On a class of minimax stochastic programs

Alexander Shapiro* **and Shabbir Ahmed†**

School of Industrial & Systems Engineering
Georgia Institute of Technology
765 Ferst Drive, Atlanta, GA 30332.

August 29, 2003

Abstract

For a particular class of minimax stochastic programming models, we show that the problem can be equivalently reformulated into a standard stochastic programming problem. This permits the direct use of standard decomposition and sampling methods developed for stochastic programming. We also show that this class of minimax stochastic programs is equivalent to a large family of mean-risk stochastic programs where risk is measured in terms of deviations from a quantile.

Key words: worst case distribution, problem of moments, Lagrangian duality, mean risk stochastic programs, deviation from a quantile.

1 Introduction

A wide variety of decision problems under uncertainty involves optimization of an expectation functional. An abstract formulation for such stochastic programming problems is

$$\operatorname*{Min}_{x \in X} \mathbb{E}_P[F(x, \omega)], \tag{1.1}$$

where $X \subseteq \mathbb{R}^n$ is the set of feasible decisions, $F : \mathbb{R}^n \times \Omega \mapsto \mathbb{R}$ is the objective function and P is a probability measure (distribution) on the space Ω equipped with a sigma algebra $\mathcal{F}$. The stochastic program (1.1) has been studied in great detail, and significant theoretical and computational progress has been achieved (see, e.g., [17] and references therein).

*Supported by the National Science Foundation under grant DMS-0073770.
†Supported by the National Science Foundation under grant DMI-0133943.

In the stochastic program (1.1) the expectation is taken with respect to the probability distribution P which is assumed to be *known*. However, in practical applications, such a distribution is not known precisely, and has to be estimated from data or constructed using subjective judgments. Often, the available information is insufficient to identify a unique distribution. In the absence of full information on the underlying distribution, an alternative approach is as follows. Suppose a set $\mathcal{P}$ of possible probability distributions for the uncertain parameters is known, then it is natural to optimize the expectation functional (1.1) corresponding to the "worst" distribution in $\mathcal{P}$. This leads to the following minimax stochastic program:

$$\operatorname*{Min}_{x \in X} \left\{ f(x) := \sup_{P \in \mathcal{P}} \mathbb{E}_P[F(x, \omega)] \right\}. \tag{1.2}$$

Theoretical properties of minimax stochastic programs have been studied in a number of publications. In that respect we can mention pioneering works of Žáčková [22] and Dupačová [3, 4], for more recent publications see [19] and references therein. These problems have also received considerable attention in the context of bounding and approximating stochastic programs [1, 7, 9]. A number of authors have proposed numerical methods for minimax stochastic program. Ermoliev, Gaivoronsky and Nedeva [5] proposed a method based on the stochastic quasigradient algorithm and generalized linear programming. A similar approach along with computational experience is reported in [6]. Breton and El Hachem [2] developed algorithms based on bundle methods and subgradient optimization. Riis and Andersen [15] proposed a cutting plane algorithm. Takriti and Ahmed [21] considered minimax stochastic programs with binary decision variables arising in power auctioning applications, and developed a branch-and-cut scheme. All of the above numerical methods require explicit solution of the inner optimization problem $\sup_{P \in \mathcal{P}} \mathbb{E}_P[F(x, \omega)]$ corresponding to the candidate solution x in each iteration. Consequently, such approaches are inapplicable in situations where calculation of the respective expectations numerically is infeasible because the set Ω although finite is prohibitively large, or possibly infinite.

In this paper, we show that a fairly general class of minimax stochastic programs can be equivalently reformulated into standard stochastic programs (involving optimization of expectation functionals). This permits a direct application of powerful decomposition and sampling methods that have been developed for standard stochastic programs in order to solve large-scale minimax stochastic programs. Furthermore, the considered class of minimax stochastic programs is shown to subsume a large family of mean-risk stochastic programs, where the risk is measured in terms of deviations from a quantile.

2 The problem of moments

In this section we discuss a variant of the problem of moments. This will provide us with basic tools for the subsequent analysis of minimax stochastic programs.

Let us denote by $\mathcal{X}$ the (linear) space of all finite signed measures on $(\Omega, \mathcal{F})$. We say that a measure $\mu \in \mathcal{X}$ is nonnegative, and write $\mu \succeq 0$, if $\mu(A) \geq 0$ for any $A \in \mathcal{F}$. For two measures $\mu_1, \mu_2 \in \mathcal{X}$ we write $\mu_2 \succeq \mu_1$ if $\mu_2 - \mu_1 \succeq 0$. That is, $\mu_2 \succeq \mu_1$ if $\mu_2(A) \geq \mu_1(A)$ for any $A \in \mathcal{F}$. It is said that $\mu \in \mathcal{X}$ is a *probability* measure if $\mu \succeq 0$ and $\mu(\Omega) = 1$. For given nonnegative measures $\mu_1, \mu_2 \in \mathcal{X}$ consider the set

$$\mathcal{M} := \big\{\mu \in \mathcal{X} : \mu_1 \preceq \mu \preceq \mu_2\big\}. \tag{2.1}$$

Let $\varphi_i(\omega)$, $i = 0, ..., q$, be real valued measurable functions on $(\Omega, \mathcal{F})$ and $b_i \in \mathbb{R}$, $i = 1, ..., q$, be given numbers. Consider the problem

$$\begin{array}{ll} \mathrm{Max}_{P \in \mathcal{M}} & \int_\Omega \varphi_0(\omega) dP(\omega) \\ \text{subject to} & \int_\Omega dP(\omega) = 1, \\ & \int_\Omega \varphi_i(\omega) dP(\omega) = b_i, \ i = 1, ..., r, \\ & \int_\Omega \varphi_i(\omega) dP(\omega) \leq b_i, \ i = r+1, ..., q. \end{array} \tag{2.2}$$

In the above problem, the first constraint implies that the optimization is performed over probability measures, the next two constraints represent moment restrictions, and the set $\mathcal{M}$ represents upper and lower bounds on the considered measures. If the constraint $P \in \mathcal{M}$ is replaced by the constraint $P \succeq 0$, then the above problem (2.2) becomes the classical problem of moments (see, e.g., [12],[20] and references therein). As we shall see, however, the introduction of lower and upper bounds on the considered measures makes the above problem more suitable for an application to minimax stochastic programming.

