

# Uncertainties in minimax stochastic programs

Jitka Dupačová

*Department of Statistics, Faculty of Mathematics and Physics  
Charles University, Prague, Czech Republic*

## Abstract

When using the minimax approach one tries to hedge against the worst possible distribution belonging to a specified class  $\mathcal{P}$ . A suitable stability analysis of results with respect to the choice of this class is an important issue. It has to be tailored to the type of the minimax problem, to the considered class of probability distributions and to the anticipated input perturbations. We shall focus on the effect of changes in input information for classes of probability distributions with support belonging to a given set and defined by (possibly perturbed) generalized moments values.

**Keywords** Minimax stochastic programs, modeling uncertainty, stability with respect to input

**AMS classification** 90C15, 90C31, 90C47

## 1 Introduction

50 years ago, stochastic programming was introduced to deal with uncertain values of coefficients which were observed in applications of linear programming. These uncertainties were modeled as random and the assumption of complete knowledge of the probability distribution  $P$  of random parameters became a standard. In practice however, complete knowledge of the probability distribution is rare. Using a hypothetical, ad hoc probability distribution  $P$  may lead to bad, costly decisions. Besides of a subsequent output analysis it pays to include the existing, possibly limited information into the model. The incomplete knowledge of  $P$  is modeled by assuming that  $P$  belongs to a specified class  $\mathcal{P}$  of probability distributions, the *ambiguity set*, and the *minimax approach* with respect to distributions belonging to the class  $\mathcal{P}$  is applied.

The minimax approach has been developed for special types of stochastic programs and special choices of the ambiguity set  $\mathcal{P}$ . To illustrate the basic ideas let us consider stochastic programs of the form

$$\text{minimize } F(\mathbf{x}, P) := E_P f(\mathbf{x}, \omega) \text{ for } \mathbf{x} \in \mathcal{X} \quad (1)$$

where  $\mathcal{X} \subseteq \mathbb{R}^n$  and  $P$  is the probability distribution of the random variable  $\omega$ . The minimax approach is applied in cases when the probability distribution  $P$  of  $\omega$  is only known to belong to a specified class  $\mathcal{P}$  of probability distributions while  $\mathcal{X}$  is assumed to be a fixed set, independent of  $P$ . To use the minimax approach means to hedge against the worst possible distribution belonging to the class  $\mathcal{P}$  by

solving the *minimax problem*

$$\min_{\mathbf{x} \in \mathcal{X}} \max_{P \in \mathcal{P}} F(\mathbf{x}, P). \quad (2)$$

The optimal value of (2) called the (upper) *minimax bound* together with the lower bound  $\min_{\mathbf{x} \in \mathcal{X}} \min_{P \in \mathcal{P}} F(\mathbf{x}, P)$  have been exploited in approximation schemes used in algorithmic procedures for solving (1). See [13, 16, 24] for an introduction and for a survey of various results. Optimal solutions of (2), called *minimax solutions*, serve as a basis for decision making. They reflect the risk aversion inherent in specific applied problems; in the energy sector, for example, social and financial consequences of blackouts are hardly tolerated and it is important to know at least the character of decisions designed to hedge against the worst possible circumstances.

To construct the class  $\mathcal{P}$  one often chooses compromises between the wish to exploit the existing, available information and the need to keep the minimax problem numerically tractable. One may rely on *sample information* to get sample moments or the empirical distribution, may use *experience* to get expert scenarios and some information about their probabilities, or to select a finite number of relevant probability distributions. Using both the sample information and experience one can make a qualified guess about the support of the probability distribution, one can incorporate a qualitative information, like symmetry or unimodality. Compact and convex classes  $\mathcal{P}$  play a key role and mathematical reasoning (& experience) may lead to a minimax problem of a manageable form, e.g. introducing bounds on probability distributions [30].

Let us mention some popular classes  $\mathcal{P}$  as given in [13] and complete the list by introducing selected recent quotations.

- $\mathcal{P}$  consists of probability distributions carried by  $\Omega \subseteq \mathbb{R}^m$  which fulfil certain *generalized moment conditions*, e.g.,

$$\mathcal{P} = \{P : E_P g_k(\omega) = y_k, k = 1, \dots, K\} \quad (3)$$

for given functions  $g_1, \dots, g_K$  and prescribed values  $y_k \forall k$ . Mostly the first and second order moments appear in (3); for a brief exposition see Section 2.

- $\mathcal{P}$  defined as above with some or all equalities replaced by inequalities. An interesting idea [6] is to identify  $\mathcal{P}$  by bounds on expectations ( $\mu$ ) and bounds on the covariance matrix, such as

$$E_P[(\omega - \mu)(\omega - \mu)^\top] \preceq \gamma \Sigma_0 \text{ for all } P \in \mathcal{P}, \quad (4)$$

and to apply approaches of semi-definite programming.

- $\mathcal{P}$  is defined as above with an additional information, such as unimodality or symmetry of  $P$  taken into account [9, 11, 23, 29];
- $\mathcal{P}$  consists of probability distributions carried by  $\Omega \subseteq \mathbb{R}^m$  with prescribed marginals [20];

- $\mathcal{P}$  consists of probability distributions  $P$  carried by a fixed *finite* set  $\Omega$ , i.e., to specify elements  $P \in \mathcal{P}$  means to fix the probabilities of the considered atoms (scenarios) taking into account a prior knowledge about their partial ordering [4] or their pertinence to an uncertainty set [32], etc.;
- $\mathcal{P}$  is a neighborhood of a (hypothetical, nominal or empirical) probability distribution  $P_0$ . This means that

$$\mathcal{P} := \{P : d(P, P_0) \leq \varepsilon\} \quad (5)$$

where  $\varepsilon > 0$  and  $d$  is a suitable distance of probability measures. Naturally, its choice influences substantially the results. See [5] for the Kullback-Leibler distance, [22, 34] for the Kantorovich distance.

- $\mathcal{P}$  consists of a finite number of probability distributions  $P_1, \dots, P_k$  and the problem is

$$\min_{\mathbf{x} \in \mathcal{X}} \max_{i=1, \dots, k} F(\mathbf{x}, P_i); \quad (6)$$

see e.g. [31].

The listed classes are not strictly separated. For example, some of moment problems, see Example 1, lead to extremal distributions carried by a finite fixed set of scenarios which does not depend on the objective function and can be thus linked with the class of distributions carried by finitely many prescribed scenarios. Also the  $\varepsilon$ -neighborhood classes (5) in [5, 22] assume discrete probability distributions concentrated at finitely many a priori fixed scenarios, or at finitely many plausible scenarios to be constructed [34]. Moreover, depending on the choice of distance  $d$  in (5),  $d(P, P_0) \leq \varepsilon$  can be treated as a generalized moment constraint. Moment conditions (3) are used in [20] to model the incomplete knowledge of marginal distributions and they are also one of ingredients for solving minimax problems under unimodality assumption or, in general, for dealing with transformed moment problems, see 2.2.3.

