KYBERNETIKA — VOLUME *48* (2012), NUMBER 1, PAGES 105–122

CHANCE CONSTRAINED PROBLEMS: PENALTY REFORMULATION AND PERFORMANCE OF SAMPLE APPROXIMATION TECHNIQUE

Martin Branda

We explore reformulation of nonlinear stochastic programs with several joint chance constraints by stochastic programs with suitably chosen penalty-type objectives. We show that the two problems are asymptotically equivalent. Simpler cases with one chance constraint and particular penalty functions were studied in [6, 11]. The obtained problems with penalties and with a fixed set of feasible solutions are simpler to solve and analyze then the chance constrained programs. We discuss solving both problems using Monte-Carlo simulation techniques for the cases when the set of feasible solution is finite or infinite bounded. The approach is applied to a financial optimization problem with Value at Risk constraint, transaction costs and integer allocations. We compare the ability to generate a feasible solution of the original chance constrained problem using the sample approximations of the chance constraints directly or via sample approximation of the penalty function objective.

Keywords: chance constrained problems, penalty functions, asymptotic equivalence, sample approximation technique, investment problem

Classification: 93E12, 62A10

1. INTRODUCTION

Stochastic programming treats problems where optimization and uncertainty appears together. Such problems arise in economy, finance, industry, agriculture, logistics and civil engineering, cf. [24, 26].

In general, we consider the following program with a random factor

$$\min \{ f(x) \,:\, x \in X, g_i(x,\omega) \le 0,\, i = 1,\dots,k \}, \tag{1}$$

where $g_i, i = 0,\dots,k$, are real functions on $\mathbb{R}^n \times \mathbb{R}^{n'}$, $X \subseteq \mathbb{R}^n$ and $\omega \in \mathbb{R}^{n'}$ is a realization of a n'-dimensional random vector defined on a probability space $(\Omega, \mathcal{F}, P)$. However, ω is unknown for us, hence a question is how to deal with the uncertain constraints. In [17], three suggestions how to deal with the stochastic constraints of the form $g_i(x,\omega) = \omega_i - h_i(x) \le 0,\ i = 1,\dots,k$, where ω_i are random bounds with marginal distributions P_i, are introduced. First, the constraints can be incorporated into the objective function of the optimization problems as the penalty function

$$\sum_{i=1}^{k} N_i \int_{h_i(x)}^{\infty} [\omega_i - h_i(x)] P_i(\mathrm{d}\omega)$$

with $N_i > 0$ being constant. Next, the reliability type model with a chance or probabilistic constraint can be considered

$$P(h_i(x) \geq \omega_i,\ i = 1, \dots, k) \geq 1 - \varepsilon$$

for some level $\varepsilon \in (0, 1)$. Finally, the constraints involving the conditional expectations can be used

$$\mathbb{E}[\omega_i - h_i(x) | \omega_i - h_i(x) > 0] \leq l_i,\ i = 1, \dots, k$$

for some small levels $l_i > 0$.

Solving the chance constrained problems is not easy. In general, the feasible region is not convex even if the functions are convex and in many cases it is even not easy to check feasibility of a point because it leads to computation of multivariate integrals. On the other hand, there are some special cases under which the convexity is preserved, e. g. the log-concave distributions [19], or it is relatively easy to check the feasibility, e. g. for the normal distribution. There are several methods for numerical solving of chance constrained problems, you can see [20]. For the problems with discretely distributed random variables, p-efficient points can be used, cf. [18]. For continuously distributed random variables the methods based on supporting hyperplanes and reduced gradients are available. In the case that the underlying distribution is continuous or discrete with many realizations, sample approximation techniques and mixed-integer programming reformulation can help us to solve the problem approximately, see [1, 13, 15].

In this paper, we will study relation between the nonlinear problems with several chance constraints and the problems with penalty function objective. We will show that the model with chance constraints and the penalty type model are asymptotically equivalent under quite mild assumptions. In [11], it was shown that equivalence between the problem with one joint chance constraint and the problem with simple penalty function holds. The approach was recently extended to a whole class of penalty functions in [6]. We propose further extension to multiple jointly chance constrained problems which cover the joint as well as the separate chance constrained problems as special cases.

The approach for solving nonlinear deterministic programs with several constraints using the penalty functions is well studied in literature. Algorithms and basic theory based on continuity and Karush–Kuhn–Tucker conditions are explained in [3] and [14]. Theoretical analysis of the penalty function method is provided by [22]. The penalized objective function epiconverges to the objective function of the nonlinear problem with several constraints, which implies “stable” behaviour of optimal values and optimal solutions.

We will show that the penalty function approach can be helpful in numerical solution of stochastic optimization problems with chance constraints. The reformulation of chance constrained problems using the penalties was applied in insurance, water-management and civil engineering, cf. [10, 11, 25]. We will draw our attention to the nonconvex case with a finite set of feasible solutions, which can appear in bounded integer programming, and with an infinite bounded set. We will extend the result on the rates of convergence for the sample approximations of the chance constrained problems and summarize the results for the problems with the expectation in the objective which cover the penalty function problems. The approach will be applied to the financial optimization problem with Value at Risk constraint, transaction costs and integer allocations. We compare the ability to generate a feasible solution of the original chance constrained problem using the sample approximations of the chance constraints directly or via sample approximation of the penalty function objective. Another possibility how to use penalty function for getting highly reliable solutions is to employ generalized integrated chance constraints where the penalized constraints are used as constraints, see [5].

The paper is organized as follows. In section 2, we formulate the multiple jointly chance constrained problem and the problem with penalty type objective and we show that they are asymptotically equivalent. In section 3, Monte Carlo techniques for solving the problems are discussed. Numerical study is included in section 4. In section 5, we will summarize our results.