We make the following assumptions throughout this section:

(A1) The functions $\varphi_i(\omega)$, $i = 0, ..., q$, are μ_2-integrable, i.e.,

$$\int_\Omega |\varphi_i(\omega)| d\mu_2(\omega) < \infty, \ \ i = 0, ..., q.$$

(A2) The feasible set of problem (2.2) is nonempty, and, moreover, there exists a probability measure $P^* \in \mathcal{M}$ satisfying the equality constraints as well as the inequality constraints as equalities, i.e.,

$$\int_\Omega \varphi_i(\omega) dP^*(\omega) = b_i, \quad i = 1, ..., q.$$

Assumption (A1) implies that $\varphi_i(\omega)$, $i = 0, ..., q$, are P-integrable with respect to all measures $P \in \mathcal{M}$, and hence problem (2.2) is well defined. By assumption (A2), we can make the following change of variables $P = P^* + \mu$, and hence to write problem (2.2) in the form

$$\begin{array}{ll} \mathrm{Max}_{\mu \in \mathcal{M}^*} & \int_\Omega \varphi_0(\omega) dP^*(\omega) + \int_\Omega \varphi_0(\omega) d\mu(\omega) \\ \text{subject to} & \int_\Omega d\mu(\omega) = 0, \\ & \int_\Omega \varphi_i(\omega) d\mu(\omega) = 0, \ i = 1, ..., r, \\ & \int_\Omega \varphi_i(\omega) d\mu(\omega) \leq 0, \ i = r+1, ..., q, \end{array} \tag{2.3}$$

where

$$\mathcal{M}^* := \{\mu \in \mathcal{X} : \mu_1^* \preceq \mu \preceq \mu_2^*\} \tag{2.4}$$

with $\mu_1^* := \mu_1 - P^*$ and $\mu_2^* := \mu_2 - P^*$.

The Lagrangian of problem (2.3) is

$$L(\mu, \lambda) := \int_\Omega \varphi_0(\omega) dP^*(\omega) + \int_\Omega \mathcal{L}_\lambda(\omega) d\mu(\omega), \tag{2.5}$$

where

$$\mathcal{L}_\lambda(\omega) := \varphi_0(\omega) - \lambda_0 - \sum_{i=1}^{q} \lambda_i \varphi_i(\omega), \tag{2.6}$$

and the (Lagrangian) dual of (2.3) is the following problem:

$$\begin{array}{ll} \mathrm{Min}_{\lambda \in \mathbb{R}^{q+1}} & \{\psi(\lambda) := \sup_{\mu \in \mathcal{M}^*} L(\mu, \lambda)\} \\ \text{subject to} & \lambda_i \geq 0,\ i = r+1, ..., q. \end{array} \tag{2.7}$$

It is straightforward to see that

$$\psi(\lambda) = \int_\Omega \varphi_0(\omega) dP^*(\omega) + \int_\Omega [\mathcal{L}_\lambda(\omega)]_+ d\mu_2^*(\omega) - \int_\Omega [-\mathcal{L}_\lambda(\omega)]_+ d\mu_1^*(\omega), \tag{2.8}$$

where $[a]_+ := \max\{a, 0\}$.

By the standard theory of Lagrangian duality we have that the optimal value of problem (2.3) is always less than or equal to the optimal value of its dual (2.7). It is possible to give various regularity conditions (constraint qualifications) ensuring that the optimal values of problem (2.3) and its dual (2.7) are equal to each other, i.e., that there is no duality gap between problems (2.3) and (2.7). For example, we have (by the theory of conjugate duality, [16]) that there is no duality gap between (2.3) and (2.7), and the set of optimal solutions of the dual problem is nonempty and bounded, if and only if the following assumption holds:

(A3) The optimal value of (2.2) is finite, and there exists a feasible solution to (2.2) for all sufficiently small perturbations of the right hand sides of the (equality and inequality) constraints.

We may refer to [18] (and references therein) for a discussion of constraint qualifications ensuring the "no duality gap" property in the problem of moments.

By the above discussion we have the following result.

Proposition 1 *Suppose that the assumptions* (A1)–(A3) *hold. Then problems* (2.2) *and* (2.3) *are equivalent and there is no duality gap between problem* (2.3) *and its dual* (2.7).

Remark 1 The preceding analysis simplifies considerably if the set Ω is finite, say $\Omega := \{\omega_1, ..., \omega_K\}$. Then a measure $P \in \mathcal{X}$ can be identified with vector

$p = (p_1, ..., p_K) \in \mathbb{R}^K$. We have, of course, that $P \succeq 0$ if and only if $p_k \geq 0$, $k = 1, ..., K$. The set $\mathcal{M}$ can be written in the form

$$\mathcal{M} = \left\{ p \in \mathbb{R}^K : \mu_k^1 \leq p_k \leq \mu_k^2,\ k = 1, ..., K \right\},$$

for some numbers $\mu_k^2 \geq \mu_k^1 \geq 0$, $k = 1, ..., K$, and problems (2.2) and (2.3) become linear programming problems. In that case the optimal values of problem (2.2) (problem (2.3)) and its dual (2.7) are equal to each other by the standard linear programming duality without a need for constraint qualifications, and the assumption (A3) is superfluous.

Let us now consider, further, a specific case of (2.2) where

$$\mathcal{M} := \left\{ \mu \in \mathcal{X} : (1 - \varepsilon_1) P^* \preceq \mu \preceq (1 + \varepsilon_2) P^* \right\}, \tag{2.9}$$

i.e., $\mu_1 = (1-\varepsilon_1)P^*$ and $\mu_2 = (1+\varepsilon_2)P^*$, for some reference probability measure P^* satisfying assumption (A2) and numbers $\varepsilon_1 \in [0, 1]$, $\varepsilon_2 \geq 0$. In that case the dual problem (2.7) takes the form:

$$\begin{array}{ll} \mathrm{Min}_{\lambda \in \mathbb{R}^{q+1}} & \mathbb{E}_{P^*} \left\{ \varphi_0(\omega) + \eta_{\varepsilon_1, \varepsilon_2} \left[\mathcal{L}_\lambda(\omega) \right] \right\} \\ \text{subject to} & \lambda_i \geq 0,\ i = r+1, ..., q, \end{array} \tag{2.10}$$

where $\mathcal{L}_\lambda(\omega)$ is defined in (2.6) and

$$\eta_{\varepsilon_1, \varepsilon_2}[a] := \begin{cases} -\varepsilon_1 a, & \text{if } a \leq 0, \\ \varepsilon_2 a, & \text{if } a > 0. \end{cases} \tag{2.11}$$

Note that the function $\eta_{\varepsilon_1, \varepsilon_2}[\cdot]$ is convex piecewise linear and $\mathcal{L}_\lambda(\omega)$ is affine in λ for every $\omega \in \Omega$. Consequently the objective function of (2.10) is convex in λ. Thus, the problem of moments (2.2) has been reformulated as a convex stochastic programming problem (involving optimization of the expectation functional) of the form (1.1).

3 A class of minimax stochastic programs

We consider a specific class of minimax stochastic programming problems of the form

$$\mathrm{Min}_{x \in X} f(x), \tag{3.1}$$

where $f(x)$ is the optimal value of the problem:

$$\begin{array}{ll} \mathrm{Max}_{P \in \mathcal{M}} & \int_\Omega F(x, \omega) dP(\omega) \\ \text{subject to} & \int_\Omega dP(\omega) = 1, \\ & \int_\Omega \varphi_i(\omega) dP(\omega) = b_i,\ i = 1, ..., r, \\ & \int_\Omega \varphi_i(\omega) dP(\omega) \leq b_i,\ i = r+1, ..., q, \end{array} \tag{3.2}$$

and $\mathcal{M}$ is defined as in (2.9). Of course, this is a particular form of the minimax stochastic programming problem (1.2) with the set $\mathcal{P}$ formed by probability measures $P \in \mathcal{M}$ satisfying the corresponding moment constraints.

