Convex relaxations of chance constrained optimization problems

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

May 12, 2011

Abstract

In this paper we develop convex relaxations of chance constrained optimization problems in order to obtain lower bounds on the optimal value. Unlike existing statistical lower bounding techniques, our approach is designed to provide deterministic lower bounds. We show that a version of the proposed scheme leads to a tractable convex relaxation when the chance constraint function is affine with respect to the underlying random vector and the random vector has independent components. We also propose an iterative improvement scheme for refining the bounds.

1 Introduction

We consider chance constrained optimization problems of the form

$$\min\{f(x) : \ \Pr[g(x,\xi) \geq 0] \geq 1-\epsilon, \ x \in S\}, \tag{1}$$

where $x \in \mathbb{R}^n$ is a vector of decision variables, ξ is a random vector with (given) probability distribution P and support $\Xi \subseteq \mathbb{R}^d$, the function $g : \mathbb{R}^n \times \Xi \mapsto \mathbb{R}$ defines a random constraint on x, $\Pr[A]$ denotes the probability of the event A, the scalar $\epsilon \in (0,1)$ is a pre-specified allowed probability of violation of the random constraint defined by g, $S \subseteq \mathbb{R}^n$ is a nonempty set defined by some deterministic side constraints on x, and $f : \mathbb{R}^n \mapsto \mathbb{R}$ is the objective function. We assume that the function f and the set S are convex, and the function g is concave in x for all $\xi \in \Xi$. Of particular interest are chance constrained linear programs where the function f is linear, the set S is defined by a finite number of linear inequalities, and the function g is affine in x. The reader is referred to [8, 9] for extensive reviews on the theory and applications of chance constrained optimization problems.

The two main difficulties in solving chance constrained optimization problems are: (i) checking feasibility of a given solution is, in general, impossible, and (ii) since the feasible region is typically non-convex, optimization poses severe challenges. As such, an important line of research has been to develop an inner approximation (restriction) of the chance constraint whose solutions are guaranteed (in some well defined sense) to satisfy the chance constraint. An additional requirement is that the approximation should be relatively easily solvable. One class of developments along this direction focuses on constructing a deterministic convex programming restriction to the chance constraint that can be optimized efficiently using standard methods [5, 7]. Another class of approaches solve deterministic optimization problems, built from samples of the random vector, that are guaranteed to produce feasible solutions with high probability [2, 4]. Various meta-heuristic methods have also been suggested for obtaining feasible solutions to chance constrained problems [1, 3]. All of these approaches are mainly designed to produce feasible solutions without any optimality guarantees. A good lower bound on the true optimal value is needed to provide an optimality certificate for the produced feasible solutions. There has been relatively less work in lower bounding chance constrained optimization problems. A commonly used approach is to use order statistics of optimal values of sampled approximations of the chance constrained problem to obtain a statistical lower bound estimate [4, 5, 6]. Such approaches require the solution of a large number of (sampled) optimization problems in order to produce a statistically valid lower bound with reasonable confidence. In this paper we propose a *deterministic* lower bounding technique by constructing a convex *relaxation* of the chance constrained problem. The approach is similar to the construction used in [5, 7] to develop a convex *restriction* of the problem. Following [5, 7] we show that a version of the proposed scheme leads to a tractable convex relaxation when the random vector ξ has independent components and the constraint function g is affine in ξ.

Remark: The chance constrained optimization problem (1) has a single individual chance constraint. The developments in this paper extend immediately to multiple individual chance constraints of the form $\Pr[g_i(x,\xi) \geq 0] \geq 1-\epsilon_i$ for $i = 1,\dots,m$. The case of *joint* chance constraints, i.e. constraints of the form $\Pr[g_i(x,\xi) \geq 0 \ \forall\ i = 1,\dots,m] \geq 1-\epsilon$, is more involved. Of course such a constraint can be represented as a single individual chance constraint $\Pr[h(x,\xi) := \min_i\{g_i(x,\xi)\} \geq 0] \geq 1-\epsilon$, however such a transformation may not preserve desirable properties such as affine-ness of the constraint function. Since we are interested in relaxations, we can relax the joint chance constraint into m individual chance constraints by noting that

$$\Pr[g_i(x,\xi) \geq 0 \ \forall\ i = 1,\dots,m] \leq \min_{i=1,\dots,m} \Pr[g_i(x,\xi) \geq 0].$$

We can then apply the developments of this paper to each of the individual m chance constraints separately.