Anyway, minimax solutions and minimax bounds depend on specification of the class  $\mathcal{P}$ . Hence, we face an *additional level of uncertainty* which influences the results. Their robustness with respect to changes of  $\mathcal{P}$  is welcome and an output analysis is important. It has to be tailored to the type of  $\mathcal{P}$ , to the structure of the solved minimax problem and to the considered input perturbations, see e.g. [25, 31]. Refinement of minimax bounds by using an additional information has been studied from the very beginning, e.g. [2, 9, 10]. Stability and sensitivity analysis of the minimax solutions with respect to perturbations of  $\mathcal{P}$  is a more demanding task.

We shall indicate the relevant output analysis techniques mainly for the class  $\mathcal{P}$  of probability distributions identified by (generalized) moment conditions (3) and a given “carrier” set  $\Omega$ . In this case, one may exploit results of parametric optimization and some of asymptotic statistics; see e.g. [7, 12, 17] for attempts in this direction. A different approach to output analysis is needed for nonparametric

types of  $\mathcal{P}$  such as (5) with empirical distribution function  $P_0$  or for analysis of the sample counterpart of (6): Using asymptotic statistics, [22] construct nonparametric confidence sets and [31] prove consistency of results under mild assumptions.

## 2 $\mathcal{P}$ defined by moment conditions

Theoretically the so called moment problems, e.g. [1], provide bounds for the expectation  $E_P f(\mathbf{x}, \omega)$  under rather general assumptions about the function  $f(\mathbf{x}, \bullet)$  and about the considered set  $\mathcal{P}$  of probability distributions on  $\Omega$  defined by generalized moment conditions, such as (3). For a convex compact (in the weak topology) set  $\mathcal{P}$  the expectation (a linear functional in  $P$ ) attains its maximal and minimal value at an extremal point of  $\mathcal{P}$ . The corresponding extremal distributions have finite supports, however, extremal distributions *independent of the form of  $f$*  (and thus independent of the decisions  $\mathbf{x}$ ) appear only exceptionally.

In the case of an incomplete knowledge of the probability distribution  $P$  in (1) the primal interest is in estimating the difference between the maximal and minimal expectation and in evaluation of bounds  $L = \min_{\mathbf{x} \in \mathcal{X}} \inf_{P \in \mathcal{P}} E_P f(\mathbf{x}, \omega)$  and  $U = \min_{\mathbf{x} \in \mathcal{X}} \sup_{P \in \mathcal{P}} E_P f(\mathbf{x}, \omega)$  for the optimal value of (1) that can be exploited in approximations. The thorough *worst case analysis* means computing minimax solutions as well.

There is a host of papers devoted to application of moment bounds in the context of stochastic programming, e.g. [3, 16], to their refinement [2] and to inclusion of qualitative information such as unimodality and/or symmetry of  $P$  by solving transformed moment problems, cf. [9, 10, 11, 23, 29].

Throughout the paper all functions and all sets are assumed Borel measurable.

### 2.1 Basic assumptions and selected results

Let  $g_1, \dots, g_K$  be given functions on  $\Omega$  and  $f$  a given function on  $\mathcal{X} \times \Omega$  where  $\mathcal{X} \subseteq \mathbb{R}^n$ ,  $\Omega \subseteq \mathbb{R}^m$ . Let  $\mathcal{Y} := \text{conv} \{g(\Omega)\}$  and assume that  $\text{int} \mathcal{Y} \neq \emptyset$ .

For  $\mathbf{y} \in \text{int} \mathcal{Y}$  denote  $\mathcal{P}_{\mathbf{y}}$  the class of probability distributions  $P$  of random vector  $\omega$  carried by  $\Omega$  such that  $g_k, k = 1, \dots, K$ , and  $f(\mathbf{x}, \bullet)$  for all  $\mathbf{x} \in \mathcal{X}$  are integrable and the moment conditions (3)

$$E_P g_k(\omega) = y_k, k = 1, \dots, K$$

are fulfilled. The class  $\mathcal{P}_{\mathbf{y}}$  is convex and the problem is to find

$$U(\mathbf{x}, \mathbf{y}) = \sup_{P \in \mathcal{P}_{\mathbf{y}}} E_P f(\mathbf{x}, \omega), \quad \text{or } L(\mathbf{x}, \mathbf{y}) = \inf_{P \in \mathcal{P}_{\mathbf{y}}} E_P f(\mathbf{x}, \omega), \quad (7)$$

the bounds for the optimal value of (1)

$$U(\mathbf{y}) = \inf_{\mathbf{x} \in \mathcal{X}} \sup_{P \in \mathcal{P}_{\mathbf{y}}} E_P f(\mathbf{x}, \omega) \quad \text{and } L(\mathbf{y}) = \inf_{\mathbf{x} \in \mathcal{X}} \inf_{P \in \mathcal{P}_{\mathbf{y}}} E_P f(\mathbf{x}, \omega) \quad (8)$$

and to compute the minimax solutions, elements of  $\mathcal{X}(\mathbf{y}) := \arg \min_{\mathbf{x} \in \mathcal{X}} U(\mathbf{x}, \mathbf{y})$ .

Under the above basic assumptions,

$$U(\mathbf{x}, \mathbf{y}) = \inf_{\mathbf{d} \in \mathcal{D}} \left\{ d_0 + \sum_{k=1}^K d_k y_k \right\}$$

where

$$\mathcal{D} := \left\{ \mathbf{d} \in \mathbb{R}^{K+1} : d_0 + \sum_{k=1}^K d_k g_k(\mathbf{z}) \geq f(\mathbf{x}, \mathbf{z}) \forall \mathbf{z} \in \Omega \right\}.$$

In many important cases, e.g. for  $\Omega$  compact,  $g_1, \dots, g_K$  continuous with  $\text{int}\mathcal{Y} \neq \emptyset$  and for  $f(\mathbf{x}, \bullet)$  upper semicontinuous, the class  $\mathcal{P}_{\mathbf{y}}$  is for  $\mathbf{y} \in \mathcal{Y}$  nonempty, compact (in weak topology) and the supremum and the infimum in (7) is achieved; see e.g. [18]. In this case, there exists probability distribution  $P^* \in \mathcal{P}_{\mathbf{y}}$  and a vector  $\mathbf{d}^* \in \mathcal{D}$  such that

$$U(\mathbf{x}, \mathbf{y}) = E_{P^*} f(\mathbf{x}, \omega) = d_0^* + \sum_{k=1}^K d_k^* y_k \quad (9)$$

and for the given  $\mathbf{y} \in \text{int}\mathcal{Y}$ , the bound is equal to the optimal value of the corresponding semiinfinite program

$$U(\mathbf{x}, \mathbf{y}) = \inf_{\mathbf{d}} \left\{ d_0 + \sum_{k=1}^K d_k y_k : d_0 + \sum_{k=1}^K d_k g_k(\mathbf{z}) \geq f(\mathbf{x}, \mathbf{z}) \forall \mathbf{z} \in \Omega \right\}. \quad (10)$$

Evidently, as a function of the parameter  $\mathbf{y}$ ,  $U(\mathbf{x}, \mathbf{y})$  is concave.

It is important from the point of view of computation that  $P^*$  in (9) is in fact a *discrete* distribution. For these and other related results see e.g. [1, 16, 19, 24]. Similar statements hold true also for the case of inequality constraints in (3), see [3, 24].

