some supply/demand relationship, where the random demand z of some good has to be met by the supply $h(x)$ depending on the decision variable x. The assumption of this specific structure is crucial for the following, since it allows to write the constraint function as a composition of two single-valued mappings in contrast to the set-valued formulation of (P). Indeed, by definition of the distribution function, (P) now writes as

$$(P') \qquad \min\{g(x) \mid x \in X,\ F_\mu(h(x)) \geq p\}.$$

Furthermore, the system $\mathcal{B}$ of closed sets figuring in the definition of the discrepancy distance $\alpha_{\mathcal{B}}$ reduce to the system of cells $z + \mathbb{R}^s_-$, $z \in \mathbb{R}^s$, since the sets $H(x)$ now are cells themselves (cf. section 2.2). Accordingly, the discrepancy $\alpha_{\mathcal{B}}$ turns into the

Kolmogorov distance

$$d_K(\nu_1,\nu_2)=\sup_{z\in\mathbb{R}^s}|F_{\nu_1}(z)-F_{\nu_2}(z)|\quad(\nu_1,\nu_2\in\mathcal{P}(\mathbb{R}^s)),$$

which equals the uniform distance between the distribution functions induced by the corresponding measures. Therefore, from now on, the stability results are formulated by using d_K.

The following lemma opens a way by means of decomposition separately to study properties of F_μ and h in the context of lower semicontinuity of Ψ.

Lemma 3.2 *Under the assumptions of Theorem 3.1 let V be an open (in the maximum norm) ball containing $\Psi(\mu)$. Set*

$$\begin{aligned}
Y_V &= [h(X\cap \mathrm{cl}\ V)+\mathbb{R}^s_-]\cap F_\mu^{-1}([p/2,1])\\
Y(\nu) &= \mathrm{argmin}\,\{\pi(y)\mid y\in Y_V, F_\nu(y)\geq p\}\quad(\nu\in\mathcal{P}(\mathbb{R}^s))\\
\pi(y) &= \inf\{g(x)\mid x\in X\cap \mathrm{cl}\ V,\ h(x)\geq y\},\\
\sigma(y) &= \mathrm{argmin}\ \{g(x)\mid x\in X\cap\mathrm{cl}\,V,\ h(x)\geq y\}\ (y\in Y_V).
\end{aligned}$$

Then it holds that

1. *Y_V is convex and compact.*

2. *π is convex, finite and lower semicontinuous on Y_V.*

3. *There is some $\delta>0$ such that for all $\nu\in\mathcal{P}(\mathbb{R}^s)$ with $d_K(\mu,\nu)<\delta$*

$$\begin{aligned}
\varphi(\nu) &= \inf\{\pi(y)\mid y\in Y_V, F_\nu(y)\geq p\} &(12)\\
\Psi(\nu) &= \sigma(Y(\nu)) &(13)
\end{aligned}$$

4. *$Y:\mathcal{P}(\mathbb{R}^s)\rightrightarrows\mathbb{R}^s$ is upper semicontinuous at μ.*

Proof.
ad 1.
The convexity of Y_V follows from the assumed convexity of X and V along with h having concave components and μ being r-concave (note that $F_\mu\leq 1$ since $\mu\in\mathcal{P}(\mathbb{R}^s)$). The compactness of $X\cap$cl V implies closedness of $h(X\cap$cl $V)+\mathbb{R}^s_-$ and, hence, closedness of Y_V due to F_μ being upper semicontinuous as a distribution function. If Y_V was not bounded, there would be a sequence $y_n\in Y_V$ with $\|y_n\|\to\infty$. By definition, $y_n\leq h(x_n)$ for $x_n\in X\cap$cl V. Since h is continuous (having concave components which are finite-valued on $\mathbb{R}^m$), each component of y_n is bounded from above. On the other hand, the condition $F_\mu(y_n)\geq p/2>0$ (due to $y_n\in Y_V$) implies all components of y_n to be bounded from below, since F_μ is a distribution function of some probability measure. This contradicts $\|y_n\|\to\infty$.
ad 2.

The convexity and finiteness of π on Y_V are immediate from the properties of g, h, X and Y_V. Now, consider any sequence $y_n \to \bar{y}$ with $y_n, \bar{y} \in Y_V$ and $\pi(y_n) \to \alpha \in I\!R \cup \{\infty\}$. Then, there are $x_n \in X \cap \mathrm{cl}\, V$ such that $y_n \leq h(x_n)$ and $g(x_n) = \pi(y_n)$ (note that the constraint set in the definition of $\pi(y)$ is nonempty and compact for $y \in Y_V$). By compactness of $X \cap \mathrm{cl}\, V$, one has $x_{n_k} \to_k \bar{x}$ for some subsequence, where $\bar{x} \in X \cap \mathrm{cl}\, V$ and $h(\bar{x}) \geq \bar{y}$ due to continuity of h. Consequently,

$$\pi(\bar{y}) \leq g(\bar{x}) = \lim_{k\to\infty} g(x_{n_k}) = \lim_{k\to\infty} \pi(y_{n_k}) = \alpha$$

by continuity of g. In particular, $\alpha > -\infty$. It results that π is lower semicontinuous on Y_V.

ad 3.

From the local nonemptiness and upper semicontinuity at μ of Ψ, which was stated in Theorem 3.1, one derives the existence of some $\delta > 0$ such that

$$\emptyset \neq \Psi(\nu) \subseteq V \quad \forall \nu \in \mathcal{P}(I\!R^s),\, d_K(\mu,\nu) < \delta. \tag{14}$$

Fix an arbitrary such ν. Then (14) and (8) yield that $\Psi_V(\nu) = \Psi(\nu) \neq \emptyset$ and $\varphi_V(\nu) = \varphi(\nu)$. Select some $\bar{x} \in \Psi(\nu)$. Then, since $\bar{x} \in \Psi_V(\nu) \subseteq \{x \in X \cap \mathrm{cl}\, V \mid F_\nu(h(x)) \geq p\}$, it follows that

$$\varphi(\nu) = g(\bar{x}) \geq \pi(h(\bar{x})) \geq \inf\{\pi(y) \mid y \in Y_V, F_\nu(y) \geq p\},$$

where the last inequality relies on $F_\mu(h(\bar{x})) \geq p/2$, which is true if δ in (14) is chosen smaller than $p/2$. For the reverse direction of (12), consider an arbitrary $\bar{y} \in Y_V$ with $F_\nu(\bar{y}) \geq p$. By definition, there is an $\tilde{x} \in X \cap \mathrm{cl}\, V$ with $h(\tilde{x}) \geq \bar{y}$. Choose some $x^* \in \sigma(\bar{y})$ (note that the constraint set in the definition of $\sigma(\bar{y})$ contains $\tilde{x}$ and, hence, is nonempty and compact such that $\sigma(\bar{y})$ is nonempty). We continue by $\pi(\bar{y}) = g(x^*) \geq \varphi_V(\nu) = \varphi(\nu)$, where the inequality follows from $x^* \in X \cap \mathrm{cl}\, V$, $h(x^*) \geq \bar{y}$ as well as $F_\nu(h(x^*)) \geq F_\nu(\bar{y}) \geq p$ (recall that distribution functions are nondecreasing). Since $\bar{y}$ was arbitrary, this establishes (12).

