Journal of Uncertain Systems
Vol.3, No.3, pp.183-191, 2009
Online at: www.jus.org.uk

Safe Approximation for Optimization with First Order Stochastic Dominance Constraints

Amir Ahmadi Javid*, Narges Kazemzadeh
Department of Industrial Engineering, Amirkabir University of Technology, Tehran, P.O. Box 15875-4413, Iran

Received 9 May 2008; Revised 5 October 2008

Abstract

Recently, there has been a significant interest in introducing stochastic dominance relations as constraints into stochastic optimization problems. Optimization with first order stochastic dominance constraints in discrete distribution case can be formulated as mixed integer programs. In this article, we present a method to safely approximate such kinds of mixed integer programs.

Keywords: stochastic dominance constraint, safe approximation, optimization

1 Introduction

The relation of stochastic dominance is a fundamental concept in statistics, decision theory, economics and finance [9, 10, 13, 14, 16, 18, 22]. This notion was originated from the theory of majorization for the discrete case [11, 17], and later extended to general distributions [20].

Recently stochastic programming under stochastic dominance constraints has been introduced in [3]-[7], [15, 19]. Most studies in this field are dedicated to the first and second order stochastic dominance relations. One of the important applications of stochastic programming with stochastic dominance constraints is in portfolio optimization. The aim of such problems is choosing investments on the available assets to maximize the expected return, under the condition that the return rate stochastically dominates a given benchmark return rate. Stochastic dominance constraints have also been used in intensity-modulated radiation therapy treatment planning. In this problem, the aim is to provide superior tumor coverage and conformity, while maintaining a low irradiation to important critical and normal tissues [12, 21]. Another application of stochastic dominance constraints can be found in risk modeling in power systems with dispersed generation [8]. In [2], stochastic dominance rules are applied to analyze the One Day International (ODI) batting performance of cricketers.

First order stochastic dominance constraints present a nonconvex feasible area [6], which makes challenges to efficiently solve optimization problems with such constraints. In [19] the authors study the problem in the discrete case and presents solutions methods based on a mixed integer formulation. In [15], a new formulation and heuristics are developed. In [1] we introduce a cutting plane method to more efficiently solve the formulation presented in [19]. In this paper, we present a method to safely approximate such optimization problems.

To continue our discussion, we need to establish some notations and definitions used throughout the paper. The triple $(\Omega, \mathbf{F}, P)$ is a probability space, where Ω is the entire space, $\mathbf{F}$ is a σ-field of subsets of Ω and P is a probability measure defined on $\mathbf{F}$. The symbol $\mathbf{X}$ denotes the space of all random variables defined on $(\Omega, \mathbf{F})$ (i.e. all finite Borel measurable functions X from Ω to $\Re$). For the random variable $X \in \mathbf{X}$, $F_X(\eta) = \Pr(X \le \eta)$, $\eta \in \Re$, is the (right-continuous) cumulative distribution function (CDF). We say that for $X, Y \in \mathbf{X}$, X dominates Y in the first order stochastic dominance, denoted by $X \succeq_{(1)} Y$, if $F_X(\eta) \le F_Y(\eta)$ for all $\eta \in \Re$.

Generally, we can define stochastic optimization model with a first order stochastic dominance constraint as

$$\begin{aligned} &\max \ f(X) \\ &\text{s.t.} \quad X \succeq_{(1)} Y,\ X \in \mathbf{C} \end{aligned} \tag{1}$$

* Corresponding author. Email: ahmadi_ javid@ aut.ac.ir (A. Ahmadi Javid).

where $\mathbf{C} \subseteq \mathbf{X}$, $f : \mathbf{C} \to \Re$, and Y is a random variable. In practice, Y is an available reference outcome, and our intention is to have the new outcome X preferable over Y in the sense of the first order stochastic dominance. Because the definition of the first order stochastic dominance only depends on the CDFs of the random variables, in practice we need not to define the probability space or to know the joint disribtion of X and Y.

Here we present a method to approximate the set of feasible solutions of an optimization problem with a first order stochastic dominance constraint in the discrete distribution case. Our approximation method is safe, i.e., the solutions obtained by this method are necessarily feasible to the original optimization problem. This approximation helps us to obtain good feasible solutions for larger instances. The above formulation and the results derived in the paper can also be generalized to the optimization models with several first order stochastic dominance constraints.