Under additional assumptions, e.g. for  $\Omega$  a bounded convex polyhedron and

$$h(\mathbf{x}, \mathbf{z}) := f(\mathbf{x}, \mathbf{z}) - \sum_{k=1}^K d_k g_k(\mathbf{z})$$

a piecewise convex or quasi-convex function on  $\Omega$  for an arbitrary  $\mathbf{x} \in \mathcal{X}$ , (10) can be reduced to a *finite-dimensional* linear program; cf. [8]. If  $h(\mathbf{x}, \mathbf{z})$  is convex in  $\mathbf{z}$  and  $\Omega = \text{conv} \{ \mathbf{z}^{(1)}, \dots, \mathbf{z}^{(H)} \}$ , we get the set of feasible solutions of (10) in the form

$$\mathcal{D} = \left\{ \mathbf{d} \in \mathbb{R}^{K+1} : d_0 + \sum_{k=1}^K d_k g_k(\mathbf{z}^{(h)}) \geq f(\mathbf{x}, \mathbf{z}^{(h)}), h = 1, \dots, H \right\}. \quad (11)$$

### 2.1.1 Special convex case

**Assumption 1** Assume that  $f(\mathbf{x}, \bullet)$  is convex for all  $\mathbf{x} \in \mathcal{X}$ ,  $g_k(\mathbf{z}) = z_k, k = 1, \dots, K$ ,  $\Omega$  is bounded convex polyhedron  $\text{conv} \{ \mathbf{z}^{(1)}, \dots, \mathbf{z}^{(H)} \}$  and  $\mathbf{y} \in \text{int}\mathcal{Y} = \text{int}\Omega$ .

This is the favorite class  $\mathcal{P}_{\mathbf{y}}$  defined by prescribed values of the first moments. The upper bound  $U(\mathbf{x}, \mathbf{y})$  reduces to the Edmundson–Madansky bound [21] and  $U(\mathbf{x}, \mathbf{y})$  is the optimal value of the linear program dual to (10):

$$U(\mathbf{x}, \mathbf{y}) = \max_{\mathbf{p}} \left\{ \sum_{h=1}^H p_h f(\mathbf{x}, \mathbf{z}^{(h)}) : \sum_{h=1}^H p_h \mathbf{z}^{(h)} = \mathbf{y}, \sum_{h=1}^H p_h = 1, p_h \geq 0 \forall h \right\}; \quad (12)$$

$L(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}, \mathbf{y})$  according to Jensen’s inequality. For an extension to piecewise linear functions  $g_k$  and an unbounded convex closed set  $\Omega$  see [3].

In a similar way, it is possible to formulate the moment problem for probability distributions carried by a *known finite set of scenarios*, i.e. for  $\Omega = \{\omega^1, \dots, \omega^I\}$ . The probability distributions are then fully identified by probabilities  $p_i, i = 1, \dots, I$ , of these scenarios and by moment conditions. The problem to solve is

$$U(\mathbf{x}, \mathbf{y}) = \max_{\mathbf{p}} \left\{ \sum_{i=1}^I p_i f(\mathbf{x}, \omega^i) : \sum_{i=1}^I p_i g_k(\omega^i) = y_k \forall k, \sum_{i=1}^I p_i = 1, p_i \geq 0 \forall i \right\}. \quad (13)$$

This situation may occur for problems with an ad hoc given finite support of  $\omega$  or may result from identification of discrete extremal distributions as in (11) or (12), may be obtained by sampling or relying on past data and on experts’ suggestions. Additional polyhedral constraints on probabilities  $p_i$ , cf. [4, 32], can be included.

Further simplifications are possible when  $f(\mathbf{x}, \omega)$  is convex separable with respect to individual components of  $\omega$ , or  $\Omega$  is a rectangle and the components of  $\omega$  are independent, or when  $\Omega$  is a simplex. Then we may even get explicit formulas for  $U(\mathbf{x}, \mathbf{y})$ , cf. [9, 10], or obtain extremal probability distributions which do not depend on the choice of the convex random objective function  $f(\bullet, \omega)$ .

**Example 1** Assume that  $f(\mathbf{x}, \mathbf{z}) = \sum_{j=1}^m f_j(\mathbf{x}, z_j)$  where for a fixed  $\mathbf{x}$ ,  $f_j \forall j$  are convex functions of  $z_j$ ,  $\mathcal{P}_{\mathbf{y}}$  is defined by the following conditions on marginal distributions of  $\omega_j$ : They are carried by given nondegenerate compact intervals  $[a_j, b_j]$  with  $\Omega$  their Cartesian product,  $E_P \omega_j = y_j$ , with given values  $y_j \in (a_j, b_j) \forall j$ .

Then

$$U(\mathbf{x}, \mathbf{y}) = \max_{P \in \mathcal{P}_{\mathbf{y}}} E_P f(\mathbf{x}, \omega) = \sum_{j=1}^m \lambda_j f_j(\mathbf{x}, a_j) + \sum_{j=1}^m (1 - \lambda_j) f_j(\mathbf{x}, b_j) \quad (14)$$

with  $\lambda_j = (b_j - y_j)/(b_j - a_j)$ .

Moreover, for arbitrary values  $y_j \in (a_j, b_j), j = 1, \dots, m$ , the minimax solution is an efficient solution of a multicriterial problem with objective functions  $f(\mathbf{x}, \mathbf{z}), \mathbf{z} \in \mathcal{Z}^*$  where  $\mathcal{Z}^*$  is the set of vertices of  $\Omega$ , cf. [11]. Specifying the expectation  $\mathbf{y}$  we get one of these efficient solutions.

**Example 2** Consider Example 1 with  $m = 1$ , and  $f(x, z) = (x - z)^+$ , which is a part of the objective function of the newsboy problem, or simple recourse or CVaR. Let us relax conditions identifying the class  $\mathcal{P} = \mathcal{P}_{\mu, V}$  as follows: assume that the

support of  $P$  is not known precisely, but that there is a known upper bound  $V$  for the range of variation of  $\omega$  and a prescribed expectation  $E_P\omega = \mu, \forall P \in \mathcal{P}_{\mu,V}$ . Then we evidently have  $\omega \in [\mu - V, \mu + V] \forall P \in \mathcal{P}_{\mu,V}$  with probability 1 and

$$\begin{aligned} U(\mathbf{x}, \mu, V) &:= \max_{P \in \mathcal{P}_{\mu,V}} E_P(x - \omega)^+ = 0 && \text{for } x < \mu - V \\ &= x - \mu && \text{for } x > \mu + V \\ &= \frac{1}{4V}(V + x - \mu)^2 && \text{for } \mu - V \leq x \leq \mu + V. \end{aligned} \quad (15)$$

To get (15) we apply (14) for  $b = V + a$  and use the value of  $a$  for which  $\max_a \{(x - a)^+ \frac{V + a - \mu}{V} : a \in [\mu - V, \mu]\}$  is attained. This worst case  $a^* = \frac{1}{2}(x + \mu - V)$  and the maximal value is (15). The corresponding extremal probability distribution  $P^* \in \mathcal{P}_{\mu,V}$  is carried by  $a^*$  and  $a^* + V$  with probabilities  $p^* = \frac{V - \mu + x}{2V}$  and  $1 - p^*$ .