Concerning (13), note that $\pi(h(x)) \leq g(x) = \varphi(\nu)$ holds for all $x \in \Psi_V(\nu) = \Psi(\nu)$, hence (12) implies

$$\pi(h(x)) = \inf\{\pi(y) \mid y \in Y_V, F_\nu(y) \geq p\} = g(x) \quad \forall x \in \Psi(\nu).$$

Therefore, $x \in \sigma(h(x))$ and $h(x) \in Y(\nu)$ for all these x. Consequently, $\Psi(\nu) \subseteq \sigma(Y(\nu))$. For the reverse inclusion, let $x \in \sigma(Y(\nu))$ be arbitrary, i.e., $x \in \sigma(y)$ for some $y \in Y(\nu)$. Then,

$$g(x) = \pi(y) \leq \pi(h(x')) \leq g(x') \quad \forall x' \in X \cap \mathrm{cl}\, V,\, F_\nu(h(x')) \geq p,$$

which amounts to $x \in \Psi_V(\nu) = \Psi(\nu)$.

ad 4.

Although it seems tempting to proove 4. via Theorem 3.1 by setting $g := \pi$, $h := id$, $X := Y_V$, this is not justified since the domain $h(X \cap \mathrm{cl}\, V) + I\!R^s_-$ of π is not the whole space in general and, hence, π - although convex on this domain - cannot be assumed to be locally Lipschitzian. Instead, we write

$$Y(\nu) = \{y \in I\!\!R^s \mid F_\nu(y) \geq p\} \cap \{y \in Y_V \mid \pi(y) \leq \varphi(\nu)\},$$

where the first part is a closed multifunction (compare 1. in Prop. 2.1 with $h := id$ and $X := I\!\!R^s$) and the second part too is closed at μ: In fact, if $\nu_n \in \mathcal{P}(I\!\!R^s)$, $y_n \in Y_V$ are sequences with $y_n \to \bar{y}$, $d_K(\mu, \nu_n) \to 0$ and $\pi(y_n) \leq \varphi(\nu_n)$, then by continuity of φ at μ, closedness of Y_V (see 1.) and lower semicontinuity of π on Y_V (see 2.) one gets

$$\bar{y} \in Y_V \text{ and } \pi(\bar{y}) \leq \liminf_{n\to\infty} \pi(y_n) \leq \varphi(\mu),$$

which is the desired closedness property. As a consequence, Y itself is closed at μ (as an intersection of closed multifunctions). On the other hand, for all ν, $Y(\nu)$ is contained in the compact set Y_V (see 1.) by definition, hence Y must be upper semicontinuous at μ.

■

Relation (13) suggests that lower semicontinuity of Ψ may be formulated in terms of the same property for the two constituents σ and Y:

Proposition 3.3 *The solution set mapping Ψ of problem (P') is lower semicontinuous at $\mu \in \mathcal{P}(I\!\!R^s)$ provided that the following two assumptions hold:*

1. *$\sigma : Y_V \rightrightarrows I\!\!R^m$ is lower semicontinuous at each $y \in Y_V$.*
2. *$Y : \mathcal{P}(I\!\!R^s) \rightrightarrows I\!\!R^s$ is lower semicontinuous at μ.*

Proof.
Let $U \subseteq I\!\!R^m$ be an arbitrary open set with $U \cap \Psi(\mu) \neq \emptyset$. By (13), there exists some $y \in Y(\mu) \subseteq Y_V$ such that $U \cap \sigma(y) \neq \emptyset$. According to assumption 1., there exists an open neighbourhood V of y such that

$$U \cap \sigma(y') \neq \emptyset \quad \text{for all } y' \in V \cap Y_V. \tag{15}$$

On the other hand, since $y \in Y(\mu) \cap V$, assumption 2. provides the existence of some $\delta > 0$ such that $Y(\nu) \cap V \neq \emptyset$ for all $\nu \in \mathcal{P}(I\!\!R^s)$ with $d_K(\mu, \nu) < \delta$. Combining this with (15) yields that (due to $Y(\nu) \subseteq Y_V$)

$$U \cap \Psi(\nu) = U \cap \sigma(Y(\nu)) \supseteq U \cap \sigma(Y(\nu) \cap V) \neq \emptyset \quad \text{for all} \nu \in \mathcal{P}(I\!\!R^s) \text{ with } d_K(\mu, \nu) < \delta.$$

This, however, is the asserted lower semicontinuity of Ψ at μ.

■

We continue by deriving verifiable conditions for the two assumptions in Proposition 3.3. First we make use of the following concept introduced in [2]: A function $\alpha : I\!\!R^n \to I\!\!R$ is called *weakly analytic* if for any $a, b \in I\!\!R^n$ with $a \neq b$, one has

$$\alpha \text{ is constant on a line segment } \operatorname{conv}\{a, b\} \Longrightarrow \alpha \text{ is constant on the entire line } \operatorname{lin}\{a, b\}.$$

Accordingly, a subset $Q \subseteq I\!R^n$ has a convex, weakly analytic description, if $Q = \{x \in I\!R^n \mid \alpha(x) \leq 0\}$ for some mapping $\alpha : I\!R^n \to I\!R^{n'}$ with convex, weakly analytic components α_i. In particular, all analytic or strictly convex functions are weakly analytic. Furthermore, each polyhedral set Q has a convex, weakly analytic description by means of the affine linear mapping $\alpha(x) := Ax + b$ for some matrix A and vector b.