The paper is organized as follows. In Section 2, we explain the optimization problem with a first order stochastic dominance constraint in the discrete distribution case. Our safe approximation for the dominance constraint is introduced in Section 3. In Section 4, we present the numerical results to study the efficiency and accuracy of the approximate method. Finally, we give concluding remarks in Section 5.

2 Optimization Problem with First Order Stochastic Dominance Constraint in Discrete Distribution Case

The stochastic optimization problem with a first order stochastic dominance constraint for discrete random variables can be formulated as a mixed integer linear program. Suppose that the entire space Ω has finitely many elementary events $\omega_1,...,\omega_T$ with probabilities of $p_t = \Pr(\{\omega_t\})$. Consider that Y is a discrete random variable in $\mathbf{X}$ with finite support, and y_i, $i = 1,...,m$, are its realizations with probability of q_i, $i = 1,...,m$. Without loss of generality assume that $y_1 < y_2 < ... < y_m$. Since F_Y is a right-continuous step function, the first order stochastic dominance constraint in (1) is equivalent to $\Pr(X < y_i) \le \Pr(Y < y_i) = F_Y(y_{i-1})$, $i = 1,...,m$, so by defining $y_0 = -\infty$, relation $X \succeq_{(1)} Y$ can be rewritten as

$$\Pr(X < y_i) \le F_Y(y_{i-1}),\ i = 1,...,m. \tag{2}$$

Thus, model (1) can be formulated as follows

$$\begin{aligned} \max\ & f(X) \\ \text{s.t.}\ & \Pr(X < y_i) \le F_Y(y_{i-1}),\ i = 1,...,m \\ & X \in \mathbf{C}. \end{aligned} \tag{3}$$

We have $F_Y(y_{i-1}) = \sum_{k=1}^{i-1} q_k$; moreover, we can write $\Pr(X < y_i) = \sum_{t=1}^{T} p_t z_{it}$ by introducing binary variables $z_{it} \in \{0,1\}$ such that for $i = 1,...,m$ and $t = 1,...,T$,

$$z_{it} = \begin{cases} 1 & \text{if } y_i - X(\omega_t) > 0 \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

Condition (4) can be expressed as linear mix integer constraints. To do this, it is required to define a big number $M \in \Re$ satisfying

$$M \ge \max_t \{y_m - X(\omega_t)\}. \tag{5}$$

Then, constraints $-M(1 - z_{it}) \le y_i - X(\omega_t) \le M z_{it}$, $i = 1,...,m$, $t = 1,...,T$ make z_{it} satisfy the condition (4). Therefore, model (3) can be reformulated as the following mixed integer program

$$\begin{aligned} \max\ & f(X) \\ \text{s.t.}\ & -M(1 - z_{it}) \le y_i - X(\omega_t) \le M z_{it},\ i = 1,...,m,\ t = 1,...,T \\ & \sum_{t=1}^{T} p_t z_{it} \le \sum_{k=1}^{i-1} q_k,\ i = 1,...,m \\ & z_{it} \in \{0,1\},\ i = 1,...,m,\ t = 1,...,T \\ & X \in \mathbf{C}. \end{aligned} \tag{6}$$

To simplify model (6) we can omit constraints $-M(1 - z_{it}) \le y_i - X(\omega_t)$, $i = 1,...,m$, $t = 1,...,T$. Hence, we get

$$\begin{aligned} \max\ & f(X) \\ \text{s.t.}\ & y_i - X(\omega_t) \le M z_{it},\ i=1,...,m,\ t=1,...,T \\ & \sum_{t=1}^{T} p_t z_{it} \le \sum_{k=1}^{i-1} q_k,\ i=1,...,m \\ & z_{it} \in \{0,1\},\ i=1,...,m,\ t=1,...,T \\ & X \in \mathbf{C}. \end{aligned} \tag{7}$$