We assume that the set X is nonempty and assumptions (A1)–(A3), of section 2, hold for the functions $\varphi_i(\cdot)$, $i = 1, ..., q$, and $\varphi_0(\cdot) := F(x, \cdot)$ for all $x \in X$. By the analysis of section 2 (see Proposition 1 and dual formulation (2.10)) we have then that the minimax problem (3.1) is equivalent to the stochastic programming problem:

$$\begin{array}{ll} \mathrm{Min}_{(x,\lambda)\in\mathbb{R}^{n+q+1}} & \mathbb{E}_{P^*}[H(x,\lambda,\omega)] \\ \text{subject to} & x \in X \ \text{ and } \ \lambda_i \geq 0,\ i = r+1, ..., q, \end{array} \tag{3.3}$$

where

$$H(x,\lambda,\omega) := F(x,\omega) + \eta_{\varepsilon_1,\varepsilon_2}\left[F(x,\omega) - \lambda_0 - \textstyle\sum_{i=1}^{q} \lambda_i\varphi_i(\omega)\right]. \tag{3.4}$$

Note that by reformulating the minimax problem (3.1) into problem (3.3), which is a standard stochastic program involving optimization of an expectation functional, we avoid explicit solution of the inner maximization problem with respect to the probability measures. The reformulation, however, introduces $q + 1$ additional variables.

For problems with a prohibitively large (or possibly infinite) support Ω, a simple but effective approach to attack (3.3) is the *sample average approximation* (SAA) method. The basic idea of this approach is to replace the expectation functional in the objective by a sample average function and to solve the corresponding SAA problem. Depending on the structure of the objective function $F(x, \omega)$ and hence $H(x, \lambda, \omega)$, a number of existing stochastic programming algorithms can be applied to solve the obtained SAA problem. Under mild assumptions, the SAA method has been shown to have attractive convergence properties. For example, a solution to the SAA problem quickly converges to a solution to the true problem as the sample size N is increased. Furthermore, by repeated solutions of the SAA problem, statistical confidence intervals on the quality of the corresponding SAA solutions can be obtained. Detailed discussion of the SAA method can be found in [17, Chapter 6] and references therein.

3.1 Stochastic programs with convex objectives

In this section, we consider minimax stochastic programs (3.1) corresponding to stochastic programs where the objective function is convex. Note that if the function $F(\cdot, \omega)$ is convex for every $\omega \in \Omega$, then the function $f(\cdot)$, defined as the optimal value of (3.2), is given by the maximum of convex functions and hence is convex. Not surprisingly, the reformulation preserves convexity.

Proposition 2 *Suppose that the function $F(\cdot, \omega)$ is convex for every $\omega \in \Omega$. Then for any $\varepsilon_1 \in [0, 1]$ and $\varepsilon_2 \geq 0$ and every $\omega \in \Omega$, the function $H(\cdot, \cdot, \omega)$ is convex and*

$$\partial H(x,\lambda,\omega) = \begin{cases} (1-\varepsilon_1)\partial F(x,\omega) \times \{\varepsilon_1\varphi(\omega)\}, & \text{if } N(x,\lambda,\omega) < 0, \\ (1+\varepsilon_2)\partial F(x,\omega) \times \{-\varepsilon_2\varphi(\omega)\}, & \text{if } N(x,\lambda,\omega) > 0, \\ \cup_{\tau\in[-\varepsilon_1,\varepsilon_2]}(1+\tau)\partial F(x,\omega) \times \{-\tau\varphi(\omega)\}, & \text{if } N(x,\lambda,\omega) = 0, \end{cases} \tag{3.5}$$

where the subdifferentials $\partial H(x,\lambda,\omega)$ and $\partial F(x,\omega)$ are taken with respect to (x,λ) and x, respectively, and

$$N(x,\lambda,\omega) := F(x,\omega) - \lambda_0 - \sum_{i=1}^{q} \lambda_i \varphi_i(\omega), \quad \varphi(\omega) := (1, \varphi_1(\omega), ..., \varphi_q(\omega)).$$

Proof. Consider the function $\psi(z) := z + \eta_{\varepsilon_1,\varepsilon_2}[z]$. We can write

$$H(x,\lambda,\omega) = \psi\big(V(x,\lambda,\omega)\big) + \lambda_0 + \sum_{i=1}^{q} \lambda_i \varphi_i(\omega),$$

where $V(x,\lambda,\omega) := F(x,\omega) - \lambda_0 - \sum_{i=1}^{q} \lambda_i \varphi_i(\omega)$. The function $V(x,\lambda,\omega)$ is convex in (x,λ), and for $\varepsilon_1 \in [0,1]$ and $\varepsilon_2 \geq 0$, the function $\psi(z)$ is monotonically nondecreasing and convex. Convexity of $H(\cdot,\cdot,\omega)$ then follows. The subdifferential formula (3.5) is obtained by the chain rule. ■

Let us now consider instances of (3.3) with a finite set of realizations of ω:

$$\begin{array}{ll} \mathrm{Min}_{(x,\lambda)\in\mathbb{R}^{n+q+1}} & \Big\{h(x,\lambda) := \sum_{k=1}^{K} p_k^* H(x,\lambda,\omega_k)\Big\} \\ \text{subject to} & x \in X \ \text{ and } \ \lambda_i \geq 0,\ i = r+1, ..., q, \end{array} \tag{3.6}$$

where $\Omega = \{\omega_1, \ldots, w_K\}$ and $P^* = (p_1^*, ..., p_K^*)$. The above problem can either correspond to a problem with finite support of ω or may be obtained by sampling as in the SAA method. Problem (3.6) has a nonsmooth convex objective function, and often can be solved by using cutting plane or bundle type methods that use subgradient information (see, e.g., [8]). By the Moreau-Rockafellar theorem we have that

$$\partial h(x,\lambda) = \sum_{k=1}^{K} p_k^* \partial H(x,\lambda,\omega_k), \tag{3.7}$$

where all subdifferentials are taken with respect to (x,λ). Together with (3.5) this gives a formula for a subgradient of $h(\cdot,\cdot)$, given subgradient information for $F(\cdot,\omega)$.

3.2 Two-stage stochastic programs

A wide variety of stochastic programs correspond to optimization of the expected value of a future optimization problem. That is, let $F(x,\omega)$ be defined as the optimal value function

$$F(x,\omega) := \operatorname*{Min}_{y\in Y(x,\omega)} G_0(x,y,\omega), \tag{3.8}$$

where

$$Y(x,\omega) := \{y \in Y : G_i(x,y,\omega) \leq 0,\ i = 1, ..., m\}, \tag{3.9}$$

Y is a nonempty subset of a finite dimensional vector space and $G_i(x, y, \omega)$, $i = 0, ..., m$, are real valued functions. Problem (1.1), with $F(x, \omega)$ given in the form (3.8), is referred to as a two-stage stochastic program, where the first-stage variables x are decided prior to the realization of the uncertain parameters, and the second-stage variables y are decided after the uncertainties are revealed. The following result show that a minimax problem corresponding to a two-stage stochastic program is itself a two-stage stochastic program.