Proposition 3.4 *Assumption 1. of Proposition 3.3 is satisfied if in problem (P') the functions g and h_i are weakly analytic and the set X has a convex, weakly analytic description.*

Proof.
Consider the multifunction $M : \Lambda \subseteq I\!R^{n_1} \rightrightarrows I\!R^{n_2}$ defined by $M(\lambda) = \{x \in I\!R^{n_2} \mid \alpha(x) \leq \beta(\lambda)\}$, with functions $\alpha : I\!R^{n_2} \to I\!R^{n_3}$ and $\beta : \Lambda \to I\!R^{n_3}$. The results in [2] (Th. 3.2.1., Th. 3.2.2. and Cor. 3.2.2.1) imply M to be lower semicontinuous at all $\lambda \in \Lambda$ under the following assumptions:

- $M(\lambda) \neq \emptyset \ \forall \lambda \in \Lambda$.
- The α_i are convex and weakly analytic.
- β is continuous.

First, put in the context of the definitions of Lemma 3.2:

$$\Lambda := Y_V, \qquad \left.\begin{array}{lcl} \alpha_i(x) & := & -h_i(x) \\ \beta_i(y) & := & -y_i \end{array}\right\} i = 1, \ldots, s \qquad \left.\begin{array}{lcl} \alpha_{s+i}(x) & := & \gamma_i(x) \\ \beta_i(y) & := & 0 \end{array}\right\} i = 1, \ldots, n'.$$

Here, γ refers to a convex, weakly analytic description of the set $X \cap \mathrm{cl}\, V$ (recall that $\mathrm{cl}\, V$ is a closed ball in the maximum norm and hence a polyhedral set). Obviously, M represents the multifunction $y \mapsto \{x \in X \cap \mathrm{cl}\, V \mid h(x) \geq y\}$ here. Now, the assumptions above are satisfied due to the definition of Y_V, and to the h_i being concave and weakly analytic. Consequently, M is lower semicontinuous at all $y \in Y_V$. Therefore, again by an argument of parametric optimization ([2], Th. 4.2.2), π defined in Lemma 3.2 is upper semicontinuous on Y_V due to the continuity of g. This yields the continuity of π on Y_V along with statement 2. of Lemma 3.2. Now, we apply the result cited in the beginning of this proof, a second time: In addition to the settings above, put $\alpha_{s+n'+1} := g$ and $\beta_{s+n'+1} := \pi$. Then, M is exactly the multifunction σ and again, the three assumptions above are met by $M = \sigma$: since, for any $y \in Y_V$ the set $\{x \in X \cap \mathrm{cl}\, V \mid h(x) \geq y\}$ is nonempty by the definition of Y_V and compact by $\mathrm{cl}\, V$ being a closed ball, the set $\sigma(y)$ of global minima of g on this set must be nonempty as well due to continuity of g. The remaining two assumptions are valid due to $\alpha_{s+n'+1} = g$ being convex (see (1)) and to the continuity of $\beta_{s+n'+1} = \pi$ just shown before. As a consequence, σ is lower semicontinuous at all $y \in Y_V$ as was to be shown.

■

A counter-example in [2] (Ex. 3.3.1.) shows that the weak analyticity assumptions in the last Proposition cannot be dispensed with.

Next, we turn to the second assumption in Proposition 3.3. At this point, the strict convexity property of the probability measure, mentioned before as a necessary requirement for lower semicontinuity of the solution set, comes into play.

Lemma 3.5 *Assumption 2. of Lemma 3.3 is satisfied in problem (P') if, in addition to the assumptions of Theorem 3.1, there exists some open convex neighbourhood U of $Y(\mu)$ such that F_μ^r is strictly convex on U (with $r < 0$ being the modulus of r- concavity of μ).*

Proof.

Setting $b_\nu(y) := F_\nu^r(y) - p^r$ for $\nu \in \mathcal{P}(\mathbb{R}^s)$, the original problem $(\widetilde{P}_\mu)$ may be written as

$$(\widetilde{P}_\mu) \quad \min \ \{\pi(y) \mid y \in Y_V, b_\mu(y) \leq 0\}.$$

Clearly, $(\widetilde{P}_\mu)$ is a convex program (since the r- concavity of μ implies F_μ^r to be convex (see Section 2.1) and due to 1. and 2. in Lemma 3.2). Also, $(\widetilde{P}_\mu)$ satisfies the Slater condition $b_\mu(\hat{y}) < 0$ for some $\hat{y} \in Y_V$. Indeed, in the proof of Theorem 3.1, the existence of some $\tilde{x} \in X \cap V$ with $F_\mu^r(h(\tilde{x})) = \mu^r(H(\tilde{x})) < p^r$ was shown, hence one may take $\hat{y} := h(\tilde{x})$. We proceed by case distinction with respect to the relation between $Y(\mu)$ and the solution set $Q := \arg\min\{\pi(y) \mid y \in Y_V\}$ of the relaxed problem:

<u>case 1:</u> $Y(\mu) \cap Q = \emptyset$.

Choose some $y^* \in Y(\mu)$ (recall that $Y(\mu) \neq \emptyset$ due to $\Psi(\mu) \neq \emptyset$ and to (13)). Since π and b_μ are finite-valued on Y_V and $\varphi(\mu) = \pi(y^*) > -\infty$, the Slater condition ensures the existence of a Lagrange multiplier $\lambda^* \geq 0$ such that (cf. [12], Cor. 28.2.1)

$$\pi(y^*) = \min \ \{\pi(y) + \lambda^* b_\mu(y) \mid y \in Y_V\} \text{ and } \lambda^* b_\mu(y^*) = 0.$$

By the case 1- assumption, one has $\lambda^* \neq 0$, hence $\lambda^* > 0$ and $\pi + \lambda^* b_\mu$ is strictly convex on $Y_V \cap U$ due to the additional assumption in this lemma. Accordingly,

$$\pi(y) + \lambda^* b_\mu(y) > \pi(y^*) + \lambda^* b_\mu(y^*) = \pi(y^*) \quad \forall y \in Y_V \cap U,\, y \neq y^*,$$

which implies that y^* is the unique minimizer of $(\widetilde{P}_\mu)$, i.e., $Y(\mu) = \{y^*\}$. Because of this uniqueness, the upper semicontinuity of Y at μ (statement 4. of Lemma 3.2) entails the desired lower semicontinuity of Y at μ.