Indeed, the constraints of model (7) do not imply (4), because $y_i - X(\omega_t) \le 0$ does not imply $z_{it} = 0$, though $y_i - X(\omega_t) > 0$ yields $z_{it} = 1$. However, we can show that the optimal solutions X^* of models (6) and (7) coincide. Consider that (X^*, z^*) is the optimal solution of model (7), where z^* is the vector of all $z_{it}^*, i=1,...,m,\ t=1,...,T$, and in z^* there are some elements z_{it}^* that do not satisfy (4), i.e. we have $y_i - X^*(\omega_t) \le 0$, but $z_{it}^* = 1$, which must be $z_{it}^* = 0$. If we change z^* to z^{**} such that all the elements of z^{**} satisfy (4), then $z^{**} \le z^*$, and so

$$\sum_{t=1}^{T} p_t z_{it}^{**} \le \sum_{t=1}^{T} p_t z_{it}^{*} \le \sum_{k=1}^{i-1} q_k,\ i=1,...,m .$$

This means that (X^*, z^{**}) is a feasible solution of model (7); furthermore, since model (7) is a relaxation of model (6), (X^*, z^{**}) is an optimal solution of model (6). Similar formulation to model (7) is also presented in [19] without mentioning the relation between models (6) and (7).

In practice we usually use the following formulation instead of the rather abstract formulation (7) to conveniently model and efficiently solve real world problems using the existing rich knowledge in numerical optimization,

$$\begin{aligned} \max\ & f(X_x) \\ \text{s.t.}\ & y_i - X_x(\omega_t) \le M z_{it},\ i=1,...,m,\ t=1,...,T \\ & \sum_{t=1}^{T} p_t z_{it} \le \sum_{k=1}^{i-1} q_k,\ i=1,...,m \\ & z_{it} \in \{0,1\},\ i=1,...,m,\ t=1,...,T \\ & x \in \mathbf{S} \end{aligned} \tag{8}$$

where $\mathbf{S} \subseteq \Re^N$ and X_x is a random variable which depends on x.

For example, in portfolio optimization $X_x \equiv R(x) = R_1 x_1 + R_2 x_2 + ... + R_N x_N$ where $R_1, R_2, ..., R_N$ are random return rates of assets $1,2,...,N$. The aim is to invest a certain capital in these assets in order to obtain some desirable characteristics of the total return rate on the investment. Denoting by $x_1, x_2, ..., x_N$ the fractions of the initial capital invested in assets $1,2,...,N$, we have $\mathbf{S} = \{x \in \Re^N : x_1 + x_2 + ... + x_N = 1, x_j \ge 0,\ j = 1,2,...,N\}$. If we consider that the return rates have a discrete joint distribution with realizations r_{jt}, $t=1,2,...,T$, $j=1,2,...,N$, attained with probabilities $p_t, t=1,2,...,T$, the objective function is the expected return rate, $f(X_x) \equiv \mathrm{E}[R(x)] = \sum_{t=1}^{T} p_t \sum_{j=1}^{N} x_j r_{jt}$, and Y is a discrete random variable with realizations y_i and probabilities q_i, $i=1,...,m$, the portfolio optimization problem becomes:

$$\begin{aligned} \max\ & \sum_{t=1}^{T} p_t \sum_{j=1}^{N} x_j r_{jt} \\ \text{s.t.}\ & y_i - \sum_{j=1}^{N} x_j r_{jt} \le M z_{it},\ i=1,...,m,\ t=1,...,T \\ & \sum_{t=1}^{T} p_t z_{it} \le \sum_{k=1}^{i-1} q_k,\ i=1,...,m \\ & \sum_{j=1}^{N} x_j = 1 \\ & x_j \ge 0,\ j=1,...,N \\ & z_{it} \in \{0,1\},\ i=1,...,m,\ t=1,...,T. \end{aligned} \tag{9}$$

3 Safe Approximation for Optimization Problem with First Order Stochastic Dominance Constraint

In the previous section, we see that an optimization problem with a first order stochastic dominance constraint can be formulated as a mixed integer program. This causes that such problems in large scales are difficult and usually impossible to be solved exactly. In the following, we present a method for providing approximate solutions for such integer programs.