2 The Relaxation Scheme

Let us denote the set of solutions to the chance constrained problem (1) as

$$X = \{x \in S \subseteq \mathbb{R}^n :\ \Pr[g(x,\xi) \geq 0] \geq 1-\epsilon\}.$$

In this section we propose a scheme to construct a convex relaxation of X. We make the following key assumption throughout:

(A0) For each $\xi \in \Xi$, there exists $L(\xi) > 0$ such that $g(x,\xi) \geq -L(\xi)$ for all $x \in S$.

We can now express X equivalently as

$$X = \{x \in S :\ \Pr[h(x,\xi) \geq 1] \geq 1-\epsilon\}, \tag{2}$$

where

$$h(x,\xi) := g(x,\xi)/L(\xi) + 1.$$

Note that by (A0), h is nonnegative for all $\xi \in \Xi$ and for all x of interest. We also make the necessary technical assumptions for the random function h to be measurable. We will work with the description (2) of X for the remainder of this paper.

Let Φ be a class of univariate functions $\phi : \mathbb{R} \mapsto \mathbb{R}$ such that

(a) $\phi(t) \geq 0$ for all $t \geq 0$,

(b) $\phi(t)$ is nondecreasing everywhere and strictly increasing on $[0,1]$, and

(c) $\phi(t)$ is concave.

Proposition 1 *For any $\phi \in \Phi$, the set*

$$X_\phi := \{x \in S :\ \mathbb{E}[\phi(h(x,\xi))] \geq (1-\epsilon)\phi(1)\} \tag{3}$$

is a convex relaxation of X.

Proof: By property (b) and (c) of ϕ and the concavity of g, and hence h, the composite function $\phi(h(x,\xi))$ is concave in x for all $\xi \in \Xi$. Thus $\mathbb{E}[\phi(h(x,\xi))]$ is concave in x and so X_ϕ is a convex set. Consider any feasible solution $x \in X$. By property (b), for any $\xi \in \Xi$, $h(x,\xi) \geq 1$ if and only if $\phi(h(x,\xi)) \geq \phi(1)$. Thus $\Pr[h(x,\xi) \geq 1] = \Pr[\phi(h(x,\xi)) \geq \phi(1)] \geq 1-\epsilon$. Note that (a) and (b) imply $\phi(1) > 0$. Moreover, $\phi(h(x,\xi)) \geq 0$ for all ξ. Thus we can apply Markov's inequality on the nonnegative random variable $\phi(h(x,\xi))$ leading to $\mathbb{E}[\phi(h(x,\xi))]/\phi(1) \geq \Pr[\phi(h(x,\xi)) \geq \phi(1)] \geq 1-\epsilon$. Thus $X_\phi \supseteq X$. $\square$

Let us now address the choice of the function ϕ. Clearly the identity $\phi(t) = t$ works. In this case $X_\phi = \{x \in S :\ \mathbb{E}[h(x,\xi)] \geq (1-\epsilon)\}$, however such a relaxation may be quite weak. For any $\phi_1, \phi_2 \in \Phi$ with $\phi_1(1) = \phi_2(1)$, if $\phi_1(t) \leq \phi_2(t)$ for all $t \geq 0$ then $X_{\phi_1} \subseteq X_{\phi_2}$. So we would like to use a function in Φ that is, when normalized with respect to its value at 1, smallest over the non-negatives.

Proposition 2 *Given any $\phi \in \Phi$, let $\phi^\circ(t) := [\phi(1) - \phi(0)]\min\{1,t\} + \phi(0)$. Then $\phi^\circ \in \Phi$, $\phi^\circ(1) = \phi(1)$ and $\phi^\circ(t) \le \phi(t)\ \forall\ t \ge 0$.*

Proof: Note that by (a) and (b), $0 \le \phi(0) < \phi(1)$. So clearly $\phi^\circ \in \Phi$. If $t > 1$ then $\phi^\circ(t) = \phi(1)$ and $\phi(t) \ge \phi(1)$, and so the claim holds. On the other hand, if $t \in [0,1]$, by concavity (c), $\phi(t) \ge t\phi(1) + (1-t)\phi(0) = [\phi(1) - \phi(0)]t + \phi(0) = \phi^\circ(t)$. □

Thus, up to translation and scaling, the function

$$\phi^*(t) = \min\{1, t\} \tag{4}$$

is the smallest function over the non-negatives in Φ, and so $X_{\phi^*} \subseteq X_\phi$ for all $\phi \in \Phi$ with $\phi(1) = 1$.

