INFORMS TutORials in Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Solving Chance-Constrained Stochastic Programs via Sampling and Integer Programming

Shabbir AhmedAlexander Shapiro

To cite this entry: Shabbir AhmedAlexander Shapiro. Solving Chance-Constrained Stochastic Programs via Sampling and Integer Programming. *In* INFORMS TutORials in Operations Research. Published online: 14 Oct 2014; 261-269.
https://doi.org/10.1287/educ.1080.0048

TUTORIALS IN
OPERATIONS RESEARCH
INFORMS 2008

 | ISBN 978-1-877640-23-0
DOI 10.1287/educ.1080.0048

Solving Chance-Constrained Stochastic Programs via Sampling and Integer Programming

Shabbir Ahmed and Alexander Shapiro
H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332 {sahmed@isye.gatech.edu, ashapiro@isye.gatech.edu}

Abstract Various applications in reliability and risk management give rise to optimization problems with constraints involving random parameters, which are required to be satisfied with a prespecified probability threshold. There are two main difficulties with such chance-constrained problems. First, checking feasibility of a given candidate solution exactly is, in general, impossible because this requires evaluating quantiles of random functions. Second, the feasible region induced by chance constraints is, in general, nonconvex, leading to severe optimization challenges. In this tutorial, we discuss an approach based on solving approximating problems using Monte Carlo samples of the random data. This scheme can be used to yield both feasible solutions and statistical optimality bounds with high confidence using modest sample sizes. The approximating problem is itself a chance-constrained problem, albeit with a finite distribution of modest support, and is an NP-hard combinatorial optimization problem. We adopt integer-programming-based methods for its solution. In particular, we discuss a family valid inequalities for a integer programming formulations for a special but large class of chance-constrained problems that have demonstrated significant computational advantages.

Keywords chance-constrained problems; stochastic programming; sampling; integer programming

1. Introduction

A large class of optimization problems arising from important planning and design applications in uncertain environments involve service level or reliability constraints. Consider, for example, the problem of locating service centers for responding to medical emergencies. Requiring 100% coverage over all possible emergency scenarios is physically and economically impractical, and, therefore, typically emergency preparedness plans calls for some minimum response reliability (Aly and White [1], Beraldi et al. [4]). Service-level agreements in telecommunication contracts require network providers to guarantee, with high probability, that packet losses will not exceed a certain percentage (Marianov and Rios [20], Ventetsanopoulos and Singh [32]). In financial portfolio planning, investors often require that, with high probability, portfolio losses do not exceed some threshold (value at risk) while maximizing expected returns (Gaivoronski and Pflug [11], Pagnoncelli et al. [25]). Mathematical models for planning/designing reliability-constrained systems such as these lead to optimization with chance constraints or probabilistic constraints.

A generic chance-constrained optimization problem can be formulated as

$$\min_{x\in X} f(x) \quad \text{subject to} \quad \Pr\{G(x,\boldsymbol{\xi}) \le \mathbf{0}\} \ge 1-\varepsilon, \tag{1}$$

where $X \subset \mathbb{R}^n$ represents a deterministic feasible region, $f\colon \mathbb{R}^n \to \mathbb{R}$ represents the objective to be minimized, $\boldsymbol{\xi}$ is a random vector whose probability distribution is supported on set $\Xi \subset \mathbb{R}^d$, $G\colon \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^m$ is a constraint mapping, $\mathbf{0}$ is an m-dimensional vector of zeroes,

and $\varepsilon \in (0,1)$ is a given risk parameter (significance level). Formulation (1) seeks a decision vector x from the feasible set X that minimizes the function $f(x)$ while satisfying the chance constraint $G(x,\boldsymbol{\xi}) \le \mathbf{0}$ with probability at least $1-\varepsilon$. It is assumed that the probability distribution of $\boldsymbol{\xi}$ is known.