In Model (8) the number of constraints is $m \times (T+1)$, and the number of binary variables is $T \times m$, while T represents the number of elementary events, and m is devoted to the number of the benchmark realizations. If the number of the benchmark realizations is reduced in a specified percentage, the number of the variables and constraints will be reduced accordingly. Denoting the reduction percentage by RP, the number of the constraints is reduced to $(1-RP) \times m \times (T+1)$, and the number of the variables is reduced to $(1-RP) \times T \times m$. According to this observation, in our proposed method we replace the benchmark random variable Y by another random variable Y_{RP} with smaller number of realizations. Hence, by solving the new problem, we can obtain an approximate solution more quickly.

In the following Y_{RP} will be called as RP -reduced benchmark, and we refer to the stochastic optimization problem (8), in which Y is replaced by Y_{RP} as the RP -reduced problem. Let m_{RP} be the number of the benchmark realizations after implementing the reduction. m_{RP} is calculated by $m_{RP} = m - \lfloor m \times RP \rfloor$, while $\lfloor a \rfloor$ refers to the integer part of $a \geq 0$. The new benchmark realizations after implementing the reduction are denoted by $y_{RP,\,i}$, $i = 1,...,m_{RP}$. Another symbol we apply in the following is $n(a)$ specifying the nearest integer number to $a \geq 0$, and if the fractional part of a is 0.5 then $n(a) = \lfloor a \rfloor + 1$.

Now we present two policies to obtain Y_{RP}, such that the feasible solutions of the RP -reduced problem remain feasible to the initial problem. We refer to such approximation as *safe approximation*.

The first policy. In the first policy the benchmark realizations are merged together regularly. The algorithm of this policy is given in the following.

For $i = 1$ *to* m_{RP}
  $y_{RP,\,i} = y_{n(i \times m / m_{RP})}$
  $F_{Y_{RP}}(y_{RP,\,i}) = F_Y\left(y_{n(i \times m / m_{RP})}\right)$
Next

For example, in the case of 50% reduction, the algorithm combines the first and second realizations into one realization with the CDF value of the second realization CDF value, third and forth realizations into one realization with the CDF value of the fourth realization CDF value, and so on.

The second policy. In the second policy we merge the benchmark realizations whose values are close to each other. The algorithm of this policy is as follows.

For $j = 1$ *to* $\lfloor m \times RP \rfloor$
  Calculate the differences between existing successive realizations: $d_i = y_{i+1} - y_i$ *for* $i = 1,...,m-j$
  Sort d_i, $i = 1,...,m-j$, *ascending, and set* i^* *as the index of the smallest* d_i
  For $i = i^*$ *to* $m - j$
    $y_i = y_{i+1}$
    $F_Y(y_i) = F_Y(y_{i+1})$
  Next
Next
For $i = 1$ *to* m_{RP}
  $y_{RP,\,i} = y_i$
  $F_{Y_{RP}}(y_{RP,\,i}) = F_Y(y_i)$
Next

In order to illustrate our safe approximate method, we present the following example of portfolio optimization.

Illustrative Example. Consider Model (9) for $N=3$, $T=m=20$. Decision variables, x_1, x_2 and x_3 refer to the fractions of the initial capital invested in assets 1, 2 and 3. Table 1 contains historical data for returns r_{jt} of three assets in 20 months, and we assume $p_t = 1/20$, $t = 1,2,...,20$. In addition, the return rates of the benchmark Y are available for each month in the last column.
We depict the CDF of the benchmark Y in Figure 1 and the CDFs of the 50%-reduced benchmark, obtained by the first and second policies, in Figure 2.

The CDF of the benchmark Y is depicted in Figure 1, and the CDFs of the 50%-reduced benchmark ($Y_{50\%}$) obtained by the first and second policies are shown in Figure 2. Using Model (9) we have solved this example in the exact case and also in the approximate cases with RP =50% and 70% by implementing the first and second policies.