Proposition 3 *If $F(x, \omega)$ is defined as in* (3.8)*, then the function $H(x, \lambda, \omega)$, defined in* (3.4)*, is given by*

$$H(x, \lambda, \omega) = \inf_{y \in Y(x,\omega)} \mathcal{G}(x, \lambda, y, \omega), \tag{3.10}$$

where

$$\mathcal{G}(x, \lambda, y, \omega) := G_0(x, y, \omega) + \eta_{\varepsilon_1, \varepsilon_2}\left[G_0(x, y, \omega) - \lambda_0 - \sum_{i=1}^{q} \lambda_i \varphi_i(\omega)\right]. \tag{3.11}$$

Proof. The result follows by noting that

$$\mathcal{G}(x, \lambda, y, \omega) = \psi\left(G_0(x, y, \omega) - \lambda_0 - \sum_{i=1}^{q} \lambda_i \varphi_i(\omega)\right) + \lambda_0 + \sum_{i=1}^{q} \lambda_i \varphi_i(\omega),$$

and the function $\psi(z) := z + \eta_{\varepsilon_1, \varepsilon_2}[z]$ is monotonically nondecreasing for $\varepsilon_1 \leq 1$ and $\varepsilon_2 \geq 0$. ∎

By the above result, if the set $\Omega := \{\omega_1, ..., \omega_K\}$ is finite, then the reformulated minimax problem (3.3) can be written as one large-scale optimization problem:

$$\begin{array}{ll} \mathrm{Min}_{x, \lambda, y_1, ..., y_K} & \sum_{k=1}^{K} p_k^* \mathcal{G}(x, \lambda, y_k, \omega_k) \\ \text{subject to} & y_k \in Y(x, \omega_k),\ k = 1, ..., K, \\ & x \in X,\ \lambda_i \geq 0,\ i = r+1, ..., q. \end{array} \tag{3.12}$$

A particularly important case of two-stage stochastic programs are the two-stage stochastic (mixed-integer) linear programs, where $F(x, \omega) := V(x, \xi(\omega))$ and $V(x, \xi)$ is given by the optimal value of the problem:

$$\begin{array}{ll} \mathrm{Min}_{y} & c^T x + q^T y, \\ \text{subject to} & Wy = h - Tx,\ y \in Y. \end{array} \tag{3.13}$$

Here $\xi := (q, W, h, T)$ represents the uncertain (random) parameters of problem (3.13), and X and Y are defined by linear constraints (and possibly with integrality restrictions). By applying standard linear programming modelling principles to the piecewise linear function $\eta_{\varepsilon_1, \varepsilon_2}$, we obtain that $H(x, \lambda, \xi(\omega))$ is given by the optimal value of the problem:

$$\begin{array}{ll} \mathrm{Min}_{y, u^+, u^-} & c^T x + q^T y + \varepsilon_1 u^- + \varepsilon_2 u^+ \\ \text{subject to} & Wy = h - Tx, \\ & u^+ - u^- = c^T x + q^T y - \varphi^T \lambda, \\ & y \in Y,\ u^+ \geq 0,\ u^- \geq 0, \end{array} \tag{3.14}$$

where $\varphi := (1, \varphi_1(\omega), \ldots, \varphi_q(\omega))^T$. As before, if the set $\Omega := \{\omega_1, ..., \omega_K\}$ is finite, then the reformulated minimax problem (3.3) can be written as one large-scale mixed-integer linear program:

$$\begin{array}{ll} \mathrm{Min}_{x,\lambda,y,u^+,u^-} & c^T x + \sum_{k=1}^K p_k^* \left(q_k^T y_k + \varepsilon_1 u_k^- + \varepsilon_2 u_k^+\right) \\ \text{subject to} & W_k y_k = h_k - T_k x,\ k = 1, \ldots, K, \\ & u_k^+ - u_k^- = c^T x + q_k^T y_k - \varphi_k^T \lambda,\ k = 1, \ldots, K, \\ & y_k \in Y,\ u_k^+ \geq 0,\ u_k^- \geq 0,\ k = 1, \ldots, K, \\ & x \in X. \end{array} \tag{3.15}$$

The optimization model stated above has a block-separable structure which can, in principle, be exploited by existing decomposition algorithms for stochastic (integer) programs. In particular, if Y does not have any integrality restrictions, then the L-shaped (or Benders) decomposition algorithm and its variants can be immediately applied (see, e.g., [17, Chapter 3]).

4 Connection to a class of mean-risk models

Note that the stochastic program (1.1) is risk-neutral in the sense that it is concerned with the optimization of an expectation objective. To extend the stochastic programming framework to a risk-averse setting, one can adopt the *mean-risk* framework advocated by Markowitz and further developed by many others. In this setting the model (1.1) is extended to

$$\underset{x \in X}{\mathrm{Min}}\ \mathbb{E}[F(x, \omega)] + \gamma \mathcal{R}[F(x, \omega)], \tag{4.1}$$

where $\mathcal{R}[Z]$ is a dispersion statistic of the random variable Z used as a measure of risk, and γ is a weighting parameter to trade-off mean with risk. Classically, the variance statistic has been used as the risk-measure. However, it is known that many typical dispersion statistics, including variance, may cause the mean-risk model (4.1) to provide inferior solutions. That is, an optimal solution to the mean-risk model may be stochastically dominated by another feasible solution. Recently, Ogryczak and Ruszczynski [14] has identified a number of statistics which, when used as the risk measure $\mathcal{R}[\cdot]$ in (4.1), guarantee that the mean-risk solutions are consistent with stochastic dominance theory. One such dispersion statistic is

$$h_\alpha[Z] := \mathbb{E}\big\{\alpha[Z - \kappa_\alpha]_+ + (1 - \alpha)[\kappa_\alpha - Z]_+\big\}, \tag{4.2}$$

where $0 \leq \alpha \leq 1$ and $\kappa_\alpha = \kappa_\alpha(Z)$ denotes the α-quantile of the distribution of Z. Recall that κ_α is said to be an α-quantile of the distribution of Z if $Pr(Z < \kappa_\alpha) \leq \alpha \leq Pr(Z \leq \kappa_\alpha)$, and the set of α-quantiles forms the interval $[a, b]$ with $a := \inf\{z : Pr(Z \leq z) \geq \alpha\}$ and $b := \sup\{z : Pr(Z \geq z) \leq \alpha\}$. In particular, if $\alpha = \frac{1}{2}$, then $\kappa_\alpha(Z)$ becomes the median of the distribution of Z and

$$h_\alpha[Z] = \tfrac{1}{2}\mathbb{E}\left|Z - \kappa_{1/2}\right|,$$

and it represents half of the mean absolute deviation from the median.

In [14], it is shown that mean-risk models (4.1), with $\mathcal{R}[\cdot] := h_\alpha[\cdot]$ and $\gamma \in [0,1]$, provide solutions that are consistent with stochastic dominance theory. In the following, we show that minimax models (3.3) subsume such mean-risk models (4.1).