Consider now the discrete distribution setting where the random vector ξ takes finitely many values $\xi^1, \ldots, \xi^k$ with $\Pr[\xi = \xi^i] = p_i$ for all $i = 1, \ldots, k$. Then the relaxation X_{ϕ^*} corresponding to $\phi^*(t)$ is given by vectors $x \in S$ satisfying:

$$\begin{array}{ll} \sum_{i=1}^k p_i y_i \ge (1-\epsilon) & \\ 0 \le y_i \le 1 & \forall\ i = 1, \ldots, k \\ L(\xi^i) y_i \le g(x, \xi^i) + L(\xi^i) & \forall\ i = 1, \ldots, k, \end{array} \tag{5}$$

where we have introduced the nonnegative variable y_i to model $\min\{1, g(x,\xi^i)/L(\xi^i) + 1\}$ for all $i = 1, \ldots, k$. Note that (5) is simply the continuous relaxation of the following standard mixed integer programming formulation of the chance constraint (2) in the discrete distribution setting:

$$\begin{array}{ll} \sum_{i=1}^k p_i y_i \ge (1-\epsilon) & \\ y_i \in \{0, 1\} & \forall\ i = 1, \ldots, k \\ L(\xi^i) y_i \le g(x, \xi^i) + L(\xi^i) & \forall\ i = 1, \ldots, k. \end{array} \tag{6}$$

The construction above shows that, in the discrete distribution setting with a large number of mass points k, the relaxation X_{ϕ^*} is a large-scale optimization problem which may be time consuming to process. Even in the simple setting when ξ has independent components and g is affine in ξ, due to the non-smooth nature of ϕ^*, the relaxation X_{ϕ^*} may not be easy to compute. In the following section we propose an alternative function from the family Φ that leads to more tractable convex programming formulations in some settings.

3 Bernstein Relaxation

In this section we consider relaxations of the form (3) developed in Proposition 1 corresponding to the function

$$\phi^\beta(t) = 1 - \exp(-t). \tag{7}$$

Following [5] and [7], where the authors used a similar exponential function to construct efficiently computable convex *restrictions* to chance constraints, we refer to the convex relaxations corresponding to ϕ^β as "Bernstein" relaxations in appreciation of the connections to the work of S. N. Bernstein in classical large deviation theory.

We first note that it is easily verified that ϕ^β satisfies conditions (a)-(c) and so $\phi^\beta \in \Phi$. Thus X_{ϕ^β} is a valid convex relaxation, which, after simple algebraic manipulations, is given by

$$X_{\phi^\beta} = \{x \in S :\ \mathbb{E}[\exp(-g(x,\xi)/L(\xi))] \le \eta(\epsilon)\}, \tag{8}$$

where $\eta(\epsilon) := (1-\epsilon) + \exp(1)\epsilon$.

In the discrete distribution setting where the random vector ξ takes finitely many values $\xi^1, \ldots, \xi^k$ with $\Pr[\xi = \xi^i] = p_i$ for all $i = 1, \ldots, k$, the relaxation X_{ϕ^β} is given by the single constraint

$$\psi(x) := \sum_{i=1}^k p_k \exp(-g(x,\xi^i)/L(\xi^i)) \le \eta(\epsilon). \tag{9}$$

Note that ψ is finite-valued and convex. If $g(x,\xi^i)$ is differentiable in x for all i, then ψ is also differentiable. Optimizing over the single convex constraint (9) can be significantly faster than solving the non-smooth relaxation given by (5), particularly when k is large. This is illustrated next.

Numerical Illustration

We compare the bounds and computing times corresponding to X_{ϕ^*} and X_{ϕ^β} for the following chance constrained linear program:

$$\min\{\sum_{j=1}^{n} x_j : \Pr[\sum_{j=1}^{n} \xi_j x_j \geq 1] \geq 1-\epsilon,\ x_j \geq \ell_j\ j=1,\ldots,n\}, \tag{10}$$

where ξ is a non-negative random vector with a discrete distribution with k realizations $\xi^1,\ldots,\xi^k$ and $\Pr[\xi=\xi^i]=p_i$ for all $i=1,\ldots,k$, and $\ell_j>0$ for all $j=1,\ldots,n$. Note that, here $g(x,\xi)=\sum_{j=1}^n \xi_j x_j - 1$ and $L(\xi^i)=1$ for all $i=1,\ldots,k$ correspond to valid lower bounds since ξ and x are nonnegative. Thus relaxation X_{ϕ^*} of (10) constructed according to (5) is the linear program,