<u>case 2:</u> $Y(\mu) \cap Q \neq \emptyset$.

In this case, $Y(\mu)$ has the simple representation

$$Y(\mu) = \{y \in Q \mid b_\mu(y) \leq 0\}. \tag{16}$$

Note also, that Q is closed and convex by the properties of π and Y_V stated in Lemma 3.2.

<u>case 2.1</u> $\exists \, \bar{y} \in Y(\mu),\, b_\mu(\bar{y}) < 0$.

Since $\bar{y} \in Y_V$, one has $F_\mu(\bar{y}) \geq p/2 > 0$ such that statement 2. of Proposition 2.1 yields the existence of $c, \delta > 0$ with

$$|b_\mu(\bar{y}) - b_\nu(\bar{y})| < c d_K(\mu, \nu) \quad \forall \nu \in \mathcal{P}(\mathbb{R}^s),\, d_K(\mu, \nu) < \delta.$$

In particular, choosing $\delta_0 := \min\{|b_\mu(\bar{y})|c^{-1}, \delta\}$, one gets $b_\nu^r(\bar{y}) < 0$ for all $\nu \in \mathcal{P}(\mathbb{R}^s)$ with $d_K(\mu,\nu) < \delta_0$. Since on the other hand, $\bar{y} \in Q$ by (16), one derives that $\bar{y} \in Y(\nu)$, hence $Y(\nu) \cap Q \neq \emptyset$ for all these ν. Consequently, the representation (16) holds as

$$Y(\nu) = \{y \in Q \mid b_\nu(y) \leq 0\}$$

for all ν close to μ.

Now, $\bar{y}$ is a Slater point of the constraint $b_\mu(y) \leq 0$ with respect to Q. According to the results in [13] (Cor. 3.7 and Th. 3.2), the multifunction Y satisfies a so-called upper Pseudo-Lipschitzian property at all $(\mu, y) \in \operatorname{Gph} Y$. This means in particular, that each $y \in Y(\mu)$ is supplied with neighbourhoods V_y of y and U_y of μ and with a constant $L_y > 0$ such that

$$d(y', Y(\nu)) \leq L_y d(\mu,\nu) \quad \forall \nu \in U_y \ \forall y' \in Y(\mu) \cap V_y.$$

The compactness of $Y(\mu) \subseteq Y_V$ (see Lemma 3.2) then allows to extract a neighbourhood $\tilde{U}$ of μ, an open set $\tilde{V}$ containing $Y(\mu)$ and a constant $L > 0$ such that

$$d(y, Y(\nu)) \leq L d(\mu,\nu) \quad \forall \nu \in \tilde{U} \ \forall y \in Y(\mu).$$

This, however, implies the lower semicontinuity of Y at μ.
case 2.2 $b_\mu(y) = 0 \ \forall y \in Y(\mu)$.
The convexity of $Y(\mu)$ along with the strict convexity of b_μ on $U \supseteq Y_V$ imply that $Y(\mu)$ reduces to a singleton. Then, as in case 1., the upper semicontinuity of Y at μ yields the lower semicontinuity at μ.

■

Combining the results of Theorem 3.1, Proposition 3.3, Proposition 3.4 and Lemma 3.5 one gets the following statement on continuity (i.e. upper- and lower semicontinuity at the same time) of the solution set mapping to problem (P'):

Theorem 3.6 *Consider problem (P') under the following assumptions:*

1. *g is convex and weakly analytic.*
2. *The h_i are concave and weakly analytic.*
3. *X has a convex, weakly analytic description.*
4. *μ is r-concave for some $r < 0$.*
5. *$\Psi(\mu)$ is nonempty and bounded.*
6. *There exists some $\hat{x} \in X$ such that $F_\mu(h(\hat{x})) > p$.*
7. *F_μ ist strictly convex on some open convex neighbourhood U of the compact set Σ with*

   $Y(\mu) \subseteq \Sigma := [h(\Psi(\mu)) + \mathbb{R}^s_-] \cap \{y \in \mathbb{R}^s \mid F_\mu(y) \geq p\} \subseteq Y_V.$

Then, the multifunction $\Psi : (\mathcal{P}(I\!R^s), d_K) \rightrightarrows I\!R^m$ is continuous at μ.

Note that all assumptions of the Theorem relate to the original data of the problem (no assumptions with perturbed measures ν are involved). We recall, that assumptions 1. and 2. are satisfied, for instance, if g and the h_i are convex (concave, respectively) and analytic (e.g. linear) or stricly convex (concave, respectively). Assumption 3. is met among others by polyhedral sets X or balls (in any of the p-norms). Assumption 4. has already been qualified in section 2.1 to hold for most of the common multivariate probability measures. Assumptions 5. and 6. were shown above to be indispensable in the context of upper semicontinuity of solutions, and they are common in general parametric programming problems. Finally, the strict convexity property of the measure μ assumed in 7. was found to be necessary when passing from upper to lower semicontinuity of solutions (see Example 4.5). In case that h is a linear mapping, it is sufficient to require U in assumption 7. to be a convex, open neighbourhood of the simpler set $\Sigma := h(\Psi(\mu))$.

Having the qualitative results obtained so far, one might ask about quantitative stability of the solution set in program (P'). This question was investigated in [7] in terms of relating the Hausdorff distance between the solution sets of the original and the perturbed problems to the Kolmogorov distance between the original and the perturbed measure. Example 4.6 demonstrates, that the assumptions of Theorem 3.1 allone are not sufficient in order to derive any Hölder rate of upper semicontinuity for solutions. The basic additional argument is to strengthen assumption 7. of Theorem 3.6 towards a strong convexity property of F_μ^r. Of course, this raises the question whether such strong convexity properties of probability measures are still as common as simple convexity of F_μ^r (coming from the r-concavity of μ). So far, some partial results have been obtained in this direction, for example the multivariate normal distribution with independent components or the uniform distribution on rectangles in $I\!R^s$ satisfy a strong convexity property.