Consider $\mathcal{L}_\lambda(\omega)$, defined in (2.6), and $\alpha := \varepsilon_2/(\varepsilon_1+\varepsilon_2)$. Observe that, for fixed λ_i, $i = 1,...,q$, and $\varepsilon_1 > 0$, $\varepsilon_2 > 0$, a minimizer $\bar{\lambda}_0$ of $\mathbb{E}_{P^*}\eta_{\varepsilon_1,\varepsilon_2}[\mathcal{L}_\lambda(\omega)]$, over $\lambda_0 \in \mathbb{R}$, is given by an α-quantile of the distribution of the random variable $Z(\omega) := \varphi_0(\omega) - \sum_{i=1}^q \lambda_i\varphi_i(\omega)$ defined on the probability space $(\Omega, \mathcal{F}, P^*)$. In particular, if $\varepsilon_1 = \varepsilon_2$, then $\bar{\lambda}_0$ is the median of the distribution of Y. It follows that if ε_1 and ε_2 are positive, then the minimum of the expectation in (3.3), with respect to $\lambda_0 \in \mathbb{R}$, is attained at an α-quantile of the distribution of $F(x,\omega) - \sum_{i=1}^q \lambda_i\varphi_i(\omega)$ with respect to the probability measure P^*. In particular, if the moments constraints are not present in (3.2), i.e., $q = 0$, then problem (3.3) can be written as follows

$$\operatorname*{Min}_{x\in X} \mathbb{E}_{P^*}\big[F(x,\omega)\big] + (\varepsilon_1+\varepsilon_2)h_\alpha[F(x,\omega)], \tag{4.3}$$

where h_α is defined as in (4.2). The above discussion leads to the following result.

Proposition 4 *The mean-risk model* (4.1) *with* $\mathcal{R}[\cdot] := h_\alpha[\cdot]$ *is equivalent to the minimax model* (3.3) *with* $\varepsilon_1 = \gamma(1-\alpha)$, $\varepsilon_2 = \alpha\gamma$ *and* $q = 0$.

The additional term $(\varepsilon_1+\varepsilon_2)h_\alpha[F(x,\omega)]$, which appears in (4.3), can be interpreted as a regularization term. We conclude this section by discussing the effect of such regularization.

Consider the case when the function $F(\cdot,\omega)$ is convex, and piecewise linear for all $\omega \in \Omega$. This is the case, for example, when $F(x,\omega)$ is the value function of the second-stage linear program (3.13). Consider the stochastic programming problem (with respect to the reference probability distribution P^*):

$$\operatorname*{Min}_{x\in X} \mathbb{E}_{P^*}[F(x,\omega)], \tag{4.4}$$

and the corresponding mean-risk or minimax model (4.3). Suppose that X is polyhedral, the support Ω of ω is finite, and both problems (4.3) and (4.4) have finite optimal solutions. Then from the discussion at the end of Section 3, the problems (4.3) and (4.4) can be stated as linear programs. Let S_0 and $S_{\varepsilon 1,\varepsilon 2}$ denote the sets of optimal solutions of (4.4) and (4.3), respectively. Then by standard theory of linear programming, we have that, for all $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$ sufficiently small, the inclusion $S_{\varepsilon_1,\varepsilon_2} \subset S_0$ holds. Consequently, the term $(\varepsilon_1+\varepsilon_2)h_\alpha[F(x,\omega)]$ has the effect of regularizing the solution set of the stochastic program (4.4). We further illustrate this regularization with an example.

Example 1 Consider the function $F(x,\omega) := |\omega - x|$, $x, w \in \mathbb{R}$, with ω having the reference distribution $P^*(\omega = -1) = p_1^*$ and $P^*(\omega = 1) = p_2^*$ for some $p_1^* > 0$, $p_2^* > 0$, $p_1^* + p_2^* = 1$. We have then that

$$\mathbb{E}_{P^*}[F(x,\omega)] = p_1^*|1+x| + p_2^*|1-x|.$$

Let us discuss first the case where $p_1^* = p_2^* = \frac{1}{2}$. Then the set S_0 of optimal solutions of the stochastic program (4.4) is given by the interval $[-1, 1]$. For $\varepsilon_2 > \varepsilon_1$ and $\varepsilon_1 \in (0, 1)$, the corresponding α-quantile $\kappa_\alpha(F(x, \omega))$, with $\alpha := \varepsilon_2/(\varepsilon_1 + \varepsilon_2)$, is equal to the largest of the numbers $|1 - x|$ and $|1 + x|$, and for $\varepsilon_2 = \varepsilon_1$ the set of α-quantiles is given by the interval with the end points $|1 - x|$ and $|1 + x|$. It follows that, for $\varepsilon_2 \geq \varepsilon_1$, the mean-risk (or minimax) objective function in problem (4.3):

$$f(x) := \mathbb{E}_{P^*}\big[F(x, \omega)\big] + (\varepsilon_1 + \varepsilon_2) h_\alpha[F(x, \omega)],$$

is given by

$$f(x) = \begin{cases} \frac{1}{2}(1 - \varepsilon_1)|1 - x| + \frac{1}{2}(1 + \varepsilon_1)|1 + x|, & \text{if } x \geq 0, \\ \frac{1}{2}(1 + \varepsilon_1)|1 - x| + \frac{1}{2}(1 - \varepsilon_1)|1 + x|, & \text{if } x < 0. \end{cases}$$

Consequently, $S_{\varepsilon_1, \varepsilon_2} = \{0\}$. Note that for $x = 0$, the random variable $F(x, \omega)$ has minimal expected value and variance zero (with respect to the reference distribution P^*). Therefore it is not surprising that $x = 0$ is the unique optimal solution of the mean-risk or minimax problem (4.3) for any $\varepsilon_1 \in (0, 1)$ and $\varepsilon_2 > 0$.

Suppose now that $p_2^* > p_1^*$. In that case $S_0 = \{1\}$. Suppose, further, that $\varepsilon_1 \in (0, 1)$ and $\varepsilon_2 \geq \varepsilon_1$, and hence $\alpha \geq \frac{1}{2}$. Then for $x \geq 0$ the corresponding α-quantile $\kappa_\alpha(F(x, \omega))$ is equal to $|1 - x|$ if $\alpha < p_2^*$, $\kappa_\alpha(F(x, \omega)) = 1 + x$ if $\alpha > p_2^*$, and $\kappa_\alpha(x)$ can be any point on the interval $\big[|1 - x|, 1 + x\big]$ if $\alpha = p_2^*$. Consequently, for $\alpha \leq p_2^*$ and $x \geq 0$,

$$f(x) = (p_1^* + \varepsilon_2 p_1^*)(1 + x) + (p_2^* - \varepsilon_2 p_1^*)|1 - x|.$$

It follows then that $S_{\varepsilon_1, \varepsilon_2} = \{1\}$ if and only if $p_1^* + \varepsilon_2 p_1^* < p_2^* - \varepsilon_2 p_1^*$. Since $\alpha \leq p_2^*$ means that $\varepsilon_2 \leq (p_2^*/p_1^*)\varepsilon_1$, we have that for ε_2 in the interval $[\varepsilon_1, (p_2^*/p_1^*)\varepsilon_1]$, the set $S_{\varepsilon_1, \varepsilon_2}$ coincides with S_0 if and only if $\varepsilon_2 < (p_2^*/p_1^* - 1)/2$. For ε_2 in this interval we can view $\bar{\varepsilon}_2 := (p_2^*/p_1^* - 1)/2$ as the breaking value of the parameter ε_2, i.e., for ε_2 bigger than $\bar{\varepsilon}_2$ an optimal solution of the minimax problem moves from the optimal solution of the reference problem.