Table 1:Asset returns for the illustrative example in Section 3

Realization (Month)	Asset 1	Asset 2	Asset 3	Benchmark returns
1	0.649	0.579	0.201	0.346
2	0.221	0.102	0.383	0.105
3	0.704	0.294	0.266	0.291
4	0.095	0.877	0.761	0.448
5	0.086	0.344	0.532	0.191
6	0.561	0.23	0.394	0.265
7	0.979	0.897	0.476	0.654
8	0.635	0.356	0.662	0.421
9	0.622	0.286	0.494	0.337
10	0.735	0.518	0.763	0.542
11	0.367	0.947	0.848	0.591
12	0.998	0.721	0.36	0.563
13	0.21	0.903	0.209	0.311
14	0.191	0.783	0.441	0.342
15	0.581	0.052	0.591	0.278
16	0.851	0.953	0.041	0.485
17	0.409	0.812	0.681	0.504
18	0.572	0.488	0.73	0.467
19	0.761	0.887	0.869	0.709
20	0.595	0.844	0.03	0.36

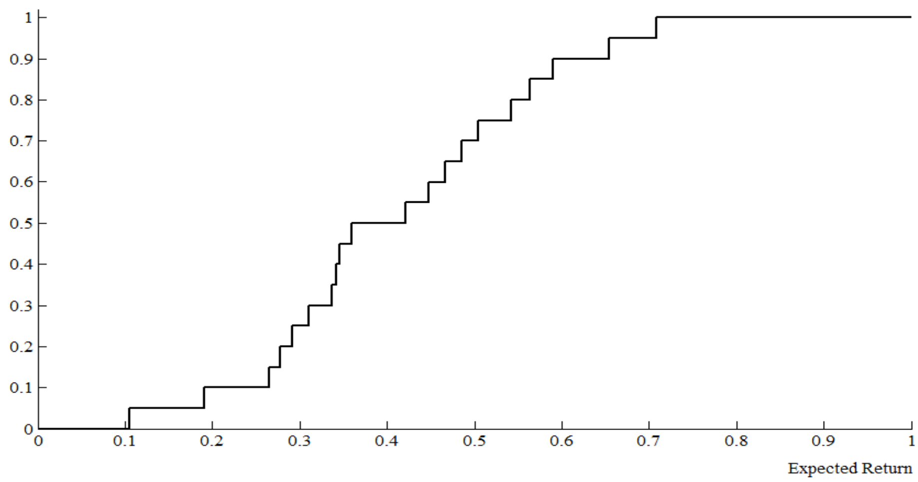

Figure 1: The CDF of the benchmark Y studied in the illustrative example in Section 3