Without compactness of  $\Omega$ , existence of the optimal solution in the inner optimization problem (10) depends much on the properties of the functions  $f(\mathbf{x}, \bullet)$  and  $g_k \forall k$ . An example is  $\Omega = \mathbb{R}^m$ ,  $f(\mathbf{x}, \bullet)$  positive conical (i.e.  $f(\mathbf{x}, \mathbf{z}) > 0 \forall \mathbf{z} \neq 0$  and  $\text{epi}f(\mathbf{x}, \bullet)$  a convex polyhedral cone in  $\mathbb{R}^{m+1}$  pointed at the origin) with moment constraints on expectations and the second order moments; see [16, 24]. A special instance (cf. [9, 10]) is obtained by a direct solution of (10):

**Example 3** Assume again that  $f(x, z) = (x - z)^+$  and define  $\mathcal{P} := \mathcal{P}_{\mu, \sigma^2}$  by moment conditions  $E_P\omega = \mu$ ,  $\text{var}_P\omega = \sigma^2$  for all  $P \in \mathcal{P}_{\mu, \sigma^2}$ .

Then

$$\max_{P \in \mathcal{P}_{\mu, \sigma^2}} E_P(x - \omega)^+ = \frac{1}{2} \left( x - \mu + \sqrt{\sigma^2 + (x - \mu)^2} \right) := U(x, \mu, \sigma^2). \quad (16)$$

For a generalization of Example 3 to piecewise linear convex functions  $f(\mathbf{x}, \bullet)$  and/or  $g_k$  see [16].

Convexity properties of  $f(\mathbf{x}, \bullet)$  play an essential role. In the framework of two-stage stochastic linear programs, convexity of  $f(\mathbf{x}, \bullet)$  means a restriction to a *fixed recourse matrix*  $\mathbf{W}$ , *fixed coefficients*  $\mathbf{q}$  in the second-stage objective function and to right-hand sides  $\mathbf{h}$  and technological matrix  $\mathbf{T}$  linear in  $\omega$ . There are also parallel results for saddle functions  $f(\mathbf{x}, \bullet)$  that allow inclusion of random coefficients  $\mathbf{q}$ , e.g. [14].

## 2.2 Stability with respect to input information

The prescribed values of moments used in definition of  $\mathcal{P}_{\mathbf{y}}$  can be considered *the input information* which influences the resulting minimax bounds and minimax solutions. However, this input information is not always completely known, it can be based on a sample or past information, on expert's opinion, etc. We shall deal with stability of minimax bounds and minimax decisions under rather simplifying assumptions postponing possible generalizations.

**Assumption 2** *Assume that*

- $\mathcal{X} \subset \mathbb{R}^n$  is a nonempty convex compact set,
- $\Omega \subset \mathbb{R}^m$  is a nonempty compact set,
- $g_1, \dots, g_K$  are given continuous functions on  $\Omega$ ,
- $f : \mathcal{X} \times \Omega \rightarrow \mathbb{R}^1$  is a continuous function on  $\Omega$  for an arbitrary fixed  $\mathbf{x} \in \mathcal{X}$  and for every  $\omega \in \Omega$  it is a closed convex function of  $\mathbf{x}$ ,
- the interior of the moment space  $\mathcal{Y} := \text{conv} \{\mathbf{g}(\Omega)\}$  is nonempty.

In this section, we shall present selected applications of parametric programming to stability and sensitivity of moment bounds (7), (8) and of minimax decisions with respect to the prescribed values of moments and/or to the choice of set  $\Omega$ .

### 2.2.1 Prescribed moments values

Being concave with respect to  $\mathbf{y} \in \mathcal{Y}$  the optimal value function  $U(\mathbf{x}, \bullet)$  of the semiinfinite program (10) has *directional derivatives* on  $\text{int}\mathcal{Y}$  in all directions. For the special problems (12) and (13), stability analysis with respect to  $\mathbf{y}$  reduces to the standard stability analysis for linear programs with respect to right-hand sides and the optimal value function  $U(\mathbf{x}, \bullet)$  is concave, piecewise linear on  $\text{int}\mathcal{Y}$ .

Concerning the optimal value  $U(\mathbf{y}) = \max_{\mathbf{x} \in \mathcal{X}} U(\mathbf{x}, \mathbf{y})$  and the minimax solutions, one can exploit results on stability for nonlinear parametric programs as in [11]; e.g.:

**Theorem 1** *Under Assumption 2,  $U(\mathbf{y}) := \min_{\mathbf{x} \in \mathcal{X}} U(\mathbf{x}, \mathbf{y})$  is concave on  $\mathcal{Y}$  and the mapping  $\mathbf{y} \rightarrow \mathcal{X}(\mathbf{y})$  is upper semicontinuous on  $\mathcal{Y}$ .*

This implies again that  $U(\mathbf{y})$  has *directional derivatives* on  $\text{int}\mathcal{Y}$  in all directions. Gradients of  $U(\mathbf{y})$  exist almost everywhere on  $\text{int}\mathcal{Y}$ , nevertheless, results on *differentiability* of  $U(\mathbf{y})$  require additional smoothness assumptions, e.g. second order differentiability of  $U(\mathbf{x}, \mathbf{y})$ , and are valid under suitable regularity and second order sufficient conditions. Under such conditions, there is a unique minimax solution, say  $\mathbf{x}(\mathbf{y})$ , and  $\nabla U(\mathbf{y}) = \nabla_{\mathbf{y}} U(\mathbf{x}(\mathbf{y}), \mathbf{y})$ . However, the assumption of the second order differentiability of  $U(\mathbf{x}, \mathbf{y})$  is not always realistic. (For an example when it is fulfilled see [7] and Example 3.) Therefore, in *postoptimality analysis* of the moment bounds one can rely mainly on the results on directional differentiability, cf. [12].

### 2.2.2 Choice of the set $\Omega$

The direct analysis of explicit formulas such as (14) shows that due to a change of  $\Omega$  the upper bound function  $U(\mathbf{x}, \mathbf{y})$  may change significantly; see also Theorem 3.1.1 of [1]. The relaxation of the assumption of a known carrier  $\Omega$  as done in Example 2 leads to a rather different moment bound as well.