By way of illustration, consider the following simple facility-sizing example. We need to decide capacities of n facilities servicing an uncertain customer demand. The cost per unit capacity installed for each facility is given, as is the joint demand distribution. The goal is to determine the cheapest capacity configuration so as to guarantee that the installed capacity exceeds demand with probability $1-\varepsilon$. This chance-constrained problem can be formulated as follows:

$$\min_{x \ge 0} \sum_{i=1}^{n} c_i x_i \quad \text{subject to} \quad \Pr\{\xi_i - x_i \le 0,\ i = 1,\dots,n\} \ge 1-\varepsilon. \tag{2}$$

Here, x_i, c_i, and ξ_i denote the capacity, cost, and random demand for facility i, respectively. It is assumed that the (joint) probability distribution of the random vector $\boldsymbol{\xi} = (\xi_1,\dots,\xi_n)$ is *known* (otherwise, the probabilistic constraint in (2) is not defined). Note that the probabilistic (chance) constraint of (2) can be considerably weaker than trying to satisfy the demand for *all* possible realizations of $\boldsymbol{\xi}$. Note also that (2) is an example of (1) with $G(x,\boldsymbol{\xi}) = \boldsymbol{\xi} - x$.

In this example, we require that the reliability requirement be applied to all facilities jointly. One could also consider the individual chance constraints $\Pr\{\xi_i \le x_i\} \ge 1-\varepsilon_i$, $i = 1,\dots,n$, applied to each facility separately. This leads to a much simpler problem because $\Pr\{\xi_i \le x_i\} \ge 1-\varepsilon_i$ is equivalent to $F_i^{-1}(x_i) \ge 1-\varepsilon_i$, where F_i is the cumulative distribution function (cdf) of ξ_i. Note, however, that to ensure the joint chance constraint by enforcing the individual chance constraints, the corresponding risk parameters ε_i should be considerably smaller than ε, especially when n is large.

Beginning with the work of Charnes et al. [7], chance-constrained stochastic programs have been studied extensively. In addition to the facility location, telecommunication, and finance examples cited earlier, chance-constrained models have been used in numerous other applications, including production planning (Murr and Prékopa [22], Lejeune and Ruszczyński [16]), chemical processing (Henrion et al. [15], Henrion and Möller [14]), and water resources management (Prékopa and Szántai [29], Takyi and Lence [31]). See Prékopa [28] for a background and an extensive list of references. Despite important theoretical progress and practical importance, chance-constrained stochastic problems of the form (1) are still largely intractable except for some special cases. There are two primary reasons for this difficulty:

1. In general, for a given $x \in X$, computing $\Pr\{G(x,\boldsymbol{\xi}) \le \mathbf{0}\}$ accurately, i.e., checking whether x is feasible to (1), can be hard. In multidimensional situations, this involves calculation of a multivariate integral, which typically cannot be computed with a high accuracy.

2. The feasible region defined by a chance constraint generally is not convex even if $G(x,\boldsymbol{\xi})$ is convex in x for every possible realization of $\boldsymbol{\xi}$. This implies that even if checking feasibility is easy, optimization of the problem remains difficult. For example, the facility-sizing example (2) with n facilities and m equiprobable realizations of the demand vector $\boldsymbol{\xi}$ is equivalent to a maximum clique problem on a graph with n nodes and m edges, and is therefore strongly NP-hard (Luedtke et al. [19]).

In light of the above difficulties, existing approaches for chance-constrained stochastic programs can be classified as follows. First are the approaches for problems where both difficulties are absent; i.e., the distribution of $\boldsymbol{\xi}$ is such that checking feasibility is easy, and the resulting feasible region is convex. A classical example of this case is when $G(x,\boldsymbol{\xi}) = v - \boldsymbol{\xi}^\top x$, and $\boldsymbol{\xi}$ has a multivariate normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Then, for $\varepsilon \in (0, 0.5)$,

$$\{x \in \mathbb{R}^n\colon \Pr\{\boldsymbol{\xi}^\top x \ge v\} \ge 1-\varepsilon\} = \{x \in \mathbb{R}^n\colon v - \boldsymbol{\mu}^\top x + z_\varepsilon \sqrt{x^\top \Sigma\, x} \le 0\},$$