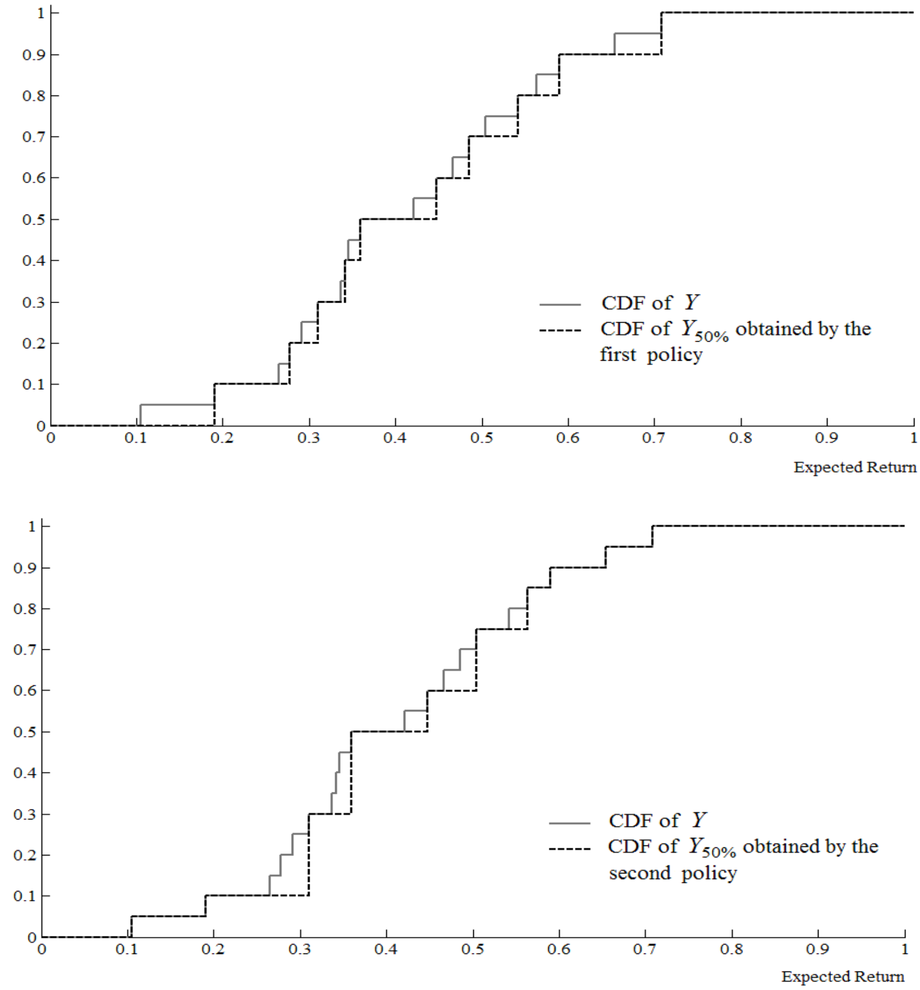

Figure 2: The CDF of the 50%-reduced benchmark ($Y_{50\%}$) obtained by the first policy and the second policy in comparison with the CDF of the original benchmark (Y) in the illustrative example in Section 3

Figure 3 shows the feasible regions for x_1 and x_2 in this example (note $x_3 = 1 - x_1 - x_2$) for the exact case in comparison with the feasible regions of 50% and 70%-reduced problems applying the first and second policies. As shown in this figure, the feasible region of the exact problem contains solutions of the approximate problems. The optimal solutions and values for the various cases of this example are presented in Table 2. We see that the second policy presents better objective values.

Table 2: Optimal solutions and values for the illustrative example in Section 3

	Exact method	First policy		Second policy	
		RP		*RP*	
		50%	70%	50%	70%
Optimum objective	58.0%	56.0%	53.2%	57.3%	56.3%
x_1^*	0.26	0	0	0.2	0.58
x_2^*	0.74	0.68	0.42	0.71	0.42
x_3^*	0	0.32	0.58	0.09	0