In the special case of probability distributions carried by a given finite set of scenarios but probabilities of their occurrence not known precisely, cf. (13), or of the class  $\mathcal{P}_{\mathbf{y}}$  determined by prescribed expectations  $\mathbf{y}$ , a convex polyhedral set  $\Omega$  and for  $f(\mathbf{x}, \bullet)$  convex, cf. (12) in section 2.1.1, the situation is relatively simple. For each fixed  $\mathbf{x} \in \mathcal{X}$  the worst case probabilities can be obtained as solutions of a linear program with a compact set of feasible solutions. Changes of scenarios or of the extremal points of  $\Omega$  influence the objective function and the matrix of coefficients of the linear program. Nevertheless, these linear programs are *stable* (small changes of the data cause only small changes of the optimal solutions, cf. [26]) provided that the sets of optimal solutions of the corresponding dual programs

$$\inf_{\mathbf{d} \in \mathcal{D}} d_0 + \sum_{k=1}^K d_k y_k \quad (17)$$

with  $\mathcal{D}$  given by

$$\mathcal{D} = \{\mathbf{d} \in \mathbb{R}^{K+1} : d_0 + \sum_{k=1}^K d_k g_k(\omega^i) \geq f(\mathbf{x}, \omega^i), i = 1, \dots, I\}$$

or

$$\mathcal{D} = \{\mathbf{d} \in \mathbb{R}^{K+1} : d_0 + \sum_{k=1}^K d_k z_k^{(h)} \geq f(\mathbf{x}, \mathbf{z}^{(h)}), h = 1, \dots, H\} \quad (18)$$

are nonempty and bounded — a condition equivalent to  $\mathbf{y} \in \text{int}\mathcal{Y}$ ; cf. [19]. The optimal value function  $U$  is then a continuous function of all coefficients on a neighborhood of the initial data  $\mathbf{x}, \mathbf{y}, \mathbf{z}^{(h)}, 1 \leq h \leq H$ , or  $\mathbf{g}(\omega^i), \forall i$ . Unique and nondegenerate optimal solution of the primal LP (13) or (12) is a special well known example. The size of the neighborhood is limited e.g. by the condition that the perturbed vector  $\mathbf{y}$  remains an interior point of the convex hull of the perturbed moment space.

Another possibility is to analyze the dual linear programs (17)–(18) allowing some uncertainty in selection of  $\omega^i$  or  $\mathbf{z}^{(h)}$ . Following [15] consider problem (12) with extremal points  $\mathbf{z}^h$  which belong to an ellipsoid around  $\mathbf{z}^{(h)}$ , say

$$\mathbf{z}^h = \mathbf{z}^{(h)} + \mathbf{E}_h \delta^h, \quad \|\delta^h\|_2 \leq \varrho, \quad (19)$$

and ask for the best solution of program (17)–(18) which is feasible for all choices of  $\mathbf{z}$  obtained by the special structure of perturbations (19). In the simplest case of  $\mathbf{E}_h = \mathbf{I}$  the  $h$ -th constraint of (18) is fulfilled if

$$\max_{\|\delta^h\|_2 \leq \varrho} \{d_0 + \mathbf{d}^\top \mathbf{z}^{(h)} + \mathbf{d}^\top \delta^h - f(\mathbf{x}, \mathbf{z}^{(h)} + \delta^h)\} \geq 0. \quad (20)$$

Lipschitz property of  $f(\mathbf{x}, \bullet)$  on the neighborhood (19) implies that there is a constant  $l$  such that

$$|f(\mathbf{x}, \mathbf{z}^h) - f(\mathbf{x}, \mathbf{z}^{(h)})| \leq l \|\delta^h\|_2 \leq l\varrho.$$

By an adaptation of results in section 5.2 of [15] to satisfy constraint (20), it is sufficient that

$$d_0 + \mathbf{d}^\top \mathbf{z}^{(h)} - f(\mathbf{x}, \mathbf{z}^{(h)}) - \varrho \sqrt{\|\mathbf{d}\|_2^2 + l^2} \geq 0. \quad (21)$$

Once more, when the optimal solution of the unperturbed linear program (17), (18) is unique and nondegenerated, then there exists  $\varrho_{\max} > 0$  such that for all problems with perturbed constraints (21) with  $0 < \varrho < \varrho_{\max}$  the optimal solutions are unique and nondegenerated, too. A similar analysis applies to (13) under suitable assumptions about the mapping  $\mathbf{g}$ .

Even for classes  $\mathcal{P}$  which do not explicitly assume a fixed known carrier  $\Omega$ , such as (4) or (5), various assumptions about  $\Omega$  are exploited in output analysis, e.g. the set  $\Omega$  is supposed to be compact convex. As a special case, existence of a ball of radius  $R$  that contains the support of the unknown probability distribution is frequently assumed where the magnitude of  $R$  may follow from “an educated and conservative guess”; cf. [6, 22].

Convergence properties can be given for finite supports which are consecutively improved to approximate the unknown support; cf. [25] and the next example.

**Example 4** Assume that  $\mathcal{P}_{\mathbf{y}}$  is the class of probability distributions on  $\Omega \subset \mathbb{R}^m$ , which fulfil the moment conditions (3) and Assumption 2 is satisfied.

Let  $\{\Omega^\nu\}_{\nu \geq 1}$  be a sequence of finite sets in  $\mathbb{R}^m$  such that  $\Omega^\nu \subseteq \Omega^{\nu+1} \subseteq \Omega$ . Choose  $\nu_0$  such that  $\mathbf{y} \in \text{int conv}\{\mathbf{g}(\Omega^{\nu_0})\}$ . For  $\nu \geq \nu_0$  consider classes  $\mathcal{P}_{\mathbf{y}}^\nu$  of probability distributions carried by  $\Omega^\nu$  for which moment conditions (3) are fulfilled. By application of Proposition 2.1 of [25] we get the following statement:

If for every  $P \in \mathcal{P}_{\mathbf{y}}$  there is a subsequence of  $\{P^\nu\}_{\nu \geq \nu_0}$ ,  $P^\nu \in \mathcal{P}_{\mathbf{y}}^\nu$  which converges weakly to  $P$ , then for  $\nu \rightarrow \infty$

$$\min_{\mathbf{x} \in \mathcal{X}} \max_{P \in \mathcal{P}_{\mathbf{y}}^\nu} E_P f(\mathbf{x}, \omega) \rightarrow \min_{\mathbf{x} \in \mathcal{X}} \max_{P \in \mathcal{P}_{\mathbf{y}}} E_P f(\mathbf{x}, \omega).$$

Moreover, also upper semicontinuity of sets of minimax solutions with respect to the considered convergence of classes  $\mathcal{P}_{\mathbf{y}}^\nu$  to  $\mathcal{P}_{\mathbf{y}}$  holds true.

### 2.2.3 Additional input information

If the class of probability distributions is defined not only by the set  $\Omega$  and the moment conditions (3) but also by other conditions such as unimodality then it is often possible to remove these conditions by a suitable transformation of probability distributions and functions and to reduce the problem to the basic moment problem. With reference to [23, 29] for more general situations, we shall delineate here only the approach for unimodal probability distributions on  $\mathbb{R}^1$  as done in [9, 10, 11].

Let  $h$  be a real function on  $\mathbb{R}^1$  that is integrable over any finite interval  $(u, M)$  and  $(M, u')$ ,  $-\infty < u < M < u' < +\infty$ . Let  $h^*$  be the transform of  $h$  defined as follows:

$$h^*(z) = \frac{1}{z - M} \int_M^z h(u) du \text{ for } z \neq M \text{ and } h^*(z) = h(z) \text{ for } z = M. \quad (22)$$

Then

$$\max_{P \in \mathcal{P}_y^M} E_P f(\mathbf{x}, \omega) = \max_P \{E_P f^*(\mathbf{x}, \omega) : E_P g_j^*(\omega) = y_j, j = 1, \dots, J\}. \quad (23)$$

Moreover, the transform  $h^*$  of a convex function  $h$  is convex. The next simple example indicates possible applications.