$$\begin{array}{lll} \min & \sum_{j=1}^n x_j & \\ \text{s.t.} & \sum_{i=1}^k p_i y_i \geq (1-\epsilon) & \\ & y_i \leq \sum_{j=1}^n \xi_j^i x_j & \forall\ i=1,\ldots,k, \\ & 0 \leq y_i \leq 1 & \forall\ i=1,\ldots,k \\ & x_j \geq \ell_j & \forall\ j=1,\ldots,n, \end{array} \tag{11}$$

and the Bernstien relaxation X_{ϕ^β} given by (8) is the convex nonlinear program,

$$\begin{array}{ll} \min & \sum_{j=1}^n x_j \\ \text{s.t.} & \sum_{i=1}^k p_i \exp(-\sum_j \xi_j^i x_j + 1) \leq \eta(\epsilon) \\ & x_j \geq \ell_j \quad \forall\ j=1,\ldots,n. \end{array} \tag{12}$$

We considered instances corresponding to various values of n and k. We generated ξ_j^i for $j=1,\ldots,n$ and $i=1,\ldots,k$ from U(0.3,1.7), and set $p_i=1/k$ for all $i=1,\ldots,k$, $\epsilon=0.05$, and $\ell_j=10^{-4}$ for all $j=1,\ldots,n$. The computations were carried out using GAMS with CONOPT 3 as the nonlinear programming solver and CPLEX 12.2 as the linear programming solver. Table 1 presents the results. As expected, the bounds from X_{ϕ^β} are weaker but are typically cheaper to compute than those from X_{ϕ^*}, especially when the cardinality of the distribution, i.e. k, is high.

		X_{ϕ^*}		X_{ϕ^β}	
n	k	Bound	Time	Bound	Time
100	1000	0.943	1.38	0.909	2.12
100	2000	0.944	5.13	0.911	1.87
100	5000	0.947	47.41	0.914	4.79
200	1000	0.939	2.21	0.907	2.39
200	2000	0.943	12.02	0.910	3.78
200	5000	0.947	123.46	0.916	9.42
500	1000	0.940	3.99	0.908	5.18
500	2000	0.943	34.43	0.909	8.26
500	5000	0.946	299.31	0.913	29.28

Table 1: Comparison of bounds and time for discrete distribution

4 An Efficiently Computable Setting

We now describe a setting where the Bernstein relaxation (8) leads to an efficiently computable convex program even when the distribution of ξ is continuous. The setting and the subsequent construction are identical to those in [5] and [7]. We make the following assumptions.

(A1) The parameter $L(\xi)$ assumed in (A0) is deterministic, i.e., $L(\xi)=L>0$ for all $\xi \in \Xi$.

(A2) The random vector ξ has independently distributed components $\xi_1, \ldots, \xi_d$. We denote the support of ξ_j by Ξ_j, so that $\Xi = \Xi_1 \times \cdots \times \Xi_d$.

(A3) For $j = 1, \ldots, d$, the moment generating function of ξ_j, denoted by $M_j(t) = \mathbb{E}[\exp(t\xi_j)]$, is finite valued for all $t \in \mathbb{R}$ and is efficiently computable. The logarithmic moment generating function of ξ_j is denoted by $\Lambda_j(t) := \log M_j(t)$.

(A4) The constraint function $g(x, \xi)$ is affine in ξ, i.e.,

$$g(x, \xi) = g_0(x) + \sum_{j=1}^{d} \xi_j g_j(x),$$

where $g_j(x)$ is real valued and concave for all $j = 0, 1, \ldots, d$, and for each j such that $\Xi_j \not\subset \mathbb{R}_+$ the function $g_j(x)$ is affine.

Proposition 3 *Under assumptions (A1)-(A4), X_{ϕ^β} is a convex set defined by an efficiently computable constraint of the form*

$$X_{\phi^\beta} = \{x \in S : \ -g_0(x)/L + \sum_{j=1}^{d} \Lambda_j(-g_j(x)/L) \leq \log \eta(\epsilon)\}. \tag{13}$$

Proof: Under assumptions (A1)-(A4),

$$\begin{aligned} \mathbb{E}[\exp(-g(x,\xi)/L(\xi))] &= \mathbb{E}[\exp(-g_0(x)/L) \times \textstyle\prod_{j=1}^d \exp(-\xi_j g_j(x)/L)] \\ &= \exp(-g_0(x)/L) \times \textstyle\prod_{j=1}^d M_j(-g_j(x)/L). \end{aligned}$$