Suppose now that $p_2^* > p_1^*$ and $\alpha \geq p_2^*$. Then for $x \geq 0$,

$$f(x) = (p_1^* + \varepsilon_1 p_2^*)(1 + x) + (p_2^* - \varepsilon_1 p_2^*)|1 - x|.$$

In that case the breaking value of ε_1, for $\varepsilon_1 \leq (p_1^*/p_2^*)\varepsilon_2$, is $\bar{\varepsilon}_1 := (1 - p_1^*/p_2^*)/2$.

5 Numerical results

In this section, we describe some numerical experiments with the proposed minimax stochastic programming model. We consider minimax extensions of two-stage stochastic linear programs with finite support of the random problem parameters. We assume that $q = 0$ (i.e., that the moment constraints are not present in the model) since, in this case, the minimax problems are equivalent

to mean-risk extensions of the stochastic programs, where risk is measured in terms of quantile deviations.

Recall that, owing to the finiteness of the support, the minimax problems reduce to the specially structured linear programs (3.15). We use an ℓ_∞–trust-region based decomposition algorithm for solving the resulting linear programs. The method along with its theoretical convergence properties is described in [11]. The algorithm has been implemented in ANSI C with the GNU Linear Programming Kit (GLPK) [13] library routines to solve linear programming subproblems. All computations have been carried out on a Linux workstation with dual 2.4 GHz Intel Xeon processors and 2 GB RAM.

The stochastic linear programming test problems in our experiments are derived from those used in [10]. We consider the problems `LandS`, `gbd`, `20term`, and `storm`. Data for these instances are available from the website:

`http://www.cs.wisc.edu/~swright/stochastic/sampling`

These problems involve extremely large number of scenarios (joint realizations of the uncertain problem parameters). Consequently, for each problem, we consider three instances each with 1000 sampled scenarios. The reference distribution P^* for these instances correspond to equal weights assigned to each sampled scenario.

Recall that a minimax model with parameters ε_1 and ε_2 is equivalent to a mean-risk model (involving quantile deviations) with parameters $\gamma := \varepsilon_1 + \varepsilon_2$ and $\alpha := \varepsilon_2/(\varepsilon_1 + \varepsilon_2)$. We consider α values of 0.5, 0.7, and 0.9, and ε_1 values of 0.1, 0.3, 0.5, 0.7, and 0.9. Note that the values of the parameters ε_2 and γ are uniquely determined by $\varepsilon_2 = \alpha\varepsilon_1/(1-\alpha)$ and $\gamma = \varepsilon_1/(1-\alpha)$. Note also that some combinations of ε_1 and α are such that $\gamma > 1$, and consequently the resulting solutions are not guaranteed to be consistent with stochastic dominance.

First, for each problem, the reference stochastic programming models (with $\varepsilon_1 = \varepsilon_2 = 0$) corresponding to all three generated instances were solved. Next, the minimax stochastic programming models for the various ε_1-α combinations were solved for all instances. Various dispersion statistics corresponding to the optimal solutions (from the different models) with respect to the reference distribution P^* were computed. Table 1 presents the results for the reference stochastic program corresponding to the four problems. The first six rows of the table displays various cost-statistics corresponding to the optimal solution with respect to P^*. The presented data is the average over the three instances. The terms "Abs Med-Dev," "Abs Dev," "Std Dev," "Abs SemiDev," and "Std SemiDev" stand for the statistics mean absolute deviation from the median, mean absolute deviation, standard deviation, absolute semi-deviation, and standard semi-deviation, respectively. The last two rows of the table display the average (over the three instances) number of iterations and CPU seconds required. Tables 2-4, 3-7, 8-10, and 11-13 present the results for the problems `LandS`, `gbd`, `20term`, and `storm`, respectively. Each table (in the set 2-13) corresponds to a particular α value and each column in a table correspond to a particular ε_1 value. The statistics are organized in the rows as in Table 1.

For a fixed level of α, increasing ε_1 corresponds to increasing the allowed perturbation of the reference distribution in the minimax model, and to increasing the weight γ for the risk term in the mean-risk model. Consequently, we observe from the tables, that this leads to solutions with higher expected costs. We also observe that the value of some of the dispersion statistics decreases indicating a reduction in risk. Similar behavior occurs upon increasing α corresponding to a fixed level of ε_1.

A surprising observation from the numerical results, is that the considered problem instances are very robust with respect to perturbations of the reference distribution P^*. Even with large perturbations of the reference distribution, the perturbations of the objective values of the solutions are relatively small.

A final observation from the tables, is the large variability of computational effort for the various ε_1-α combinations. This can be somewhat explained by the regularization nature of minimax (or mean-risk) objective function as discussed in Section 4. For certain ε_1-α combinations, the piece-wise linear objective function may become very sharp resulting in faster convergence of the algorithm.

	`LandS`	`gbd`	`20term`	`storm`
Expected cost	225.524231	1655.544680	254147.150217	15498557.910287
Abs Med-Dev	46.631711	502.017789	10022.597583	304941.126223
Abs Dev	46.950206	539.633584	10145.862901	313915.600392
Std Dev	59.263949	715.331904	12079.769991	371207.137372
Abs SemiDev	23.475103	269.816792	5072.931451	156957.800196
Std SemiDev	44.55075	605.012796	8824.368440	261756.118948
Iterations	47.333333	57.333333	275.333333	5000.000000
CPU seconds	0.666667	0.666667	32.333333	2309.333333

Table 1: Statistics corresponding to the reference stochastic program

	$\varepsilon_1 = 0.1$	$\varepsilon_1 = 0.3$	$\varepsilon_1 = 0.5$	$\varepsilon_1 = 0.7$	$\varepsilon_1 = 0.9$
Expected cost	225.572591	225.741331	225.992122	226.394508	226.950087
Abs Med-Dev	45.907760	45.031670	44.408435	43.743833	43.040004
Abs Dev	46.239633	45.384829	44.787747	44.151339	43.465447
Std Dev	58.277697	57.158371	56.408928	55.625918	54.838189
Abs SemiDev	23.119817	22.692415	22.393873	22.075670	21.732724
Std SemiDev	43.783344	42.972997	42.477392	41.969667	41.451160
Iterations	3357.333333	3357.000000	75.000000	70.000000	67.333333
CPU seconds	196.333333	195.333333	1.000000	1.000000	1.000000