where $z_\varepsilon = \Phi^{-1}(1-\varepsilon)$ is the $(1-\varepsilon)$-quantile of the standard normal distribution. In this case, under convexity of X, the chance-constrained problem reduces to a deterministic convex optimization problem. The second class of approaches are for problems where only the second difficulty is absent; i.e., the feasible region of the chance constraint is guaranteed to be convex. The best known example of this case is when $G(x,\boldsymbol{\xi}) = \boldsymbol{\xi} - \mathbf{A}x$, where $\mathbf{A}$ is a deterministic matrix and $\boldsymbol{\xi}$ has a *log-concave* distribution. In this case, the chance constraint feasible set is convex (Prékopa [26]). However, it may still be difficult to compute $\Pr\{G(x,\boldsymbol{\xi}) \le \mathbf{0}\}$ exactly. Solution methods in this class are primarily based on classical nonlinear programming techniques adapted with suitable approximations of the chance constraint function and its gradients (see Prékopa [27]). The third class of approaches are for problems where the first difficulty is absent; i.e., computing $\Pr\{G(x,\boldsymbol{\xi}) \le \mathbf{0}\}$ is easy, e.g., when $\boldsymbol{\xi}$ has a finite distribution with a modest number of realizations (in this case the feasible region is typically nonconvex). A number of approaches based on integer programming and global optimization have been developed for this class of problems (Cheon et al. [8], Dentcheva et al. [10], Ruszczyński [30]). Finally, more recently, a number of approaches have been proposed to deal with both difficulties (Calafiore and Campi [5, 6], Nemirovski and Shapiro [23, 24], Ben-Tal and Nemirovski [3]). The common theme in these approaches is that they all propose convex approximations of the nonconvex chance constraint that yield solutions that are feasible, or at least highly likely to be feasible, to the original problem. Thus, the difficulty of checking feasibility as well as nonconvexity is avoided. Unfortunately, often, the solutions produced by these approaches are quite conservative.

In this tutorial, we consider an approximation of the chance-constrained problem (1) where the true distribution of $\boldsymbol{\xi}$ is replaced by an empirical distribution corresponding to a Monte Carlo sample. The resulting *sample average approximation* problem can be used to provide good candidate solutions along with optimality gap estimates. The sampled approximation problem is a chance-constrained problem with a discrete distribution and can be quite difficult. We discuss integer-programming-based approaches for solving it.

2. Sample Average Approximation

To simplify the presentation we assume, without loss of generality, that the constraint function $G\colon \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}$ in (1) is scalar valued. Of course, a number of constraints $G_i(x,\xi) \le 0,\ i = 1,\dots,m$, can be equivalently replaced by one constraint $G(x,\xi) := \max_{1\le i\le m} G_i(x,\xi) \le 0$. The chance-constrained stochastic program (1) can be rewritten as

$$\min_{x\in X} f(x) \quad \text{subject to} \quad q(x) \le \varepsilon, \tag{3}$$

where $q(x) := \Pr\{G(x,\boldsymbol{\xi}) > 0\}$.

Now let $\boldsymbol{\xi}^1,\dots,\boldsymbol{\xi}^N$ be an *independent identically distributed* (iid) sample of N realizations of random vector $\boldsymbol{\xi}$. Given $x \in X$, let

$$\hat{q}_N(x) := N^{-1} \sum_{j=1}^{N} \mathbb{1}_{(0,\infty)}(G(x,\boldsymbol{\xi}^j)),$$

where $\mathbb{1}_{(0,\infty)}\colon \mathbb{R} \to \mathbb{R}$ is the indicator function of $(0,\infty)$. That is, $\hat{q}_N(x)$ is equal to the proportion of realizations with $G(x,\boldsymbol{\xi}^j) > 0$ in the sample. For some given $\gamma \in (0,1)$, consider the following optimization problem associated with a sample $\boldsymbol{\xi}^1,\dots,\boldsymbol{\xi}^N$,

$$\min_{x\in X} f(x) \quad \text{subject to} \quad \hat{q}_N(x) \le \gamma. \tag{4}$$

We refer to problems (3) and (4) as the true and *sampled average approximate* (SAA) problems, respectively, at the respective risk levels ε and γ.

The SAA problem is a chance-constrained stochastic problem with a different (discrete) distribution and a different risk level than (3). Unless N is prohibitively large, the chance-constrained problem SAA does not suffer from the first difficulty (computing $\hat{q}_N(x)$) mentioned in §1; however, it may still be difficult to solve. Assuming we have a scheme for solving the SAA problem, what can we say about an optimal solution and the optimal value of the SAA in relation to that of the true problem (3)? Intuitively, assuming N is large enough, if $\gamma \le \varepsilon$, then a feasible solution of the SAA is likely to be feasible to the true problem, and if $\gamma \ge \varepsilon$, then the optimal value of the SAA is likely to be a lower bound to that of the true problem. Thus, the SAA problem can be used to obtain both candidate feasible solutions to the true problem as well as optimality gap estimates. Next, we discuss these concepts slightly more rigorously.