Using the above expression for the right-hand-side of the inequality describing (8) and taking logs on both sides we arrive at the representation (13). Since $-g_j(\cdot)$ and $\Lambda_j(\cdot)$ are convex for all $j = 0, 1, \ldots, d$ and Λ_j is monotone nondecreasing for all j with $\Xi_j \subset \mathbb{R}_+$, the left hand side of the inequality defining (13) is a convex function of x, and hence X_{ϕ^β} is a convex set. Moreover, the derived constraint function is efficiently computable due to (A3). □

Numerical Illustration

We revisit the chance constrained linear program (10) where the constraint coefficient ξ_j for $j = 1, \ldots, n$ are now assumed to be independent and uniformly distributed with nonnegative support. Let $[a_j, b_j]$, with $0 \leq a_j \leq b_j$, be the support of ξ_j, then the moment generating function of ξ_j is

$$M_j(t) = \begin{cases} \frac{\exp(tb_j) - \exp(ta_j)}{t(b_j - a_j)} & \text{for } t \neq 0 \\ 1 & \text{for } t = 0. \end{cases}$$

Taking $g(x, \xi) = \sum_{j=1}^n \xi_j x_j - 1$ with $g_0(x) = -1$ and $g_j(x) = x_j$ for $j = 1, \ldots, n$; and $L = 1$, we note that problem (10) now satisfies assumptions (A1)-(A4). Using Proposition 3, the Bernstein relaxation of the chance constrained linear program (10) is the convex nonlinear program:

$$\begin{array}{ll} \min & \sum_{j=1}^n x_j \\ \text{s.t.} & \sum_{j=1}^n \log\left(\frac{\exp(-x_j a_j) - \exp(-x_j b_j)}{x_j(b_j - a_j)}\right) \leq \log \eta(\epsilon) - 1 \\ & x_j \geq \ell_j \quad \forall\, j = 1, \ldots, n. \end{array} \tag{14}$$

We shall compare the deterministic bounds obtained from the above Bernstein relaxation to the following statistical bounding technique prescribed in [5]. Choose positive integers L, M, N and let $\{\xi^{1i}, \xi^{2i}, \ldots, \xi^{Ni}\}$ for $i = 1, \ldots, M$, be M independent samples, each of size N, of the underlying random vector ξ. For each $i = 1, \ldots, M$, solve the linear program:

$$\begin{array}{lll} \min & \sum_{j=1}^n x_j & \\ \text{s.t.} & \sum_{j=1}^n \xi_j^{ik} x_j - 1 \geq 0 & \forall\, k = 1, \ldots, N \\ & x_j \geq \ell_j & \forall\, j = 1, \ldots, n, \end{array} \tag{15}$$

and let v^i be the corresponding optimal value. Let $v^{(1)} \leq v^{(2)} \leq \cdots \leq v^{(M)}$ be the nondecreasingly sorted values of the $\{v^i\}_{i=1}^M$. Then, if

$$\sum_{r=0}^{L-1} \binom{M}{r} (1-\epsilon)^{Nr}[1-(1-\epsilon)^N]^{M-r} \leq \alpha \tag{16}$$

for some $\alpha \in (0,1)$, then, with probability at least $1-\alpha$, the quantity $v^{(L)}$ is a valid lower bound on the optimal value of (10). Note that the approach requires solving M linear programs (15) each with N constraints. For a given L and α, higher value of N means higher values of M. On the other hand, if N is small, the obtained bounds are typically weak.

Table 2 compares the bounds and times for the Bernstein relaxation (14) and the statistical bounding procedure of [5] described above for the chance constrained linear program (10) where $\xi_j \sim U(0.3, 1.7)$ independently for all $j = 1, \ldots, n$. That is, $a_j = 0.3$ and $b_j = 1.7$ for all $j = 1, \ldots, n$ in (14). Recall that $\ell_j = 10^{-4}$ for all $j = 1, \ldots, n$ and $\epsilon = 0.05$. We considered problems with $n = 100, 200$ and 500 variables. The Bernstein bounds are nearly identical across the problem sizes due to the identical distribution of the constraint coefficients. For the statistical bounding procedure, we assumed $\alpha = 0.001$ and $L = 1$ and then computed the required M for different values of N according to (16). Thus the statistical lower bounds in columns 3, 5 and 7 (except those in the last row) of Table 2 correspond to a confidence level of at least 0.999. We observe that the quality of the statistical bounds is poor in comparison to the deterministic Bernstein bound unless N is large and the problem size is small. We also observe that the statistical bounding technique is significantly more expensive than the Bernstein bounding procedure, particularly for large problems. Arguably, we did not make any attempt to warm-start the M linear programs (15) which perhaps could cut down the computing time. However, for large N, even a single solve of the linear program (15) is more expensive than that of the single-constrained convex nonlinear program (14).