4 Examples

Example 4.1 *In the program (P') put $m = s = 2$, $g(x_1, x_2) = x_2 - x_1$, $h = \mathrm{id}$, $X = \{(x_1, x_2) \mid x_1 + x_2 = 3/2\}$, $p = 1/4$ and let $\mu \in \mathcal{P}(I\!R^2)$ be defined as the uniform distribution on the right upper triangle* conv $\{(1,0),(0,1),(1,1)\}$. *According to section 2.1, μ is logconcave as a uniform distribution on a bounded, convex set, hence it is r-concave for any $r < 0$. The distribution function F_μ of μ is given by*

$$F_\mu(x_1,x_2) = \begin{cases} 1 & , \ x_1, x_2 \geq 1 \\ (x_1 + x_2 - 1)^2 & , \ x_1 + x_2 \geq 1 \text{ and } \mathrm{x}_1, \mathrm{x}_2 \in [0,1], \\ x_1^2 & , \ x_2 \geq 1 \text{ and } \mathrm{x}_1 \in [0,1], \\ x_2^2 & , \ x_1 \geq 1 \text{ and } \mathrm{x}_2 \in [0,1] \\ 0 & , \ \text{else} \end{cases}$$

Hence, F_μ is constant on the line segments $\{(x_1,x_2) \in [0,1]^2 \mid x_1 + x_2 = a\}$ with $a \in [0,1]$ (see Fig. 1). Now, let $\tilde{\mu}$ be the uniform distribution on $[1/2,1]^2$ and consider the perturbed

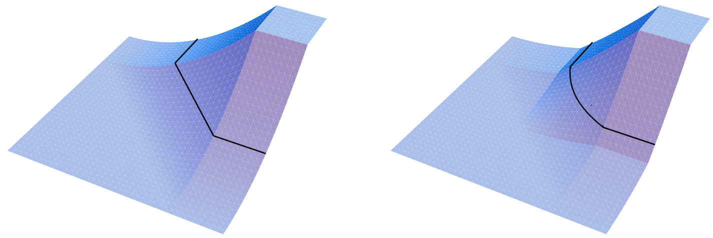

Figure 1: Distribution function F_μ for the uniform distribution on the right upper triangle conv $\{(1,0),(0,1),(1,1)\}$ (left) and distribution function F_{μ_λ} ($\lambda = 0.5$) for the perturbed measure. The level lines $F_\mu(x_1,x_2) = p$ and $F_{\mu_\lambda}(x_1,x_2) = p$ are indicated on both graphs.

probability measures $\mu_\lambda = (1-\lambda)\mu + \lambda\tilde{\mu}$, $\lambda \in [0,1]$. The induced perturbed distribution function F_{μ_λ} is illustrated in Fig. 1 for $\lambda = 1/2$. By definition, the discrepancy distance (which reduces to the Kolmogorov distance here) between μ and μ_λ computes as $d_K(\mu,\mu_\lambda) = \lambda d_K(\mu,\tilde{\mu})$, hence μ_λ converges towards μ with $\lambda \downarrow 0$.

Some calculation shows that the perturbed level set $F^{-1}_{\mu_\lambda}(p)$ is given by

$$\left\{ (x_1,x_2) \;\middle|\; \begin{array}{lll} x_1 = \alpha_\lambda & \textit{if} & x_2 \geq 1 \\ x_2 = \alpha_\lambda & \textit{if} & x_1 \geq 1 \\ x_2 = \beta(x_1,\lambda) & \textit{if} & x_1 \in [\alpha_\lambda, 1] \end{array} \right\},$$

where

$$\alpha_\lambda = \frac{2\lambda - \sqrt{1+3\lambda}}{2(\lambda-1)}; \; \beta(x,\lambda) = \frac{-2+2x+2\lambda x - \sqrt{1+3\lambda-16\lambda x+16\lambda x^2}}{2\,(\lambda-1)}.$$

Note, that $\beta(3/4,\lambda) = 3/4$ for all $\lambda \in [0,1]$ but $\beta(x,\lambda) > 3/2 - x$ for all $x \in [\alpha_\lambda, 1]$ with $x \neq 3/4$ and for all $\lambda \in (0,1]$. Consequently, for arbitrarily small $\lambda > 0$, the perturbed constraint set $\{x \in X \mid F_{\mu_\lambda}(x) \geq p\}$ reduces to the singleton $\{(3/4,3/4)\}$ (compare Fig. 2). In particular, $\Psi(\mu_\lambda) = \{(3/4,3/4)\}$ for any $\lambda > 0$. On the other hand, the constraint set of the unperturbed problem is given by the line segment conv $\{(1/2,1),(1,1/2)\}$ *(see Fig. 2) and compare definition of α_0 and $\beta(x,0)$). Consequently, the original solution set is $\Psi(\mu) = \{(1,1/2)\}$ by definition of g. Therefore, Ψ is not upper semicontinuous at μ. At the same time, one has $\varphi(\mu) = -1/2$ but $\varphi(\mu_\lambda) = 0$ for any $\lambda > 0$, hence φ is not continuous at μ. The reason for the failure of Theorem 3.1 with respect to the continuity properties of Ψ and φ is the absence of a Slater point.*

Example 4.2 *In the program (P'), set $m = s = 1$, $h(x) = g(x) = x$, $X = (-\infty, 0]$, $p = 1/2$ and let μ be the uniform distribution on the interval $[-1,1]$. Then, μ is log-concave and, hence, r-concave for all $r < 0$ (see Section 2.1). Since $F_\mu(x) = \min\{\max\{0,(x+1)/2\},1\}$, the constraint set and, hence, the solution set reduce to the singleton $\{0\}$, whereas, after an arbitrary small perturbation of μ in the sense of Kolmogorov distance (uniform distribution on $[-1+\varepsilon, 1+\varepsilon]$), the constraint and, hence, the solution set become empty. Again, all assumptions of Theorem 3.1 except Slater's condition are satisfied.*