Table 2: Problem: `LandS`, $\alpha = 0.5$

	$\varepsilon_1 = 0.1$	$\varepsilon_1 = 0.3$	$\varepsilon_1 = 0.5$	$\varepsilon_1 = 0.7$	$\varepsilon_1 = 0.9$
Expected cost	225.604863	225.860693	226.308477	226.924267	227.732520
Abs Med-Dev	45.691352	44.759124	43.910281	43.143102	42.325219
Abs Dev	46.007275	45.108807	44.283083	43.536913	42.757779
Std Dev	57.941208	56.790168	55.787518	54.917470	54.011196
Abs SemiDev	23.003638	22.554403	22.141542	21.768456	21.378890
Std SemiDev	43.475330	42.691880	42.022589	41.445042	40.850585
Iterations	5000.000000	72.666667	64.666667	70.666667	68.000000
CPU seconds	293.000000	1.333333	1.000000	1.000000	1.000000

Table 3: Problem: `LandS`, $\alpha = 0.7$

	$\varepsilon_1 = 0.1$	$\varepsilon_1 = 0.3$	$\varepsilon_1 = 0.5$	$\varepsilon_1 = 0.7$	$\varepsilon_1 = 0.9$
Expected cost	225.657561	226.231321	227.158529	228.239015	228.723862
Abs Med-Dev	45.442289	44.060225	42.928119	42.065348	41.752750
Abs Dev	45.767970	44.452844	43.357957	42.497211	42.175589
Std Dev	57.607866	55.952645	54.636028	53.623360	53.263650
Abs SemiDev	22.883985	22.226422	21.678979	21.248606	21.087795
Std SemiDev	43.208970	42.131097	41.267697	40.541459	40.277103
Iterations	65.666667	63.333333	59.666667	60.000000	1700.333333
CPU seconds	1.000000	1.000000	1.000000	1.000000	95.666667

Table 4: Problem: `LandS`, $\alpha = 0.9$

	$\varepsilon_1 = 0.1$	$\varepsilon_1 = 0.3$	$\varepsilon_1 = 0.5$	$\varepsilon_1 = 0.7$	$\varepsilon_1 = 0.9$
Expected cost	1655.544680	1656.696416	1659.855302	1682.355861	1685.929064
Abs Med-Dev	502.017789	496.352624	488.853714	450.792925	446.110925
Abs Dev	539.633584	536.241586	531.229091	501.621083	498.632665
Std Dev	715.331904	711.500466	705.983292	678.924336	675.058785
Abs SemiDev	269.816792	268.120793	265.614546	250.810542	249.316332
Std SemiDev	605.012796	602.666522	599.863871	586.674962	583.868317
Iterations	79.000000	70.000000	70.666667	63.333333	66.000000
CPU seconds	1.000000	0.666667	0.333333	1.000000	1.000000

Table 5: Problem: `gbd`, $\alpha = 0.5$

	$\varepsilon_1 = 0.1$	$\varepsilon_1 = 0.3$	$\varepsilon_1 = 0.5$	$\varepsilon_1 = 0.7$	$\varepsilon_1 = 0.9$
Expected cost	1655.570090	1657.065720	1663.667081	1685.507685	1771.068164
Abs Med-Dev	501.772028	496.543213	483.940329	452.686268	396.222865
Abs Dev	539.433816	536.230530	523.960900	504.791237	442.632554
Std Dev	715.147566	711.256157	702.313425	678.743484	619.951050
Abs SemiDev	269.716908	268.115265	261.980450	252.395618	221.316277
Std SemiDev	604.913488	602.289640	598.708725	584.678270	536.764625
Iterations	75.666667	71.333333	71.666667	63.333333	60.666667
CPU seconds	1.000000	0.333333	1.000000	0.666667	0.666667

Table 6: Problem: `gbd`, $\alpha = 0.7$

	$\varepsilon_1 = 0.1$	$\varepsilon_1 = 0.3$	$\varepsilon_1 = 0.5$	$\varepsilon_1 = 0.7$	$\varepsilon_1 = 0.9$
Expected cost	1656.742634	1668.311115	1701.823473	2054.695770	2104.715572
Abs Med-Dev	497.768654	496.665323	493.446982	411.528582	416.847130
Abs Dev	536.887996	533.105354	531.749429	437.668526	434.673818
Std Dev	711.572070	703.357694	685.771387	523.260522	518.807426
Abs SemiDev	268.443998	266.552677	265.874714	218.834263	217.336909
Std SemiDev	601.247020	588.428289	568.798933	412.283174	401.524723
Iterations	82.333333	73.333333	68.000000	63.000000	65.333333
CPU seconds	1.000000	1.000000	1.000000	1.000000	0.666667

Table 7: Problem: `gbd`, $\alpha = 0.9$

	$\varepsilon_1 = 0.1$	$\varepsilon_1 = 0.3$	$\varepsilon_1 = 0.5$	$\varepsilon_1 = 0.7$	$\varepsilon_1 = 0.9$
Expected cost	254156.095167	254172.391617	254272.736467	254338.015533	254368.068283
Abs Med-Dev	9865.177167	9780.642083	9547.653600	9424.043778	9387.429861
Abs Dev	10026.802482	9935.445188	9700.126819	9550.311715	9505.935648
Std Dev	11909.728572	11827.617116	11506.971142	11323.282363	11267.261816
Abs SemiDev	5013.401241	4967.722594	4850.063410	4775.155857	4752.967824
Std SemiDev	8654.320482	8590.418211	8289.983061	8106.878737	8049.794913
Iterations	284.333333	304.333333	277.000000	279.000000	273.333333
CPU seconds	35.666667	36.000000	32.666667	33.000000	32.666667

Table 8: Problem: `20term`, $\alpha = 0.5$

	$\varepsilon_1 = 0.1$	$\varepsilon_1 = 0.3$	$\varepsilon_1 = 0.5$	$\varepsilon_1 = 0.7$	$\varepsilon_1 = 0.9$
Expected cost	254157.838900	254329.222033	254545.403950	254655.101626	254670.894863
Abs Med-Dev	9850.254233	9437.472500	9220.598261	9120.799565	9114.212673
Abs Dev	10008.715887	9566.587326	9360.187746	9345.966887	9337.985153
Std Dev	11891.616364	11344.804887	11002.470097	10958.635626	10947.621409
Abs SemiDev	5004.357943	4783.293663	4680.093873	4672.983444	4668.992577
Std SemiDev	8637.923902	8129.071863	7767.960224	7738.486426	7726.880046
Iterations	314.000000	273.666667	281.333333	275.666667	1834.000000
CPU seconds	37.666667	32.333333	34.000000	32.666667	431.333333

Table 9: Problem: `20term`, $\alpha = 0.7$

	$\varepsilon_1 = 0.1$	$\varepsilon_1 = 0.3$	$\varepsilon_1 = 0.5$	$\varepsilon_1 = 0.7$	$\varepsilon_1 = 0.9$
Expected cost	254310.491039	254508.075023	254523.789306	255016.024362	255924.359895
Abs Med-Dev	9521.522703	9330.217798	9323.709859	9457.586872	9301.167534
Abs Dev	9649.482731	9469.459564	9461.571942	9534.906547	9353.787342
Std Dev	11414.001982	11115.029895	11104.118907	11187.286221	11114.273955
Abs SemiDev	4824.741366	4734.729782	4730.785971	4767.453273	4676.893671
Std SemiDev	8174.885367	7857.867726	7846.314786	7817.800413	7648.477100
Iterations	290.000000	311.333333	351.000000	443.000000	475.666667
CPU seconds	34.000000	36.666667	41.333333	53.000000	57.666667