We assume that X is compact, $f(\cdot)$ is continuous, $G(x,\cdot)$ is measurable for every $x \in \mathbb{R}^n$, and $G(\cdot,\boldsymbol{\xi})$ is continuous for almost every $\boldsymbol{\xi}$. Then, the functions $q(x)$ and $\hat{q}_N(x)$ are lower-semicontinuous, and the true problem (3) and the SAA problem (4) are guaranteed to have optimal solutions if they are feasible. Let $X^*(\varepsilon)$ and $\hat{X}_N(\gamma)$ denote the set of optimal solutions of the true and SAA problems, respectively, and $v(\varepsilon)$ and $\hat{v}_N(\gamma)$ denote the optimal value of the true and SAA problems, respectively.

2.1. Convergence Properties

Under reasonable regularity assumptions it can be shown that for $\gamma = \varepsilon$ the optimal value $\hat{v}_N(\gamma)$ and optimal solutions of set $\hat{X}_N(\gamma)$ of the SAA problem converge to their true counterparts $v(\varepsilon)$ and $X^*(\varepsilon)$ with one probability as N approaches infinity (Pagnoncelli et al. [25]). Suppose now that $\gamma > \varepsilon$. Then, we may expect that with increase of the sample size N, an optimal solution of the SAA problem will approach an optimal solution of the true problem with the risk level γ rather than ε. Of course, increasing the risk level leads to enlarging the feasible set of the true problem, which in turn may result in decreasing the optimal value of the true problem. For a point $\bar{x} \in X$ we have that $\hat{q}_N(\bar{x}) \le \gamma$; i.e., $\bar{x}$ is a feasible solution to the SAA problem if and only if no more than γN times the event "$G(\bar{x},\boldsymbol{\xi}^j) > 0$" happens in N trials. Because the probability of the event "$G(\bar{x},\boldsymbol{\xi}^j) > 0$" is $q(\bar{x})$, it follows that

$$\Pr\{\hat{q}_N(\bar{x}) \le \gamma\} = B(\lfloor \gamma N \rfloor; q(\bar{x}), N), \tag{5}$$

where

$$B(k;q,N) := \sum_{i=0}^{k} \binom{N}{i} q^i (1-q)^{N-i}, \quad k = 0,\dots,N, \tag{6}$$

denotes the cumulative density function (cdf) of binomial distribution. Recall that by the Chernoff [9] inequality for $k > Np$,

$$B(k;q,N) \ge 1 - \exp\{-N(k/N - q)^2/(2q)\}.$$

It follows that if $q(\bar{x}) \le \varepsilon$ and $\gamma > \varepsilon$, then $1 - \Pr\{\hat{q}_N(\bar{x}) \le \gamma\}$ approaches zero at a rate of $\exp(-\kappa N)$, where $\kappa := (\gamma - \varepsilon)^2/(2\varepsilon)$. Of course, if $\bar{x}$ is an optimal solution of the true problem and $\bar{x}$ is a feasible point of the SAA problem, then $\hat{v}_N(\gamma) \le v^*(\varepsilon)$. That is, if $\gamma > \varepsilon$, then the probability of the event "$\hat{v}_N(\gamma) \le v^*(\varepsilon)$" approaches one exponentially fast. By similar analysis it can be shown that if $q(\bar{x}) \le \gamma$, i.e., $\bar{x}$ is a feasible solution of SAA and $\gamma < \varepsilon$, then the probability that $\bar{x}$ is a feasible solution of the true problem approaches one exponentially fast (see Luedtke and Ahmed [18]). Based on this analysis, we can compute a priori the sample size required in the SAA problem so that it produces a feasible solution to the true problem with high probability (typically such estimates of a required sample size are quite conservative). Next we discuss techniques for assessing the quality (feasibility and optimality gap) of an arbitrary candidate solution to the true problem.

2.2. Solution Validation

For a given candidate point $\bar{x} \in X$, say obtained as a solution of an SAA problem, we would like to validate its quality as a solution of the true problem. This involves two questions, namely, whether $\bar{x}$ is a feasible point of the true problem, and, if yes, then what is the optimality gap $f(\bar{x}) - v(\varepsilon)$. Of course, if $\bar{x}$ is a feasible point of the true problem, then $f(\bar{x}) - v(\varepsilon)$ is nonnegative and is zero if and only if $\bar{x}$ is an optimal solution of the true problem.