**Example 5** *As in Example 1 with  $m = 1$  let  $f(\mathbf{x}, z)$  be a convex function of  $z$ . Let  $\mathcal{P}_y^M$  be the class of unimodal probability distributions with the given mode  $M$ , carried by a compact interval  $[a, b]$  and such that for all  $P \in \mathcal{P}_y^M$ ,  $E_P \omega = y$ . For  $g(u) = u$  the transform (22) gives  $g^*(z) = \frac{1}{2}(z + M)$  and equation (23) then reads*

$$\max_{P \in \mathcal{P}_y^M} E_P f(\mathbf{x}, \omega) = \max_P \{E_P f^*(\mathbf{x}, \omega) : E_P g^*(\omega) = y, P(\omega \in [a, b]) = 1\} := \tilde{U}(\mathbf{x}, y, M). \quad (24)$$

Define  $\mu = 2y - M$ ; then (24) is nothing else as the usual moment problem with the class  $\mathcal{P}_\mu = \{P : E_P \omega = \mu, P(\omega \in [a, b]) = 1\}$ ; moreover, if the expectation  $y$  and the mode  $M$  coincide, then  $\mu = y$ . The transformed objective  $f^*(\mathbf{x}, z)$  is convex in  $z$ . This means that the maximal expectation  $E_P f^*(\mathbf{x}, \omega)$  over the class  $\mathcal{P}_\mu$  is

$$U(\mathbf{x}, \mu) = \lambda f^*(\mathbf{x}, a) + (1 - \lambda) f^*(\mathbf{x}, b) = \tilde{U}(\mathbf{x}, y, M)$$

with  $\lambda = \frac{b - 2y + M}{b - a} = \frac{b - \mu}{b - a}$ . Substituting for  $f^*(\mathbf{x}, z)$  according to (22) we get

$$\tilde{U}(\mathbf{x}, y, M) = \frac{b - 2y + M}{(b - a)(M - a)} \int_a^M f(\mathbf{x}, u) du + \frac{2y - M - a}{(b - a)(b - M)} \int_M^b f(\mathbf{x}, u) du;$$

it is easy to recognize two densities of uniform distributions on  $[a, M]$  and  $[M, b]$  weighted by  $\lambda$  and  $(1 - \lambda)$ , respectively.

If the mode is not known, additional maximization with respect to  $M \in [a, b]$  is possible. As a result, the worst case probability distribution is uniform on  $[a, b]$  if  $y = \frac{1}{2}(a + b)$  or is a mixture of the uniform distribution over  $[a, b]$  and the degenerated one concentrated at  $a$  or at  $b$  if  $y > \frac{1}{2}(a + b)$  or  $y < \frac{1}{2}(a + b)$ , respectively.

The bound  $\tilde{U}(\mathbf{x}, y, M)$  matches the general form of extremal distributions of  $\mathcal{P}_y^M$ , which are mixtures of uniform distributions over intervals  $(u, M)$  and  $(M, u')$  and the degenerated distribution concentrated at  $M$ .

### 3 Stability with respect to estimated moments values

Assume now that a sample information was used to estimate the moments values or other parameters which identify the class  $\mathcal{P}_y$ . Assume that these parameters were consistently estimated using a sequence of i.i.d. observations of  $\omega$ . Let  $\mathbf{y}^\nu$  be based

on the first  $\nu$  observations. Using continuity of function  $U(\mathbf{x}, \bullet)$  and theorems about transformed random variables, cf. [28], we get for consistent estimates  $\mathbf{y}^\nu$  of the true parameter  $\mathbf{y}$  the pointwise convergence

$$u^\nu(\mathbf{x}) := U(\mathbf{x}, \mathbf{y}^\nu) \rightarrow U(\mathbf{x}, \mathbf{y}) \text{ a.s.} \quad (25)$$

valid at an arbitrary element  $\mathbf{x} \in \mathcal{X}$ .

As in general the pointwise convergence does neither imply consistency of the minimax bounds, i.e. of optimal values  $U(\mathbf{y}^\nu) := \min_{\mathbf{x} \in \mathcal{X}} U(\mathbf{x}, \mathbf{y}^\nu)$ , nor consistency of minimax solutions, we shall discuss epi-convergence of the approximating objectives  $u^\nu(\mathbf{x}) := U(\mathbf{x}, \mathbf{y}^\nu)$ ; for epi-convergence consult e.g. Chapter 7 of [27].

**Definition 1** *A sequence of functions  $\{u^\nu : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}, \nu = 1, \dots\}$  is said to epi-converge to  $u : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$  if for all  $\mathbf{x} \in \mathbb{R}^n$  the two following properties hold true:*

$$\liminf_{\nu \rightarrow \infty} u^\nu(\mathbf{x}^\nu) \geq u(\mathbf{x}) \text{ for all } \{\mathbf{x}^\nu\} \rightarrow \mathbf{x} \quad (26)$$

and for some  $\{\mathbf{x}^\nu\}$  converging to  $\mathbf{x}$

$$\limsup_{\nu \rightarrow \infty} u^\nu(\mathbf{x}^\nu) \leq u(\mathbf{x}). \quad (27)$$

Whereas the pointwise convergence implies condition (27), additional assumptions are needed to get validity of condition (26). Fortunately, pointwise convergence of closed, convex functions  $u, u^\nu$  with  $\text{int dom}(u) \neq \emptyset$  implies epi-convergence; see e.g. Corollary 4 of [33]. In such case, we also have  $\limsup \{\arg \min u^\nu\} \subset \arg \min u$ .

Let us apply epi-convergence to  $U(\mathbf{x}, \mathbf{y}^\nu)$  under Assumption 2: For all fixed  $\mathbf{x} \in \mathcal{X}$ ,  $U(\mathbf{x}, \bullet)$  is concave and finite on  $\mathcal{Y}$  which implies its continuity on  $\text{int} \mathcal{Y}$ , hence, the almost sure pointwise convergence of  $u^\nu(\mathbf{x}) \rightarrow U(\mathbf{x}, \mathbf{y})$  follows. Boundedness, continuity and convexity of  $f(\mathbf{x}, \omega)$  with respect to  $\mathbf{x}$  was assumed, expectations  $E_P f(\mathbf{x}, \omega)$  are bounded convex functions on the compact set  $\mathcal{X}$  for all  $P \in \mathcal{P}_{\mathbf{y}}$  and  $u(\mathbf{x}) := U(\mathbf{x}, \mathbf{y})$ ,  $u^\nu(\mathbf{x}) := U(\mathbf{x}, \mathbf{y}^\nu)$  are closed convex functions on  $\mathcal{X}$ . Hence, the sought consistency result holds true:

**Theorem 2** *Under Assumption 2 (that implies continuity of  $U(\mathbf{x}, \mathbf{y})$  with respect to  $\mathbf{y}$ , convexity with respect to  $\mathbf{x}$ ) and under consistency of estimates  $\mathbf{y}^\nu$ , the approximate objectives  $u^\nu(\mathbf{x})$  epi-converge almost surely to  $U(\mathbf{x}, \mathbf{y})$  as  $\nu \rightarrow \infty$ . It implies that with probability 1 all cluster points of sequence of minimizers  $\mathbf{x}^\nu$  of  $u^\nu(\mathbf{x})$  on  $\mathcal{X}$  are minimizers of  $U(\mathbf{x}, \mathbf{y})$  on  $\mathcal{X}$  and  $\min_{\mathbf{x} \in \mathcal{X}} u^\nu(\mathbf{x}) \rightarrow \min_{\mathbf{x} \in \mathcal{X}} U(\mathbf{x}, \mathbf{y})$ .*

An example is the class  $\mathcal{P}_{\mathbf{y}}$  defined by moment conditions (3) and by a fixed compact set  $\Omega$  with  $f(\bullet, \mathbf{y})$  convex on a compact convex set  $\mathcal{X}$ , or the special convex problem with perturbed both  $\mathbf{y}$  and  $\Omega$  treated in Section 2.2.2.