		$n = 100$		$n = 200$		$n = 500$	
N	M	Bound	Time	Bound	Time	Bound	Time
10	8	0.815	2.51	0.781	2.94	0.758	3.09
20	16	0.870	5.66	0.858	6.20	0.832	6.97
30	29	0.910	10.66	0.887	12.28	0.871	18.08
40	51	0.916	21.73	0.905	26.10	0.885	39.63
50	87	0.944	46.54	0.924	61.82	0.896	109.33
X_{ϕ^β}		0.918	0.36	0.918	0.42	0.918	0.45

Table 2: Comparison of bounds and time for continuous distribution

5 An Improvement Procedure

The relaxation scheme outlined previously depends on the quality of the lower bound on $g(x,\xi)$ assumed in Assumptions (A0) and (A1). In this section we discuss how we can improve these bounds, and hence the quality of the relaxation. We first consider the deterministic lower bound case, i.e. Assumption (A1).

Proposition 4 *Suppose for all $x \in S$ and $\xi \in \Xi$, we have that $g(x,\xi) \geq -L_2 \geq -L_1$ for some $0 < L_2 \leq L_1$. For $t = 1, 2$, let $h_i(x,\xi) := g(x,\xi)/L_i + 1$ and $X_\phi^i := \{x \in \mathbb{R}^n : \mathbb{E}[\phi(h_i(x,\xi))] \geq (1-\epsilon)\phi(1)\}$ for some $\phi \in \Phi$. Then*

$$X \subseteq X_\phi^2 \subseteq X_\phi^1.$$

Proof: Since L_1 and L_2 correspond to valid lower bounds, it follows from Proposition 1 that $X \subseteq X_\phi^i$ for $i = 1, 2$. Let $\alpha := L_2/L_1$ and note that $\alpha \in (0,1]$. Then $h_1(x,\xi) = \alpha h_2(x,\xi) + (1-\alpha)$ for any x and ξ. By concavity of ϕ we have

$$\phi(h_1(x,\xi)) \geq \alpha\phi(h_2(x,\xi)) + (1-\alpha)\phi(1)$$

for any x and ξ. Taking expectations on both sides of the inequality, we have for any x,

$$\mathbb{E}[\phi(h_1(x,\xi))] \geq \alpha\mathbb{E}[\phi(h_2(x,\xi))] + (1-\alpha)\phi(1).$$

Consider $\widehat{x} \in X_\phi^2$. Then $\mathbb{E}[\phi(h_2(\widehat{x},\xi))] \geq (1-\epsilon)\phi(1)$. Thus

$$\mathbb{E}[\phi(h_1(\widehat{x},\xi))] \geq \alpha(1-\epsilon)\phi(1) + (1-\alpha)\phi(1) = (1-\alpha\epsilon)\phi(1) \geq (1-\epsilon)\phi(1),$$

where the last inequality follows from the fact that $\alpha \in (0,1]$. Thus $\widehat{x} \in X_\phi^1$. □

Thus we get tighter valid relaxations with better lower bounds on $g(x,\xi)$. The above result suggests the following iterative improvement scheme.

Corollary 1 *Assume (A1). Let $L_0 = L > 0$, and set*

$$L_{i+1} = -\min_{\xi\in\Xi}\min_{x\in X_\phi^i}\{g(x,\xi)\} \text{ for } i = 0,1,2,\ldots, \tag{17}$$

as long as $L_{i+1} > 0$, where $X_\phi^i := \{x \in S: \ \mathbb{E}[\phi(h_i(x,\xi))] \geq (1-\epsilon)\phi(1)\}$ and $h_i(x,\xi) := g(x,\xi)/L_i + 1$. Then

$$X \subseteq \ldots X_\phi^{i+1} \subseteq X_\phi^i \subseteq \ldots X_\phi^0.$$

Proof: It is sufficient to show that for all $i = 0,1,2,\ldots,$ such that $L_i > 0$, we have $g(x,\xi) \geq -L_{i+1} \geq -L_i$, since the result then follows from Proposition 4. We proceed by induction on i. Note that

$$-L_1 = \min_{\xi\in\Xi}\min_{x\in X_\phi^0}\{g(x,\xi)\} \geq -L_0$$

by the fact that $X_\phi^0 \subseteq S$ and $g(x,\xi) \geq -L = -L_0$ for all $x \in S$ and $\xi \in \Xi$. Suppose now that $-L_\tau \geq -L_{\tau-1} \geq \cdots \geq -L_0$. Then

$$-L_\tau = \min_{\xi\in\Xi}\min_{x\in X_\phi^{\tau-1}}\{g(x,\xi)\} \leq \min_{\xi\in\Xi}\min_{x\in X_\phi^\tau}\{g(x,\xi)\} = -L_{\tau+1},$$

where the inequality follows from $X_\phi^\tau \subseteq X_\phi^{\tau-1}$ due to Proposition 4. □