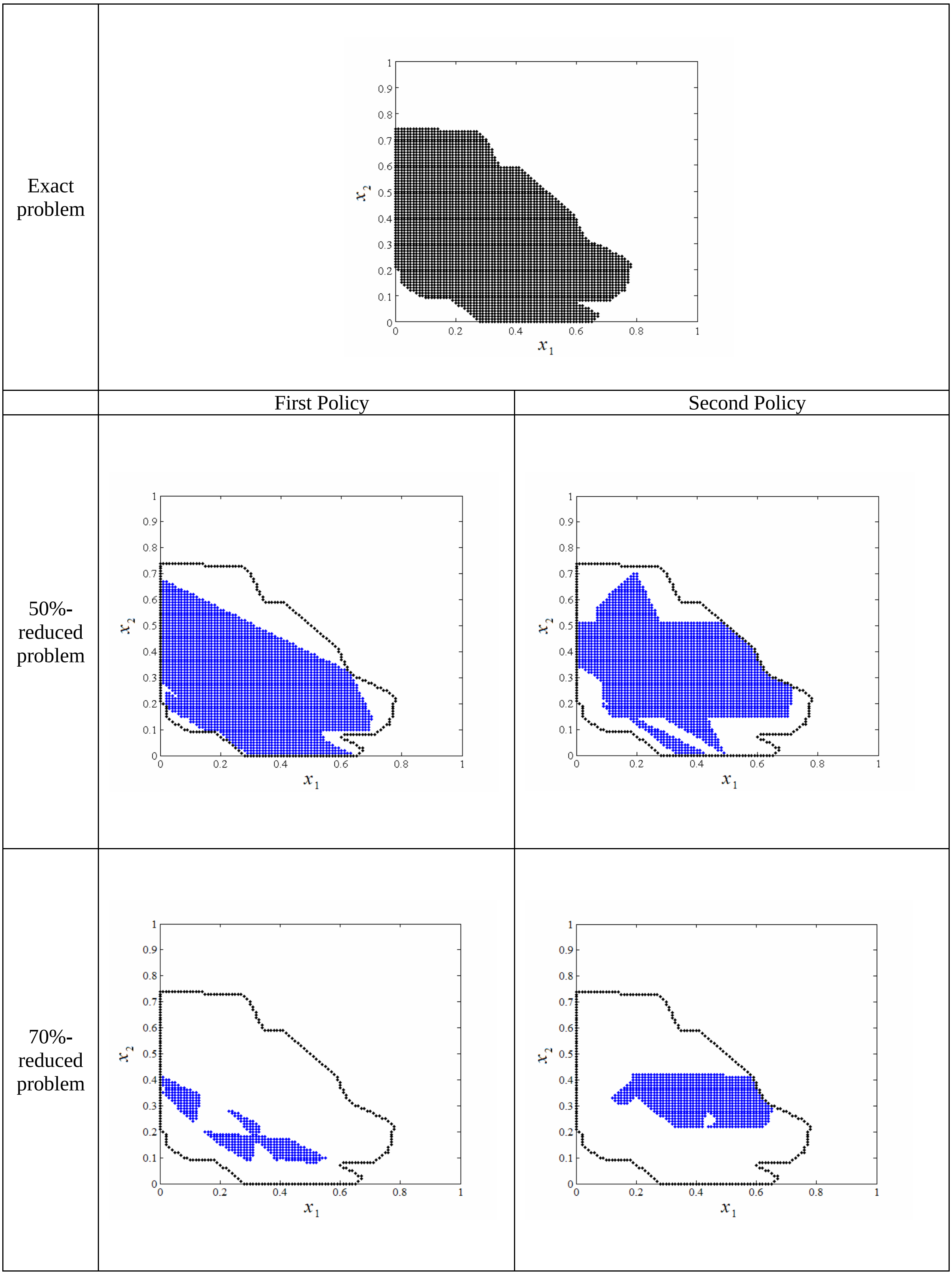

Figure 3: The feasible regions of the exact problem and the problems with *RP* of 50% and 70% by applying the first and second policies for the illustrative example in Section 3.
(The dashed black lines in the last four figures specify the boundary of the exact feasible region.)

4 Numerical Results

To assess the computational efficiency of the presented safe approximation method, we consider a portfolio optimization problem with finitely many assets formulated in (9). We define benchmark return Y as the average of the return rates of the assets, i.e., each realization y_i is defined as:

$$y_i = \frac{1}{N}\sum_{j=1}^{N} r_{ji},\ i = 1,...,m\ .$$

To compare the proposed policies and their properties, we generate instances of randomly produced data with 200 assets and 50 realizations. For each scenario in Table 3, we have solved 25 instances using Lingo 8.0 and applying the cutting plane method introduced in [1].

Table 3: The averages of reduction percentages in CPU time and objective value obtained by applying the first and second policies

Scenario	Reduction policy	RP (%)	Average of reduction percentages in CPU time	Average of reduction percentages in objective value
1	First policy	50	77.5443%	0.0503%
2	Second policy	50	53.5477%	0.0053%
3	First policy	70	94.3343%	0.1147%
4	Second policy	70	82.7023%	0.0350%
5	First policy	90	97.7906%	0.3711%
6	Second policy	90	95.8712%	0.2487%

Table 3 shows the averages of reduction percentages in CPU time and in objective value for the instances with RP of 50%, 70% and 90% by exploiting the two policies in comparison with the exact results. As it could be seen from this table, applying the first policy saves time more than the second one, but the objective values obtained by the first policy are worse than those obtained by the second policy. Also, it is interesting that for relatively high values of reduction percentage, e.g. see Scenario 4, the reduction percentages in objective value are small in comparison with the reduction percentages, particularly when we apply the second policy.

5 Conclusions

Optimization problems with first order stochastic dominance constraints in the case of discrete distribution can be formulated as mixed integer programs. In such problems the desirable outcome must dominate an available random variable called benchmark in the sense of the first order stochastic dominance. In this paper, we presented a method for safely approximating such mixed integer programs by reducing the number of benchmark realizations. We used two policies to reduce the number of the benchmark realizations such that the feasible solutions of the new problems, attained after implementing these reductions, remain feasible to the initial problem. The numerical results showed that in reasonable time our method could lead to good solutions.