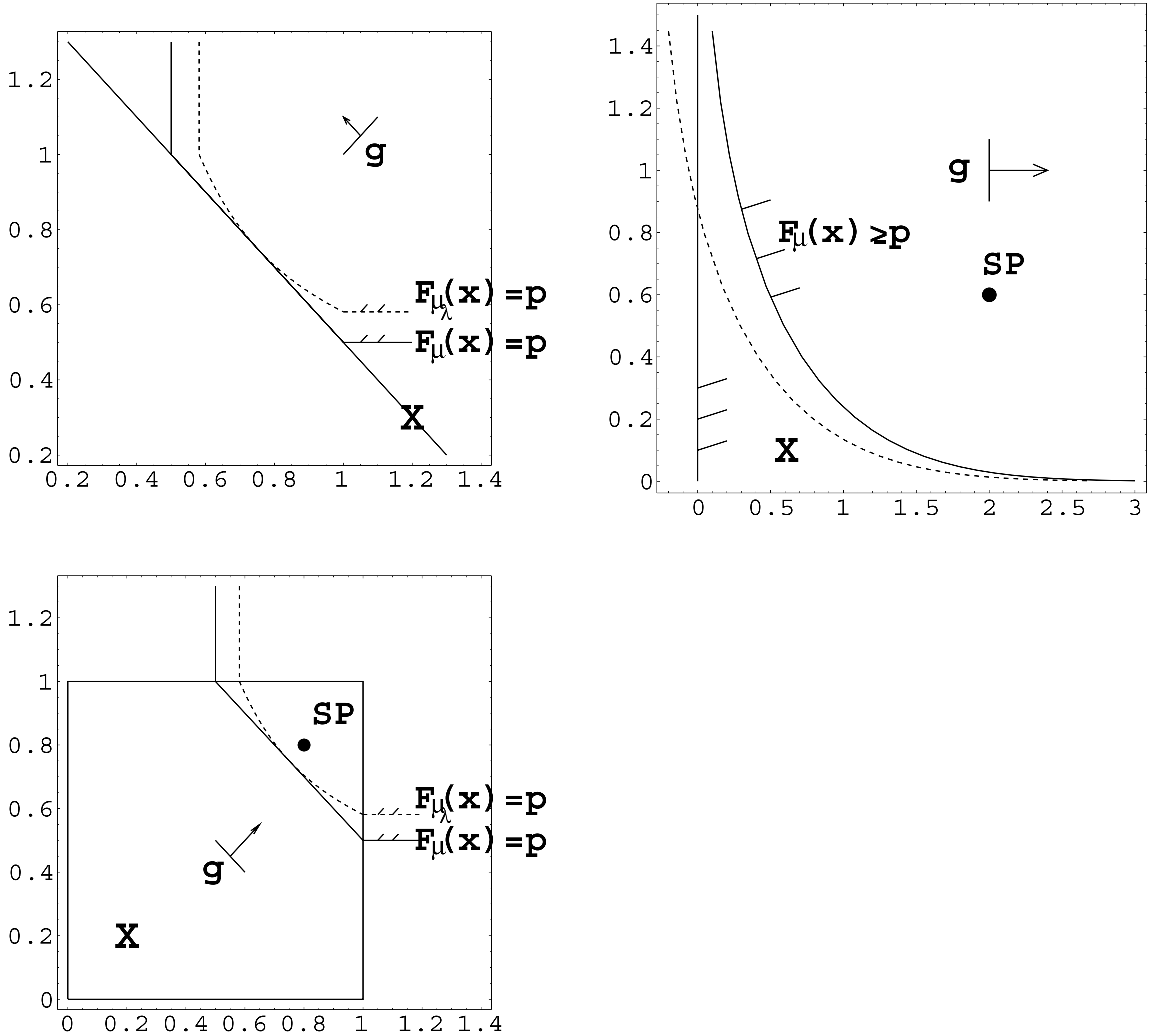

Figure 2: Illustration of original and perturbed constraint sets (dashed boundaries) in the examples 4.1 (top left), 4.3 (top right), and 4.5 (bottom left). 'SP' refers to Slater point.

Example 4.3 *In the program (P'), set $m = s = 2$, $h = \mathrm{id}$, $g(x_1, x_2) = x_1$, $X = \{(x_1, x_2) \mid x_1 = 0\}$, $p = 1/2$ and let $\mu \sim \mathcal{N}((0,0); I_2)$ be the bivariate standard normal distribution. Then, μ is log-concave and, hence, r-concave for all $r < 0$ (see Section 2.1). The corresponding distribution function writes as $F_\mu(x_1, x_2) = \Phi(x_1)\Phi(x_2)$, where Φ is the univariate standard normal distribution. The level line $F_\mu(x_1, x_2) = p$ is the graph of the function $x_2 = \Phi^{-1}(1/(2\Phi(x_1)))$, which asymptotically reaches the boundary of X (see Fig. 2). Obviously, the Slater condition is satisfied. However, $\Psi(\mu) = \emptyset$. On the other hand, for the perturbed measure $\nu_\varepsilon \sim \mathcal{N}((-\varepsilon, 0); I_2)$ ($\varepsilon > 0$) - where the distribution is slightly shifted to the left and the Kolmogorov distance $d_K(\mu, \nu_\varepsilon)$ becomes arbitrarily small - the level line $F_{\nu_\varepsilon}^{-1}(p)$ now intersects the boundary of X, and the perturbed solution set becomes $\Psi(\nu_\varepsilon) = [\Phi^{-1}(1/(2\Phi(\varepsilon))), \infty)$. Consequently, the upper semicontinuity of Ψ at μ is violated due to emptiness.*

Example 4.4 *In the program (P'), set* $m = s = 1$, $h(x) = g(x) = x$, $X = \mathbb{R}$, $p = 1/2$ *and let* μ *be the uniform distribution on the set* $[0,1] \cup [2,3]$. *The resulting distribution function is*

$$F_\mu(x) = \min\{\max\{\min\{\max\{0,(x/2\},1/2\},(x-1)/2\},1\}$$

The constraint set is the interval $[1,\infty)$, *hence* $\Psi(\mu) = \{1\}$ *which is nonempty and bounded. Furthermore, for* $\hat{x} = 3$, *one has* $F_\mu(\hat{x}) = 1$, *hence the 'Slater condition' is satisfied. We define a perturbed measure* ν_ε *via the density*

$$f_{\nu_\varepsilon}(x) = \begin{cases} 1/2-\varepsilon & \text{if} \quad x \in [0,1] \\ 1/2+\varepsilon & \text{if} \quad x \in [2,3] \\ 0 & \quad\quad \text{else} \end{cases}$$

The perturbed distribution function becomes

$$F_{\nu_\varepsilon}(x) = \min\{\max\{\min\{\max\{0,(1/2-\varepsilon)x\},1/2-\varepsilon\},(1/2+\varepsilon)x-1/2-3\varepsilon\},1\}.$$