**Example 6** *Assume that parameters  $a, b, \mu$  identifying the class of one-dimensional probability distributions on the interval  $[a, b]$  with mean value  $\mu$  are known to belong to the interior of a compact set and their values can be obtained by an estimation*

procedure based on a sample path of *i.i.d.* observations of  $\omega$  from the true probability distribution  $P$ . Their consistent estimates based on a sample size  $\nu$  are the minimal/maximal sample values and the arithmetic mean, *i.e.*  $\omega_{\nu:1}, \omega_{\nu:\nu}$  and  $\mu^\nu = \frac{1}{\nu} \sum_{i=1}^{\nu} \omega_i$ . In this case, we know the explicit form of all approximate objective functions

$$u^\nu(\mathbf{x}) := \lambda^\nu f(\mathbf{x}, \omega_{\nu:1}) + (1 - \lambda^\nu) f(\mathbf{x}, \omega_{\nu:\nu})$$

with  $\lambda^\nu = (\omega_{\nu:\nu} - \mu^\nu) / (\omega_{\nu:\nu} - \omega_{\nu:1})$ ; see Example 1 for  $J = 1$ . This is a continuous function of parameters provided that  $\omega_{\nu:1} < \omega_{\nu:\nu}$ . For convex  $f(\bullet, \omega)$ ,  $u^\nu(\mathbf{x})$  are convex in  $\mathbf{x}$ . For compact set  $\mathcal{X}$ , existence of the true minimax solution  $\mathbf{x}$  follows from continuity of  $f(\bullet, a)$  and  $f(\bullet, b)$ . Hence, the consistency statement of Theorem 2 holds true. An extension to the corresponding “data-driven” version of Example 1 is obvious.

A similar consistency result holds true also for the “data-driven” version of Example 3. Using the second order differentiability of function  $U(x, \mu, \sigma^2)$  we can get also the rate of convergence.

**Example 7** – continuation of Example 3. Assume that  $\sigma$  is known and  $\mu$  is replaced by arithmetic mean  $\mu^\nu = \frac{1}{\nu} \sum_i \omega^i$  of *i.i.d.* realizations of  $\omega$ , which is an asymptotically normal estimate.

Using differentiability of the function  $U(x, \mu, \sigma^2)$  in (16) with respect to  $\mu, x$  and theorems about distributions of transformed sequences of random variables, here of  $U(x, \mu^\nu, \sigma^2)$ , we get their asymptotically normal distribution. The rate of convergence  $O(\nu^{-1/2})$  follows from the Berry-Esséen inequality; see [7].

Additional assumptions are needed to get asymptotic normality of the estimated minimax bounds and minimax decisions.

## 4 Extensions

Up to now we have assumed that the set of feasible solutions does not depend on the probability distribution  $P$ . Let us remove this assumption and consider stochastic program

$$\text{minimize } F(\mathbf{x}, P) := E_P f(\mathbf{x}, \omega) \text{ on the set } \mathcal{X} \cap \mathcal{X}(P) \quad (28)$$

where  $\mathcal{X}$  does not depend on  $P$  and  $\mathcal{X}(P) = \{\mathbf{x} \in \mathcal{X} : G_i(\mathbf{x}, P) \leq 0, i = 1, \dots, k\}$ ; stochastic programs with probabilistic constraints or portfolio optimization with risk constraints are problems of this type. Theoretically, it is enough to deal just with one constraint  $G(\mathbf{x}, P) \leq 0$ .

Under incomplete knowledge of the probability distribution  $P$ , [22, 34] suggest to solve the “robustified” version of (28):

$$\min_{\mathbf{x} \in \mathcal{X}} \max\{F(\mathbf{x}, P) : P \in \mathcal{P}\} \quad (29)$$

subject to  $G(\mathbf{x}, P) \leq 0 \forall P \in \mathcal{P}$  or equivalently, subject to

$$\max_{P \in \mathcal{P}} G(\mathbf{x}, P) \leq 0. \quad (30)$$

Assume that  $G(\mathbf{x}, P)$  is convex in  $\mathbf{x}$  and linear in  $P$ . Then for convex, compact classes  $\mathcal{P}$  and for a fixed  $\mathbf{x}$ , the maxima in (29), (30) are attained at extremal points of  $\mathcal{P}$ ; hence for the class  $\mathcal{P}_{\mathbf{y}}$  identified by moment conditions (3) and under assumptions of section 2.1, it is possible to work with discrete distributions  $P \in \mathcal{P}$ . This property carries over also to  $G(\mathbf{x}, P)$  in (30) and/or  $F(\mathbf{x}, P)$  in (29) *convex* in  $P$ .

Whereas expected utility functions or CVaR( $\mathbf{x}, P$ ) are linear in  $P$ , other popular portfolio characteristics are even not convex in  $P$ : the variance is concave in  $P$ , mean absolute deviation is neither convex nor concave in  $P$ . This means that extensions to risk functionals nonlinear in  $P$  carry through only under special circumstances delineated in the next example.

**Example 8** Denote by  $\omega$  the random vector of unit returns of assets included to portfolio and let  $f(\mathbf{x}, \omega) = -\omega^\top \mathbf{x}$  quantify the random loss of the investment  $\mathbf{x}$ . The probability distribution  $P$  of  $\omega$  is known to belong to a class  $\mathcal{P}$  of distributions for which, *inter alia*, the expectation  $E_P \omega = \mu$  is fixed (independent of  $P$ ). Then for a fixed  $\mathbf{x}$   $\text{var}_P f(\mathbf{x}, \omega) = E_P(\omega^\top \mathbf{x})^2 - (\mu^\top \mathbf{x})^2$  is a linear functional in  $P$  and the mean absolute deviation  $\text{MAD}_P f(\mathbf{x}, \omega) = E_P |\omega^\top \mathbf{x} - \mu^\top \mathbf{x}|$  is linear in  $P$  as well.