References

[1] Ahmadi Javid, A., and N. Kazemzadeh, A cutting plane method for optimization with first order stochastic dominance constraints, submitted for publication, 2008.

[2] Damodaran, U., Stochastic dominance and analysis of ODI batting performance: The Indian Cricket Team, 1989-2005. *Journal of Sports Science and Medicine*, vol.5, pp.503-508, 2006.

[3] Dentcheva, D., and A. Ruszczyński, Optimization with stochastic dominance constraints, *SIAM Journal on Optimization*, vol.14, pp.548-566, 2003.

[4] Dentcheva, D., and A. Ruszczyński, Optimality and duality theory for stochastic optimization problems with nonlinear dominance constraints, *Mathematical Programming*, vol.99, pp.329-350, 2004.

[5] Dentcheva, D., and A. Ruszczyński, Convexification of stochastic ordering constraints, *Comptes Rendus de l_Academie Bulgare des Sciences*, vol.57, no.4, pp.11-16, 2004.

[6] Dentcheva, D., and A. Ruszczyński, Semi-infinite probabilistic optimization: first order stochastic dominance constraints, *Optimization*, vol.53, pp.583-601, 2004.

[7] Dentcheva, D., and A. Ruszczyński, Portfolio optimization with stochastic dominance constraints, *Journal of Banking & Finance*, vol.30, pp.433-451, 2006.

[8] Gollmer, R., U. Gotzes, *et al.*, Risk modeling via stochastic dominance in power systems with dispersed generation, Tech. Report Preprint 651/2007, Department of Mathematics, University of Duisburg-Essen, www.uniduisburg.de/FB11/disma/ preprints.shtml, 2007.

[9] Hadar, J., and W. Russell, Rules for ordering uncertain prospects, *American Economic Review*, vol.59, pp.25-34, 1969.

[10] Hanoch, G., and H. Levy, The efficiency analysis of choices involving risk, *Review of Economic Study*, vol.36, pp.335-346, 1969.

[11] Hardy, G.H., J.E. Littlewood, and G. Pólya, *Inequalities*, Cambridge University Press, Cambridge, MA, 1934.

[12] Lee, E.K., T. Fox, and I. Crocker, Integer programming applied to intensity-modulated radiation therapy treatment planning, *Annals of Operations Research*, vol.119, pp.165-181, 2003

[13] Lehmann, E., Ordered families of distributions, *The Annals of Mathematical Statistics*, vol.26, pp.399-419, 1955.

[14] Levy, H., Stochastic dominance and expected utility: Survey and analysis, *Management Science*, vol.38, pp.555-593, 1992.

[15] Luedtke, J., New formulations for optimization under stochastic dominance constraints, to appear in *SIAM Journal on Optimization*, 2008.

[16] Mann, H.B., and D.R. Whitney, On a test of whether one of two random variables is stochastically larger than the other, *The Annals of Mathematical Statistics*, vol.18, pp.50-60, 1947.

[17] Marshall, A.W., and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*, Academic Press, San Diego, 1979.

[18] Mosler, K., and M. Scarsini, *Stochastic Orders and Decision under Risk*, Institute of Mathematical Statistics, Hayward, California, 1991.

[19] Noyan, N., and A. Ruszczyński, Valid inequalities and restrictions for stochastic programming problems with first order stochastic dominance constraints, *Mathematical Programming*, vol.115, pp.249-275, 2008.

[20] Quirk, J.P., and R. Saposnik, Admissibility and measurable utility functions, *Review of Economic Studies*, vol.29, pp.140-146, 1962.

[21] Romeijn, H.E., R.K. Ahuja, *et al*. Dempsey and A. Kumar, A new linear programming approach to radiation therapy treatment planning problems, *Operation Research*, vol.54, pp.201-216, 2006.

[22] Rothschild, M., and J.E. Stiglitz, Increasing risk: I. A definition, *Journal of Economic Theory*, vol.2, pp.225-243, 1969.