This results in the perturbed solution set $\Psi(\nu_\varepsilon) = \{(1+3\varepsilon)/(1/2+\varepsilon)\}$ *which is a value always larger than 2. Clearly,* $d_K(\mu,\nu_\varepsilon) \to_{\varepsilon\downarrow 0} = 0$. *Therefore, neither* Ψ *is upper semicontinuous nor* φ *is continuous at* μ *caused by the fact that* μ *is not* r*-concave for any* $r < 0$ *(the support of the density of* μ *is not convex).*

Example 4.5 *Take the same data as in Example 4.1, but now with* $X := [0,1]^2$ *and* $g(x_1,x_2) = x_1 + x_2$. *Now, all assumptions of Theorem 3.1 including Slater's condition are satisfied (see Fig. 2), but the solution set mapping is not lower semicontinuos. Indeed,* $\Psi(\mu_\lambda) = \{(3/4,3/4)\}$ *for any* $\lambda > 0$ *and* $\Psi(\mu) = \mathrm{conv}\{(1/2,1),(1,1/2)\}$ *(compare discussion of Example 4.1).*

Example 4.6 *In the program (P'), set* $m = s = 2$, $h = \mathrm{id}$, $g(x_1,x_2) = x_2$, $X = [-1,1] \times [1/2,1]$, $p = 1/2$. *We are going to construct a probability measure* μ *on* $\mathbb{R}^2$ *such that all three assumptions of Theorem 3.1 are satisfied but* Ψ *fails to be upper Hölder continuous at* μ *with any rate. To this aim, put*

$$f_1(x) := \begin{cases} \alpha \cdot e^{1/x - x^2} & x < 0 \\ 0 & x \geq 0 \end{cases}, \quad (\alpha \text{ such that } \int_{-\infty}^{\infty} f_1(x)dx = 1).$$

$$f_2(x) := \begin{cases} 1 & x \in [0,1] \\ 0 & \text{else} \end{cases}$$

Clearly, both f_1 *and* f_2 *are log-concave probability densities on* $\mathbb{R}$. *Hence,* $f(x_1,x_2) := f_1(x_1)f_2(x_2)$ *is a log-concave probability density on* $\mathbb{R}^2$. *It is illustrated in Fig. 3 (left part). As a consequence of a Theorem by Prékopa ([11], Th. 4.2.1.), the probability measure* μ *induced by* f *is log-concave as well, in particular it is* r*-concave for all* $r < 0$ *(see section 2.1), hence the third assumption of Theorem 3.1 is met. Now, denote by* F_1, F_2 *the one-dimensional distribution functions induced by* f_1, f_2*, respectively. Obviously, the distribution function* F_μ *belonging to* μ *is given then by* $F_\mu(x_1,x_2) = F_1(x_1)F_2(x_2)$. *It is*

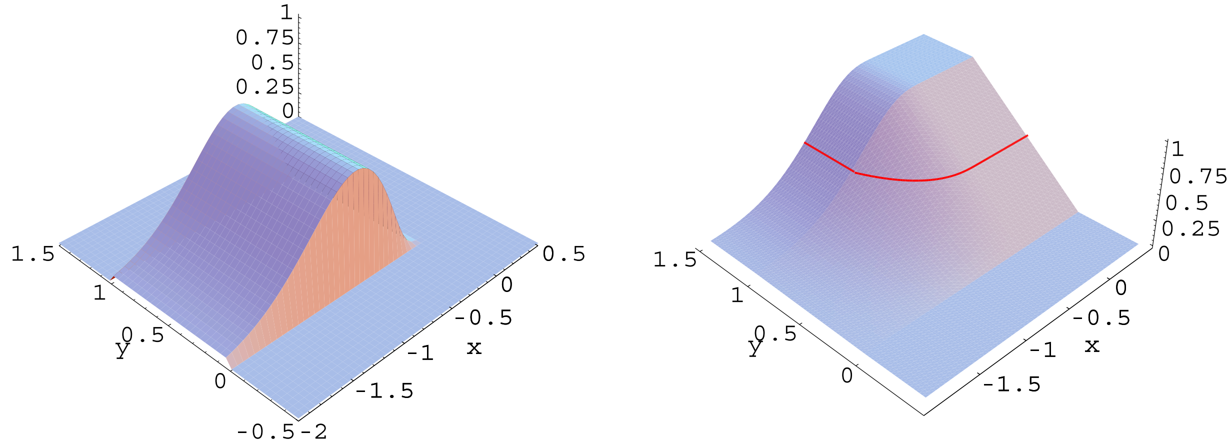

Figure 3: Illustration of the probability density (left) and of the induced distribution function (right) including the level line $F_\mu(x, y) = p$ (lifted to the graph).

illustrated in Fig. 3 (right part). Setting $c := F_1^{-1}(1/2)$ *and* $\varphi(x) := 1/2/F_1(x)$*, it is elementary to check that the constraint set is given by (see Figure 4)*

$$\{(x_1, x_2) \in [-1, 1] \times [1/2, 1] \quad | \quad \begin{array}{ll} x_1 \geq c & \textit{if} \quad x_2 \geq 1 \\ x_2 \geq \varphi(x_1) & \textit{if} \quad x_1 \geq c\}. \end{array}$$

Obviously, the solution set is $\Psi(\mu) = \{(x, 1/2) \mid x \in [0, 1]\}$*. Finally,* $(0, 3/4)$ *is a possible candidate for a Slater point (SP in Fig. 4). Summarizing, all three assumptions of Theorem 3.1 are met.*

Now, suppose that Ψ *was upper semicontinuous at* μ *with some Hölder rate* $1/k$ $(k \in N)$*. Then, there exist constants* $L, \delta > 0$ *such that*

$$\sup_{y \in \Psi(\nu)} d(y, \Psi(\mu)) \leq L[d(\mu, \nu)]^{1/k} \quad \forall \nu \in \mathcal{P}(I\!R^s), \; d(\mu, \nu) < \delta. \tag{17}$$

In order to lead this assumption to a contradiction, we may assume without loss of generality that k *is an odd number. We define perturbed probability measures* $\nu_\varepsilon \in \mathcal{P}(I\!R^s)$ *via a perturbed density by* $f^\varepsilon(x_1, x_2) = f_1(x_1) f_2^\varepsilon(x_2)$*, where*

$$f_2^\varepsilon(x) := \left\{ \begin{array}{ll} 1 & x \in [-\varepsilon, 1 - \varepsilon] \\ 0 & \textit{else} \end{array} \right. \qquad (\varepsilon > 0)$$