Let us start with verification of feasibility of $\bar{x}$. For that we need to estimate the probability $q(\bar{x})$. We proceed by employing again the Monte Carlo sampling technique. Generate an iid sample $\xi^1, \ldots, \xi^{N'}$ and estimate $q(\bar{x})$ by $\hat{q}_{N'}(\bar{x})$. Note that this random sample should be generated independently of a random procedure, which produced the candidate solution $\bar{x}$, and that we can use a very large sample, of size N', because we do not need to solve any optimization problem here. The estimator $\hat{q}_{N'}(\bar{x})$ of $q(\bar{x})$ is unbiased, and for large N' and not "too small" $q(\bar{x})$ its distribution can be approximated reasonably well by the normal distribution with mean $q(\bar{x})$ and variance $q(\bar{x})(1 - q(\bar{x}))/N'$. This leads to the following approximate $(1-\beta)$-confidence upper bound on $q(\bar{x})$:

$$U_{\beta, N'}(\bar{x}) := \hat{q}_{N'}(\bar{x}) + z_\beta \sqrt{\hat{q}_{N'}(\bar{x})(1 - \hat{q}_{N'}(\bar{x}))/N'}, \tag{7}$$

where $z_\beta = \Phi^{-1}(1-\beta)$. We can now compare $U_{\beta, N}(\bar{x})$ to ε to check the feasibility of $\bar{x}$ (see Nemirovski and Shapiro [24] for a slightly more accurate confidence upper bound).

To get a lower bound for the optimal value $v(\varepsilon)$, we proceed as follows. Let us choose two positive integers M and N, and let

$$\theta_N := B(\lfloor \gamma N \rfloor; \varepsilon, N)$$

and L be the *largest* integer such that

$$B(L-1; \theta_N, M) \le \beta. \tag{8}$$

Next, generate M independent samples $\xi^{1,m}, \ldots, \xi^{N,m}$, $m = 1, \ldots, M$, each of size N, of random vector ξ. For each sample, solve the associated SAA problem and record the corresponding optimal objective values $\hat{v}_N^m(\gamma)$, $m = 1, \ldots, M$. It may happen that the SAA problem is either infeasible or unbounded from below, in which case we assign its optimal value as $+\infty$ or $-\infty$, respectively. We can view $\hat{v}_N^m(\gamma)$, $m = 1, \ldots, M$, as an iid sample of the random variable $\hat{v}_N(\gamma)$. Next, we sort the calculated optimal values in nondecreasing order, i.e., $\hat{v}_N^{(1)}(\gamma) \le \cdots \le \hat{v}_N^{(M)}(\gamma)$. It is possible to show that with probability at least $1-\beta$, the random quantity $\hat{v}_N^{(L)}(\gamma)$ is a lower bound of the true optimal value $v(\varepsilon)$ (Luedtke and Ahmed [18], Nemirovski and Shapiro [24]).

Extensive computational results on the performance of the above-mentioned validation schemes on various classes of chance-constrained problems are reported in Luedtke and Ahmed [18], Nemirovski and Shapiro [24], and Pagnoncelli et al. [25].

3. Solving Sample Approximations

We have seen that we can generate as well as validate candidate solutions to the chance-constrained problem (3) by solving (several) sampled approximations (4). In this section, we explore approaches for solving these problems.

If $\gamma = 0$, then the SAA problem reduces to

$$\min_{x \in X} f(x) \quad \text{subject to } G(x, \xi^j) \le 0 \quad j = 1, \ldots, N. \tag{9}$$

When the functions $f(\cdot)$ and $G(\cdot, \xi^j)$ for $j = 1, \ldots, N$ are convex (linear) and the set X is convex (polyhedral), then (9) is a convex (linear) program and can usually be solved

efficiently using off-the-shelf software. From the convergence discussion in §2.1 we can see that the candidate solutions generated from (9) can often be overly conservative in terms of the objective function. We can then consider increasing the risk level γ in the SAA problem. However, with $\gamma > 0$, the SAA problem is a chance-constrained optimization problem (with a finite distribution) and is NP-hard even in very simple settings (such as the facility-sizing example discussed in §1) (Luedtke et al. [19]). A wide variety of approaches have been proposed to solve different classes of chance-constrained optimization problems under finite distributions (cf. Cheon et al. [8], Dentcheva et al. [10], Prékopa [28] and references therein). In this tutorial, we consider an integer-programming-based approach.