The scheme outlined above requires solving an optimization problem of the form (17) in every iteration. In general such a problem may be difficult. At the end of this section, we consider a class of chance consider linear programs for which the optimization problem (17) is a tractable convex program.

We now consider the case of stochastic lower bounds on $g(x,\xi)$, i.e. assumption (A0). Similar to the case of (A1), we can now consider improving lower bounds $L(\xi)$ correspond to each $\xi \in \Xi$. This is particularly relevant in the case when Ξ has finite support.

Proposition 5 *Suppose for all $x \in S$ and $\xi \in \Xi$, we have that $g(x,\xi) \geq -L_2(\xi) \geq -L_1(\xi)$ where $0 < L_2(\xi) \leq L_1(\xi)$ for all $\xi \in \Xi$. For $i = 1,2$, let $h_i(x,\xi) := g(x,\xi)/L_i(\xi) + 1$ and $X_\phi^i := \{x \in \mathbb{R}^n: \ \mathbb{E}[\phi(h_i(x,\xi))] \geq (1-\epsilon)\phi(1)\}$ for some $\phi \in \Phi$. Suppose also that $\phi \in \Phi$ satisfies $\phi(t) = \phi(1)$ for all $t \geq 1$. Then*

$$X \subseteq X_\phi^2 \subseteq X_\phi^1.$$

Proof: Since $L_1(\xi)$ and $L_2(\xi)$ correspond to valid lower bounds on $g(x,\xi)$, it follows from Proposition 1 that $X \subseteq X_\phi^t$ for $t = 1,2$. Consider $\widehat{x} \in X_\phi^2$, and let $\Xi^+ = \{\xi \in \Xi: \ g(\widehat{x},\xi) \geq 0\}$ and $\Xi^- = \{\xi \in \Xi: \ g(\widehat{x},\xi) < 0\}$. Then $h_2(\widehat{x},\xi) \geq h_1(\widehat{x},\xi) \geq 1$ for all $\xi \in \Xi^+$ and $h_2(\widehat{x},\xi) \leq h_1(\widehat{x},\xi) < 1$ for all $\xi \in \Xi^-$. Thus $\phi(h_2(\widehat{x},\xi)) = \phi(h_1(\widehat{x},\xi))$ for all $\xi \in \Xi^+$ and $\phi(h_2(\widehat{x},\xi)) \leq \phi(h_1(\widehat{x},\xi))$ for all $\xi \in \Xi^-$. It then follows that

$$\mathbb{E}[\phi(h_1(\widehat{x},\xi))] \geq \mathbb{E}[\phi(h_2(\widehat{x},\xi))] \geq (1-\epsilon)\phi(1),$$

and so $\widehat{x} \in X_\phi^1$. □

Realization	Probability	$g(\widehat{x},\xi)$	$L_1(\xi)$	$L_2(\xi)$	$h_1(\widehat{x},\xi)$	$h_2(\widehat{x},\xi)$
ξ^a	0.9996	-0.5	1.0	-0.5	0.5	0.0
ξ^b	0.0004	1.0	0.002	0.0005	501.0	2001.0

Table 3: Counter example to improving stochastic lower bounds

Following is a counter example showing that the result of Proposition 5 is not guaranteed to hold in the absence of the requirement $\phi(t) = \phi(1)$ for all $t \geq 1$. Suppose $\phi(t) = t$. Table 3 presents the data for the counter example. In this example, ξ has two possible realizations and $\widehat{x}$ is some fixed vector. Suppose $(1-\epsilon) = 0.8$. Then $\mathbb{E}[\phi(h_2(x,\xi))] = \mathbb{E}[h_2(x,\xi)] = 0.8 \geq (1-\epsilon)\phi(1)$. Thus $\widehat{x} \in X^2_\phi$. However $\mathbb{E}[\phi(h_1(x,\xi))] = \mathbb{E}[h_1(x,\xi)] = 0.7001 < (1-\epsilon)\phi(1)$. Thus $\widehat{x} \notin X^1_\phi$.