## 5 Conclusions

The presented approach to stability analysis of minimax stochastic programs with respect to the input information was elaborated for the class  $\mathcal{P}$  defined by generalized moment conditions (3) and a given carrier set  $\Omega$ . It is suitable also for other “parametric” classes  $\mathcal{P}$  whereas stability for “nonparametric” classes, e.g. (5), would require different techniques. We did not aim at the most general statements and results on stability and sensitivity of minimax bounds and minimax decisions with respect to the model input. Specifically, various convexity assumptions were exploited, such as convexity and compactness of the class  $\mathcal{P}_{\mathbf{y}}$ , convexity of the random objective function  $f(\mathbf{x}, \omega)$  with respect to the decision variable  $\mathbf{x}$  on a compact convex set of feasible decisions, convexity of functionals  $F(\mathbf{x}, P)$ ,  $G(\mathbf{x}, P)$  with respect to the probability distribution  $P$ .

Convexity of the random objective with respect to  $\mathbf{x}$  can be replaced by a saddle property and under suitable conditions, also unbounded sets  $\mathcal{X}$  can be treated. An open question is under what general assumptions the presented approach can be applied to minimax problems with functionals nonconvex in  $P$ .

**Acknowledgments** This research is supported by the project “Methods of modern mathematics and their applications” – MSM 0021620839 and by the Grant Agency of the Czech Republic (grants 402/08/0107 and 201/09/0133).

## References

- [1] G. A. Anastassiou, *Moments in Probability and Approximation Theory*. Longman Sci. & Tech, Harlow, UK, 1993.
- [2] A. Ben-Tal and E. Hochman, *More bounds on the expectation of a convex function of a random variable*, J. Appl. Prob. 9 (1972), pp. 803–812.
- [3] J. R. Birge and R. J.-B. Wets, *Computing bounds for stochastic programming problems by means of a generalized moment problem*, Math. Oper. Res. 12 (1987), pp. 149–162.
- [4] W. Bühler, *Characterization of the extreme points of a class of special polyhedra*, Zeitschrift für Operations Research 19 (1975), pp. 131–137.
- [5] G. C. Calafiore, *Ambiguous risk measures and optimal robust portfolios*, SIAM J. Optim. 18 (2007), pp. 853–877.
- [6] E. Delage and Y. Ye, *Distributionally robust optimization under moment uncertainty with application to data-driven problems*, submitted to *Operations Research*.
- [7] J. Dupačová, *Stability in stochastic programming with recourse – Estimated parameters*, Math. Program. 28 (1984), pp. 72–83.
- [8] J. Dupačová, *Minimax stochastic programs with nonconvex nonseparable penalty functions*, in *Progress in Operations Research*, A. Prékopa, ed., J. Bolyai Math. Soc./North Holland, Budapest, 1976, pp. 303–316.
- [9] J. Dupačová, *Minimax approach to stochastic linear programming and the moment problem* (in Czech), EMO, 13 (1977), pp. 279–307; extended abstract ZAMM, 58 (1977), pp. T466–T467.
- [10] J. Dupačová, *On minimax decision rule in stochastic linear programming*, in *Studies on Mathematical Programming*, A. Prékopa, ed., Akadémiai Kiadó, Budapest, 1980, pp. 47–60.
- [11] J. Dupačová, *The minimax approach to stochastic programming and an illustrative application*, Stochastics 20 (1987), pp. 73–88.
- [12] J. Dupačová, *Moment bounds for stochastic programs in particular for recourse problems*, in: *Distributions with given Marginals and Moment Problems*, V. Beneš and J. Štěpán, eds., Kluwer Academic Publ., Dordrecht, 1997, pp. 199–204.
- [13] J. Dupačová, *Stochastic programming: minimax approach*, in *Encyclopedia of Optimization, Vol. V.*, Ch. A. Floudas and P. M. Pardalos, eds., Kluwer Academic Publ., Dordrecht 2001, pp.327–330.

- [14] N. C. P. Edirisinghe and W. T. Ziemba, *Bounds for two-stage stochastic programs with fixed recourse*, Math. Oper. Res. 19 (1994), pp. 292–313.
- [15] L. El Ghaoui, F. Oustry, and H. Lebret, *Robust solutions to uncertain semi-definite programs*, SIAM J. Optim. 9 (1998), pp.33–52.
- [16] P. Kall, *Stochastic programming with recourse: Upper bounds and moment problems - a review*, in *Advances in Mathematical Optimization*, J. Guddat, et al., eds., Akademie-Verlag, Berlin, 1988, pp. 86–103.
- [17] V. Kaňková, *On an  $\varepsilon$ -solution of minimax problem in stochastic programming*, in: *Distributions with given Marginals and Moment Problems*, V. Beneš and J. Štěpán, eds., Kluwer Academic Publ., Dordrecht, 1997, pp. 211–216.
- [18] A. F. Karr, *Extreme points of certain sets of probability measures with applications*, Math. Oper. Res. 8 (1983), pp. 74–85.
- [19] J. M. B. Kemperman, *The general moment problem, a geometric approach*, Ann. Math. Statist. 39 (1968), pp. 93–122.
- [20] W. Klein Haneveld, *Robustness against dependence in PERT: An application of duality and distributions with known marginals*, Math. Program. Study 27 (1986), pp. 153–182.
- [21] A. Madansky, *Inequalities for stochastic programming problems*, Manag. Sci. 6 (1960), pp. 197–204.
- [22] G. Pflug and D. Wozabal, *Ambiguity in portfolio selection*, Quantitative Finance 7 (2007), pp. 435–442.
- [23] I. Popescu, *A semidefinite programming approach to optimal moment bounds for convex classes of distributions*, Math. Oper. Res. 30 (2005), pp. 632–657.
- [24] A. Prékopa, *Stochastic Programming*, J. Bolyai Math. Soc./Kluwer Academic Publ., Budapest, 1995.
- [25] M. Riis and K. A. Andersen, *Applying the minimax criterion in stochastic recourse programs*, European J. of Oper. Res. 165 (2005), pp. 569–584.
- [26] S. M. Robinson, *A characterization of stability in linear programming*, Operations Research 25 (1977), pp. 435–447.
- [27] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer, Berlin, 1997.
- [28] R. J. Serfling, *Approximation Theorems in Mathematical Statistics*, Wiley, New York, 1980.
- [29] A. Shapiro, *Worst-case distribution analysis of stochastic programs*, Math. Program. B 107 (2006), pp. 91–96.



- [30] A. Shapiro and S. Ahmed, *On a class of minimax stochastic programs*, SIAM J. Optim. 14 (2004), pp. 1237–1249.
- [31] A. Shapiro and A. Kleywegt, *Minimax analysis of stochastic programs*, GOMS 17 (2002), pp. 532–542.
- [32] A. Thiele, *Robust stochastic programming with uncertain probabilities*, IMA J. of Manag. Math. 19 (2008), pp. 289–321.
- [33] R. J.-B. Wets, *Convergence of convex functions, variational inequalities and convex optimization problems*, in *Variational Inequalities and Complementarity Problems. Theory and Applications*, R. W. Cottle et al., eds., Wiley, New York, 1980, pp. 375–403.
- [34] D. Wozabal, *A framework for optimization under ambiguity*, TR2008-08, Dept of Statistics and Decision Support Systems, University of Vienna, August 2008; submitted.