The induced perturbed distribution function F_{ν_ε} *satisfies* $\|F_{\nu_\varepsilon} - F_\mu\|_\infty < \varepsilon$ *and, consequently,* $d_K(\mu, \nu_\varepsilon) < \varepsilon$*. The perturbed constraint set now becomes (see dashed line in Fig.*

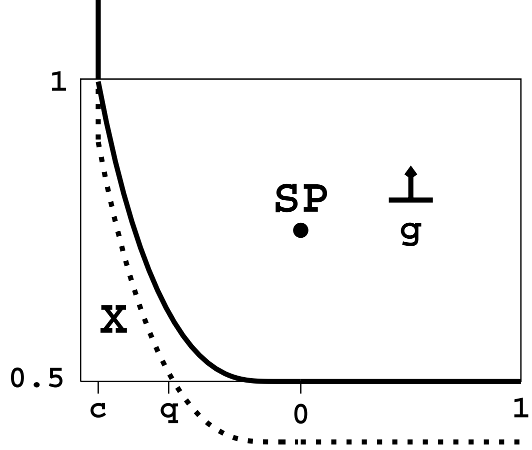

Figure 4: Illustration of the original and the perturbed constraint set: thick line = boundary of the original constraint $F_\mu(x,y) \geq p$, dashed line = boundary of the perturbed constraint $F_{\nu_\varepsilon}(x,y) \geq p$, SP = Slater point.

4)

$$\{(x_1,x_2) \in [-1,1] \times [1/2,1] \quad | \quad \begin{array}{l} x_1 \geq c \quad \text{if} \quad x_2 \geq 1-\varepsilon \\ x_2 \geq \varphi(x_1) - \varepsilon \quad \text{if} \quad x_1 \geq c\}. \end{array}$$

Accordingly, the solution set of the perturbed problem becomes $\Psi(\nu_\varepsilon) = \{(x,1/2) \mid x \in [q,1]\}$, *where* $q = \varphi^{-1}(\varepsilon + 1/2)$ *(see Fig. 4), hence* $\sup_{y \in \Psi(\nu_\varepsilon)} d(y, \Psi(\mu)) = |q|$.

Since, by definition, one has $f_1 \in \mathcal{C}^\infty(I\!R)$ *with* $f_1^{(k)}(0) = 0$ *for* $k = 0,1,2,\ldots$, *it follows that* $F_1 \in \mathcal{C}^\infty(I\!R)$ *with* $F_1^{(k)}(0) = 0$ *for* $k = 1,2,\ldots$ *and, hence,* $\varphi \in \mathcal{C}^\infty(I\!R)$ *with* $\varphi^{(k)}(0) = 0$ *for* $k = 1,2,\ldots$. *Consequently,*

$$\left((\cdot)^{k+1} + 1/2 - \varphi\right)^{(j)}(0) = \left\{ \begin{array}{ll} 0 & j = 0,\ldots,k \\ > 0 & j = k+1 \end{array} \right.$$

and one gets the relation $x^{k+1} + 1/2 \geq \varphi(x)$ *for* x *close to* 0. *In particular, we may insert the point* $x := -\varepsilon^{1/(k+1)}$ *for* ε *sufficiently close to* 0, *and it follows that* $\varphi(-\varepsilon^{1/(k+1)}) \leq \varepsilon + 1/2$ *(recall that* k *was odd). More generally, one has*

$$\varphi(x) \leq x^{k+1} + 1/2 \leq \varepsilon + 1/2 \quad \forall x \in [-\varepsilon^{1/(k+1)}, 0],$$

which implies that $q = \varphi^{-1}(\varepsilon + 1/2) < -\varepsilon^{1/(k+1)}$, *whence for all small enough* $\varepsilon > 0$ *the contradiction (see (17))*

$$\varepsilon^{1/(k+1)} < |q| = \sup_{y \in \Psi(\nu_\varepsilon)} d(y, \Psi(\mu)) \leq L[d(\mu,\nu)]^{1/k} < L\varepsilon^{1/k}$$

Acknowledgment

The author wishes to thank W. Römisch (Humboldt University Berlin) for instructive discussion which essentially influenced this paper.

References

[1] Z. Artstein, "Sensitivity with respect to the underlying information in stochastic programs," *Journal of Computational and Applied Mathematics* 56 (1994) 127-136.

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

[3] C. Borell, "Convex set functions in d-space," *Periodica Mathematica Hungarica* 6 (1975) 111-136.

[4] J. Dupačová, "Stability and Sensitivity Analysis for Stochastic Programming," *Annals of Operations Research* 27 (1990) 115-142.

[5] J. Elstrodt, *Maß- und Integrationstheorie*, Springer, Berlin, 1996.

[6] N. Gröwe, "Estimated stochastic programs with chance constraints," *European Journal of Operational Research* 101 (1997) 285-305.

[7] R. Henrion and W. Römisch, "Metric regularity and quantitative stability in stochastic programs with probabilistic constraints," *Mathematical Programming*, to appear.

[8] R. Henrion and W. Römisch, "Stability of solutions to chance constrained stochastic programs," Preprint No. 397, Weierstrass Institute Berlin, 1998.

[9] V. Kaňková, "On the Stability in Stochastic Programming: The Case of Individual Probability Constraints" *Kybernetika* 33 (1997) 525-546.

[10] D. Klatte, "A note on quantitative stability results in nonlinear optimization," in: K. Lommatzsch, ed., *Proceedings of the 19. Jahrestagung Mathematische Optimierung* (Seminarbericht No. 90, Humboldt University Berlin, Department of Mathematics, 1987) pp. 77-86.

[11] A. Prékopa, *Stochastic Programming*, Kluwer, Dordrecht, 1995.

[12] R.T.R. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970.

[13] W. Römisch and R. Schultz, "Stability analysis for stochastic programs," *Annals of Operations Research* 30 (1991) 241-266.

[14] W. Römisch and R. Schultz, "Distribution sensitivity in stochastic programming," *Mathematical Programming* 50 (1991) 197-226.

[15] R.J.-B. Wets, "Stochastic programming," in: G.L. Nemhauser, A.H.G. Rinnoy Kan and M.J. Todd eds., *Handbooks in operations research and management science. Volume 1: Optimization* (North-Holland, 1989) pp. 573-629.