The SAA problem (4) can be formulated as the following mixed-integer problem (MIP):

$$\begin{aligned} \min \quad & f(x) \\ \text{subject to} \quad & G(x,\xi^j) \le M_j z_j \quad j = 1,\dots,N \\ & \sum_{j=1}^{N} z_j \le \gamma N \\ & z_j \in \{0,1\} \quad j = 1,\dots,N \\ & x \in X, \end{aligned} \tag{10}$$

where z_j is a binary variable and M_j is a large positive number such that $M_j \ge \max_{x\in X} G(x,\xi^j)$ for all $j = 1,\dots,N$. Note that if z_j is 0, then the constraint $G(x,\xi^j) \le 0$ corresponding to the realization j in the sample is enforced. However, $z_j = 1$ does not pose any restriction on $G(x,\xi^j)$. The cardinality constraint $\sum_{j=1}^{N} z_j \le \gamma N$ requires that at least γN of the N constraints $G(x,\xi^j) \le 0$ for $j = 1,\dots,N$ are enforced.

Even in a linear setting (i.e., the functions f and G are linear in x and the set X is polyhedral), moderate-sized instances of the MIP (10) are typically very difficult to solve as is by state-of-the-art MIP solvers. The difficulty is due to the fact that the continuous relaxation of (10) (obtained by dropping the integrality restriction on the z variables) provides a weak relaxation, and, hence, slows down the branch-and-bound algorithm that is the work horse of MIP solvers. This difficulty can be alleviated by strengthening the formulation (10) by addition of valid inequalities or reformulation. Such improved formulations have tighter continuous relaxation gaps and can serve to significantly cut down solve times.

A variety of approaches for strengthening special classes of the MIP (10) have been proposed. Here we discuss an approach for the case of joint probabilistic constraints where the uncertain parameters only appear on the right-hand side; i.e.,

$$G(x,\xi) = \max_{i=1,\dots,m} \{\xi_i - G_i(x)\}.$$

Note that the facility sizing example (2) is of this form. By appropriately translating, we assume that $\xi_j^i \ge 0$ for all i and j. The MIP (10) can then be written as

$$\begin{aligned} \min \quad & f(x) \\ \text{subject to} \quad & G_i(x) \ge v_i \quad i = 1,\dots,m \\ & v_i + \xi_i^j z_j \ge \xi_i^j \quad i = 1,\dots,m, \;\; j = 1,\dots,N \\ & \sum_{j=1}^{N} z_j \le \gamma N \\ & z_j \in \{0,1\} \quad j = 1,\dots,N \\ & x \in X, \quad v_i \ge 0 \quad i = 1,\dots,m. \end{aligned} \tag{11}$$

Note that we have introduced the auxiliary variables v_i for $i = 1, \ldots, m$ to conveniently represent $G_i(x)$. As before, if z_j is 0, then the constraints $G_i(x) \geq \xi_i^j$ for $i = 1, \ldots, m$ corresponding to the realization j in the sample are enforced. Consider now the following subsystem corresponding to the ith row of the probabilistic constraint system:

$$F_i := \left\{ (v_i, z_1, \ldots, z_N) \in \mathbb{R}_+ \times \{0,1\}^N \colon v_i + \xi_i^j z_j \geq \xi_i^j \;\; j = 1, \ldots, N, \sum_{j=1}^{N} z_j \leq \gamma N \right\}.$$