Corollary 2 *Assume (A0) and $\phi(t) = \phi(1)$ for all $t \geq 1$. Let $L_0(\xi) = L(\xi) > 0$ for all $\xi \in \Xi$, and, for all $\xi \in \Xi$, set*

$$L_{i+1}(\xi) = -\min_{x \in X^i_\phi} \{g(x,\xi)\} \text{ for } i = 0,1,2,\ldots, \tag{18}$$

as long as $L_{i+1}(\xi) > 0$, where $X^i_\phi := \{x \in S : \mathbb{E}[\phi(h_i(x,\xi))] \geq (1-\epsilon)\phi(1)\}$ and $h_i(x,\xi) := g(x,\xi)/L_i(\xi) + 1$. Then

$$X \subseteq \ldots X^{i+1}_\phi \subseteq X^i_\phi \subseteq \ldots X^0_\phi.$$

Proof: Analogous to the proof of Corollary 1 □

The above iterative improvement scheme for the case of stochastic bounds requires solving an optimization problem (18) for each realization of ξ in every iteration, and can be computationally prohibitive.

Numerical Illustration

We revisit the chance constrained linear program (10) with independent coefficients $\xi_j \sim U(a_j, b_j)$ for $j = 1,\ldots,n$, considered in Section 4. Note that in this case, for any $x \in S$,

$$\min_{\xi \in \Xi}\{g(x,\xi)\} = \min_{\xi \in \Xi}\{\sum_{j=1}^n \xi_j x_j - 1\} = \sum_{j=1}^n a_j x_j - 1.$$

Thus we can apply the iterative improvement scheme outlined in Corollary 1 to improve the Bernstein relaxation by solving the convex optimization problem

$$L_{i+1} = -\min\{\sum_{j=1}^n a_j x_j - 1 : \; x \in X^i_{\phi_\beta}\} \text{ for } i = 0,1,\ldots$$

We can stop either when L_i does not change significantly or when it gets sufficiently close to zero. Table 4 shows the effect and time consumed by the improvement procedure for problems of size $n = 100, 200$ and 500. In each case we started with $L_0 = 1$. We observe that the improvement scheme was able to improve the bounds by over 3%.

Acknowledgements

The improvement scheme of Section 5 was inspired by on-going joint research with Santanu S. Dey, Feng Qiu and Laurence Wolsey on a coefficient tightening scheme for integer programming formulations of chance constrained linear programs under discrete distributions. The author thanks Arkadi Nemirovski for some helpful comments. Partial support for this research was provided by the Air Force Office of Scientific Research under Grant FA9550-08-1-0117 and the National Science Foundation under Grant CMMI-0758234.

	$n = 100$	$n = 200$	$n = 500$
Initial L	1.0000	1.0000	1.0000
Initial bound	0.9182	0.9179	0.9177
Iterations	3	3	3
Final L	0.7174	0.7176	0.7177
Final bound	0.9419	0.9414	0.9410
Total time	2.24	3.18	3.92

Table 4: Effect of improvement scheme

References

[1] R. Aringhieri. A tabu search algorithm for solving chance-constrained programs. *Journal of the ACM*, 5:1–14, 2004.

[2] G. Calafiore and M. C. Campi. Uncertain convex programs: Randomized solutions and confidence levels. *Mathematical Programming*, 102:25–46, 2005.

[3] K. Iwamura and B. Liu. A genetic algorithm for chance constrained programming. *Journal of Information and Optimization Sciences*, 17:409–422, 1996.

[4] J. Luedtke and S. Ahmed. A sample approximation approach for optimization with probabilistic constraints. *SIAM Journal on Optimization*, 19:674–699, 2008.

[5] A. Nemirovski and A. Shapiro. Convex approximations of chance constrained programs. *SIAM Journal on Optimization*, 17:969–996, 2006.

[6] B. Pagnoncelli, S. Ahmed, and A. Shapiro. The sample average approximation method for chance constrained programming: Theory and applications. *Journal of Optimization theory and Applications*, 142:399–416, 2009.

[7] J. Pinter. Deterministic approximations of probability inequalities. *Mathematical Methods of Operations Research*, 33: 219–239, 1989.

[8] A. Prekopa. *Stochastic Programming.* Kluwer Academic Publishers, Dordrecht, The Netherlands, 1995.

[9] A. Shapiro, D. Dentcheva and A. Ruszczynski. *Lectures on Stochastic Programming: Modeling and Theory.* MOS-SIAM Series on Optimization, vol.9, SIAM, Philadelphia, 2009.