Table 10: Problem: `20term`, $\alpha = 0.9$

	$\varepsilon_1 = 0.1$	$\varepsilon_1 = 0.3$	$\varepsilon_1 = 0.5$	$\varepsilon_1 = 0.7$	$\varepsilon_1 = 0.9$
Expected cost	15498566.363404	15498566.363404	15499088.268081	15499180.567248	15499502.750174
Abs Med-Dev	304768.076897	304768.076898	303407.515564	303257.959564	302867.196699
Abs Dev	313743.532720	313743.532725	312276.310249	312125.258689	311544.608977
Std Dev	371020.831673	371020.831676	369656.245604	369528.782376	369239.025550
Abs SemiDev	156871.766360	156871.766363	156138.155125	156062.629345	155772.304488
Std SemiDev	261623.460914	261623.460918	260734.509580	260641.989695	260124.766710
Iterations	5000.000000	1718.666667	80.666667	1719.666667	1713.333333
CPU seconds	2344.333333	785.000000	16.666667	802.333333	793.666667

Table 11: Problem: `storm`, $\alpha = 0.5$

	$\varepsilon_1 = 0.1$	$\varepsilon_1 = 0.3$	$\varepsilon_1 = 0.5$	$\varepsilon_1 = 0.7$	$\varepsilon_1 = 0.9$
Expected cost	15498597.009358	15498912.463425	15499225.250463	15499412.008793	15501297.484860
Abs Med-Dev	304768.076896	304264.549632	303585.261410	303345.546078	303651.417922
Abs Dev	313693.250991	313144.675810	312532.808424	312284.406979	312788.985008
Std Dev	371019.219394	370329.421866	369731.473411	369444.593031	368414.762740
Abs SemiDev	156846.625495	156572.337905	156266.404212	156142.203490	156394.492504
Std SemiDev	261574.859566	260945.658373	260501.734375	260321.919740	259467.084176
Iterations	3359.666667	3358.000000	1718.333333	3358.666667	1713.333333
CPU seconds	1565.333333	1643.000000	807.333333	1574.000000	788.666667

Table 12: Problem: `storm`, $\alpha = 0.7$

	$\varepsilon_1 = 0.1$	$\varepsilon_1 = 0.3$	$\varepsilon_1 = 0.5$	$\varepsilon_1 = 0.7$	$\varepsilon_1 = 0.9$
Expected cost	15498693.299136	15499682.020559	15507795.606679	15517578.176729	15522145.320553
Abs Med-Dev	305049.775029	304785.970204	302892.249844	302161.828558	304312.662029
Abs Dev	314177.306262	313993.856585	312331.074727	310811.051642	311205.223802
Std Dev	370960.125688	370450.826583	367610.550262	366045.144715	366481.769877
Abs SemiDev	157088.653131	156996.928293	156165.537364	155405.525821	155602.611901
Std SemiDev	261570.538153	260717.002761	256621.384854	254270.119605	253850.982098
Iterations	3360.000000	5000.000000	76.333333	3362.333333	1701.666667
CPU seconds	1610.333333	2361.000000	15.333333	1623.333333	822.666667

Table 13: Problem: `storm`, $\alpha = 0.9$

References

[1] J. R. Birge and R. J.-B. Wets. Computing bounds for stochastic programming problems by means of a generalized moment problem. *Mathematics of Operations Research*, 12(1):149–162, 1987.

[2] M. Breton and S. El Hachem. Algorithms for the solution of stochastic dynamic minimax problems. *Computational Optimization and Applications*, 4:317–345, 1995.

[3] J. Dupačová. Minimax stochastic programs with nonseparable penalties. In *Optimization techniques (Proc. Ninth IFIP Conf., Warsaw, 1979), Part 1*, volume 22 of *Lecture Notes in Control and Information Sci.*, pages 157–163, Berlin, 1980. Springer.

[4] J. Dupačová. The minimax approach to stochastic programming and an illustrative application. *Stochastics*, 20:73–88, 1987.

[5] Y. Ermoliev, A. Gaivoronsky, and C. Nedeva. Stochastic optimization problems with partially known distribution functions. *SIAM Journal on Control and Optimization*, 23:697–716, 1985.

[6] A. A. Gaivoronski. A numerical method for solving stochastic programming problems with moment constraints on a distribution function. *Annals of Operations Research*, 31:347–370, 1991.

[7] H. Gassmann and W. T. Ziemba. A tight upper bound for the expectation of convex function of a multi-variate random variable. *Mathematical Programming Study*, 27:39–53, 1986.

[8] J.-B. Hiriart-Urruty and C. Lemarchal *Convex Analysis and Minimization Algorithms II.* Springer-Verlag, Berlin, 1996.

[9] P. Kall. An upper bound for SLP using first and total second moments. *Annals of Operations Research*, 30:670–682, 1991.

[10] J. Linderoth, A. Shapiro and S. Wright The empirical behavior for sampling methods for stochastic programming. *Annals of Operations Research*, to appear.

[11] J. Linderoth and S. Wright Decomposition algorithms for stochastic programming on a computational grid. *Computational Optimization and Applications*, 24:207–250, 2003.

[12] H.J. Landau (ed.) *Moments in mathematics*, Proc. Sympos. Appl. Math., 37, Amer. Math. Soc., Providence, RI, 1987.

[13] A. Makhorin. GNU Linear Progamming Kit, Reference Manual, Version 3.2.3. `http://www.gnu.org/software/glpk/glpk.html`, 2002.

[14] W. Ogryczak and A. Ruszczynski. Dual stochastic dominance and related mean-risk models. *SIAM Journal on Optimization*, 13:60–78, 2002.

[15] M. Riis and K. A. Andersen. Applying the minimax criterion in stochastic recourse programs. Technical Report 2002/4, University of Aarhus, Department of Operations Research, Aarhus, Denmark, 2002.

[16] R. T. Rockafellar. *Conjugate Duality and Optimization*. Regional Conference Series in Applied Mathematics, SIAM, Philadelphia, PA, 1974.

[17] A. Rusczynski and A. Shapiro, editors. *Stochastic Programming*, volume 10 of *Handbooks in Operations Research and Management Science*. North-Holland, 2003.

[18] A. Shapiro. On duality theory of conic linear problems. In *Semi-Infinite Programming,* M.A. Goberna and M.A. López (eds), Kluwer Academic Publishers, pages 135–165, 2001.

[19] A. Shapiro and A. Kleywegt. Minimax analysis of stochastic programs. *Optimization Methods and Software*, 17:523–542, 2002.

[20] J.E. Smith. Generalized Chebychev inequalities: Theory and applications in decision analysis. *Operations Research*, 43: 807-825, 1995.

[21] S. Takriti and S. Ahmed. Managing short-term electricity contracts under uncertainty: A minimax approach. Technical report, School of Industrial & Systems Engineering, Georgia Institute of Technology, Atlanta, GA, 2002.

[22] J. Žáčková. On minimax solution of stochastic linear programming problems. *Cas. Pest. Mat.*, 91:423–430, 1966.