Note that the set of feasible vectors $(v_i, z_1, \ldots, z_N)$ for the MIP (11) are contained in F_i and, hence, in its convex hull $\text{conv}(F_i)$. Thus, any valid inequality for $\text{conv}(F_i)$ will also be valid for (11) and can potentially strengthen its continuous relaxation. Next, we describe a family of such inequalities. Without loss of generality, we can assume that ξ_j^i are indexed such that $\xi_i^1 \geq \xi_i^2 \geq \ldots \geq \xi_i^N$. Let $p := \lfloor \gamma N \rfloor$. The cardinality constraint $\sum_{j=1}^{N} z_j \leq \gamma N$ implies that we cannot have $z_j = 1$ for all $j = 1, \ldots, p+1$; thus, $v_i \geq \xi_i^{p+1}$. This also implies that the constraints $v_i + \xi_i^j z_j \geq \xi_i^j$ for $j = p+1, \ldots, N$ are redundant. We can therefore tighten the formulation of F_i to

$$F_i' := \left\{ (v_i, z_1, \ldots, z_N) \in \mathbb{R}_+ \times \{0,1\}^N \colon v_i + (\xi_i^j - \xi_i^{p+1}) z_j \geq \xi_i^j \;\; j = 1, \ldots, p, \sum_{j=1}^{N} z_j \leq \gamma N \right\}.$$

The subsystem obtained by dropping the cardinality constraint $\sum_{j=1}^{N} z_j \leq \gamma N$ from F_i' is the well-studied *mixing set* (Atamtürk et al. [2], Günlük and Pochet [13], Guan et al. [12], Miller and Wolsey [21])

$$M_i := \left\{ (v_i, z_1, \ldots, z_p) \in \mathbb{R}_+ \times \{0,1\}^p \colon v_i + \xi_i^j z_j \geq \xi_i^j \;\; j = 1, \ldots, N \right\}.$$

It is known that the convex hull of M_i is completely characterized by the so-called *star inequalities*

$$v_i + \sum_{k=1}^{l} (\xi_i^{j_k} - \xi_i^{j_{k+1}}) z_{j_k} \geq \xi_i^{j_1} \quad \forall\, J := \{j_1, \ldots, j_l\} \subseteq \{1, \ldots `, p\}, \tag{12}$$

where $\xi_i^{j_{l+1}} := \xi_i^{p+1}$. Clearly, the star inequalities (12) are valid for F_i and, hence, for (11). It turns out that the above star inequalities define some but not all facets of $\text{conv}(F_i)$. Even though there are exponentially many such star inequalities (one for each subset J), these can be separated very efficiently (Atamtürk et al. [2]). It has been observed that the addition of the above star inequalities within a branch-and-cut framework for solving linear instances of (11) can have tremendous computational benefits (Luedtke et al. [19]). A number of additional classes of valid inequalities and an extended reformulation for F_i that also provide a significant computational advantage in solving (10) have been developed (Luedtke [17]).

4. Conclusion

The stochasticity and nonconvexity associated with chance-constrained stochastic programs make these extremely hard to solve. In this tutorial, we discussed a sampling-based approach wherein we approximated the problem by replacing the distribution of the uncertain parameters by an empirical distribution corresponding to a Monte Carlo sample. The approximate problem, called a sample average approximation, is still a chance-constrained problem, albeit with a finite distribution of modest support. The SAA problem can serve to provide both candidate solutions as well as solution quality estimates. Because the SAA problem is an NP-hard combinatorial problem, we adopt integer-programming-based methods for its solution. In particular, we discussed a family valid inequalities for a integer programming formulations for a special but large class of chance-constrained problems that have demonstrated significant computational advantages.

References

[1] A. Aly and J. White. Probabilistic formulation of the emergency service location problem. *The Journal of the Operational Research Society* 29:1167–1179, 1978.

[2] A. Atamtürk, G. L. Nemhauser, and M. W. P. Savelsbergh. The mixed vertex packing problem. *Mathematical Programming* 89:35–53, 2000.

[3] A. Ben-Tal and A. Nemirovski. Robust convex optimization. *Mathematics of Operations Research* 23:769–805, 1998.

[4] P. Beraldi, M. Bruni, and D. Conforti. Designing robust emergency medical service via stochastic programming. *European Journal of Operational Research* 158:183–193, 2004.

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

[6] G. C. Calafiore and M. C. Campi. The scenario approach to robust control design. *IEEE Transactions on Automatic Control* 51:742–753, 2006.

[7] A. Charnes, W. W. Cooper, and G. H. Symonds. Cost horizons and certainty equivalents: An approach to stochastic programming of heating oil. *Management Science* 4:235–263, 1958.

[8] M.-S. Cheon, S. Ahmed, and F. Al-Khayyal. A branch-reduce-cut algorithm for the global optimization of probabilistically constrained linear programs. *Mathematical Programming* 108:617–634, 2006.

[9] H. Chernoff. A measure of asymptotic efficiency for tests of a hypothesis based on the sum observations. *Annals of Mathematical Statistics* 23:493–507, 1952.

[10] D. Dentcheva, A. Prékopa, and A. Ruszczyński. Concavity and efficient points of discrete distributions in probabilistic programming. *Mathematical Programming* 89:55–77, 2000.

[11] A. Gaivoronski and G. Pflug. Value-at-risk in portfolio optimization: Properties and computational approach. *The Journal of Risk* 7:1–33, 2005.

[12] Y. Guan, S. Ahmed, and G. L. Nemhauser. Sequential pairing of mixed integer inequalities. *Discrete Optimization* 4:21–39, 2007.

[13] O. Günlük and Y. Pochet. Mixing mixed-integer inequalities. *Mathematical Programming* 90:429–457, 2001.

[14] R. Henrion and A. Möller. Optimization of a continuous distillation process under random inflow rate. *Computers and Mathematics with Applications* 45:247–262, 2003.

[15] R. Henrion, P. Li, A. Möller, M. C. Steinbach, M. Wendt, and G. Wozny. Stochastic optimization for operating chemical processes under uncertainty. M. Grötschel, S. Krunke, and J. Rambau, eds. *Online Optimization of Large Scale Systems*. Springer, Berlin, 457–478, 2001.

[16] M. A. Lejeune and A. Ruszczyński. An efficient trajectory method for probabilistic inventory-production-distribution problems. *Operations Research* 55:378–394, 2007.

[17] J. Luedtke. Integer programming approaches to two non-convex optimization problems. Ph.D. thesis, Georgia Institute of Technology, Atlanta, 2007.

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

[19] J. Luedtke, S. Ahmed, and G. Nemhauser. An integer programming approach for linear programs with probabilistic constraints. M. Fischetti and D. Williamson, eds. *Integer Programming and Combinatorial Optimization, 12th International IPCO Conference, Ithaca, NY, Proceedings, Lecture Notes in Computer Science*, Vol 4513. Springer-Verlag, Berlin, 410–423, 2007.

[20] V. Marianov and M. Rios. A probabilistic quality of service constraint for a location model of switches in ATM communications networks. *Annals of Operations Research* 96:237–243, 2000.

[21] A. J. Miller and L. A. Wolsey. Tight formulations for some simple mixed integer programs and convex objective integer programs. *Mathematical Programming* 98:73–88, 2003.

[22] M. R. Murr and A. Prékopa. Solution of a product substitution problem using stochastic programming. S. P. Uryasev, ed. *Probabilistic Constrained Optimization: Methodology and Applications*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 252–271, 2000.

[23] A. Nemirovski and A. Shapiro. Scenario approximation of chance constraints. G. Calafiore and F. Dabbene, eds. *Probabilistic and Randomized Methods for Design Under Uncertainty*. Springer, London, 3–48, 2005.

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

[25] B. Pagnoncelli, S. Ahmed, and A. Shapiro. Computational study of a chance constrained portfolio selection problem. Technical report, School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta. E-print available at http://www.optimization-online.org, 2008.

[26] A. Prékopa. On probabilistic constrained programmming. H. W. Kuhn, ed. *Proceedings of the Princeton Symposium on Mathematical Programming*, Princeton University Press, Princeton, NJ, 113–138, 1970.

[27] A. Prékopa. *Stochastic Programming*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1995.

[28] A. Prékopa. Probabilistic programming. A. Ruszczyński and A. Shapiro, eds. *Stochastic Programming*, Vol. 10 of *Handbooks in Operations Research and Management Science*. Elsevier, London, 267–351, 2003.

[29] A. Prékopa and T. Szántai. Flood control reservoir system design. *Mathematical Programming Study* 9:138–151, 1978.

[30] A. Ruszczyński. Probabilistic programming with discrete distributions and precedence constrained knapsack polyhedra. *Mathematical Programming* 93:195–215, 2002.

[31] A. K. Takyi and B. J. Lence. Surface water quality management using a multiple-realization chance constraint method. *Water Resources Research* 35:1657–1670, 1999.

[32] A. N. Ventetsanopoulos and I. Singh. Topological optimization of communication networks subject to reliability constraints. *Problem of Control and Information Theory* 15:63–78, 1986.