2. REFORMULATION

Let $g_{ji}(x,\omega), i = 0,\dots,k_j, j = 1,\dots,m$, be real functions on $\mathbb{R}^n \times \mathbb{R}^{n'}$ measurable in ω for all $x \in X$. Then the multiple chance constrained problem can be formulated as follows:

$$\begin{array}{rlcl} \psi_\epsilon = & \min_{x\in X} f(x), & & \\ \text{s.t.} & & & \\ & P\big(g_{11}(x,\omega) \le 0,\dots,g_{1k_1}(x,\omega) \le 0\big) & \ge & 1-\varepsilon_1, \\ & & \vdots & \\ & P\big(g_{m1}(x,\omega) \le 0,\dots,g_{mk_m}(x,\omega) \le 0\big) & \ge & 1-\varepsilon_m, \end{array} \tag{2}$$

with an optimal solution x_ϵ, where $\epsilon = (\varepsilon_1,\dots,\varepsilon_m)$, with the levels $\varepsilon_j \in (0,1)$. The formulation covers the joint ($k_1 > 1$ and $m = 1$) as well as the separate ($k_j = 1$ and $m > 1$) chance constrained problems as special cases.

In [11], asymptotic equivalence between the problem with one joint chance constraint and the problem with simple penalty function is shown. The approach by [11] can be extended to a whole class of penalty functions with desirable properties which was done in [6]. We propose further extension to the multiple jointly chance constrained problems (2).

Below, we will consider the penalty functions $\vartheta_j : \mathbb{R}^{k_j} \to \mathbb{R}_+,\ j = 1,\dots,m$, which are continuous nondecreasing in their components, equal to 0 on $\mathbb{R}^{k_j}_-$ and positive otherwise. Two special penalty functions are readily available: $\vartheta^{1,o}(u) = \sum_{i=1}^k ([u_i]^+)^o, o > 0$, where $\vartheta^{1,1}(u) = \sum_{i=1}^k [u_i]^+$ was applied in [11], and $\vartheta^2(u) = \max_{1\le i\le k}[u_i]^+$ applied in [10]. Both functions preserve convexity, ϑ^2 is usually used for the joint chance constraints. The ideal (perfect) penalty function is closely connected to the duality in nonlinear programming:

$$\vartheta^3(u) = \sup_{y \geq 0} \sum_{i=1}^{k} y_i u_i,$$

where $y \in \mathbb{R}^k$. For any nonpositive u it holds $\vartheta^3(u) = 0$, and $\vartheta^3(u) = \infty$ otherwise.

We denote

$$p_j(x, \omega) = \vartheta_j(g_{j1}(x, \omega), \ldots, g_{jk_j}(x, \omega)) : \mathbb{R}^n \times \mathbb{R}^{n'} \to \mathbb{R}$$

the penalized constraints. Our choice is appropriate, because it holds

$$P\big(g_{ji}(x, \omega) \leq 0,\ i = 1, \ldots, k_j\big) \geq 1 - \varepsilon_j \iff P\big(p_j(x, \omega) > 0\big) \leq \varepsilon_j. \tag{3}$$

The corresponding penalty function problem can be formulated as follows:

$$\varphi_N = \min_{x \in X} \left[f(x) + N \cdot \sum_{j=1}^{m} \mathbb{E}[p_j(x, \omega)] \right] \tag{4}$$

with N a positive parameter. We denote x_N an optimal solution of (4).

A rigorous proof of the relationship between the optimal values of (2) and those of (4) for a special additive penalty function and one chance constraint was given by [11]. The following main theorem states the asymptotic equivalence of the models in generalized settings.

Theorem 2.1. Consider the two problems (2) and (4) and assume: $X \neq \emptyset$ is compact, $f(x)$ is a continuous function, $\vartheta_j : \mathbb{R}^{k_j} \to \mathbb{R}_+$, $j = 1, \ldots, m$, are continuous functions, nondecreasing in their components, which are equal to 0 on $\mathbb{R}^{k_j}_-$ and positive otherwise, denote

$$p_j(x, \omega) = \vartheta_j(g_{j1}(x, \omega), \ldots, g_{jk_j}(x, \omega)),\ j = 1, \ldots, m,$$

and assume

(i) $g_{ji}(\cdot, \omega),\ i = 1, \ldots, k_j, j = 1, \ldots, m$, are almost surely continuous;

(ii) there exists a nonnegative random variable $C(\omega)$ with $\mathbb{E}[C^{1+\kappa}(\omega)] < \infty$ for some $\kappa > 0$, such that $|p_j(x, \omega)| \leq C(\omega),\ j = 1, \ldots, m$, for all $x \in X$;

(iii) $\mathbb{E}[p_j(x', \omega)] = 0,\ j = 1, \ldots, m$, for some $x' \in X$;

(iv) $P(g_{ji}(x, \omega) = 0) = 0,\ i = 1, \ldots, k_j, j = 1, \ldots, m$, for all $x \in X$.

Denote $\eta = \kappa/(2(1 + \kappa))$, and for arbitrary $N > 0$ and $\epsilon \in (0, 1)^m$ put

$$\begin{aligned}
\varepsilon_j(x) &= P\big(p_j(x, \omega) > 0\big),\quad j = 1, \ldots, m, \\
\alpha_N(x) &= N \cdot \sum_{j=1}^{m} \mathbb{E}[p_j(x, \omega)], \\
\beta_\epsilon(x) &= \varepsilon_{\max}^{-\eta} \sum_{j=1}^{m} \mathbb{E}[p_j(x, \omega)],
\end{aligned}$$

where $\varepsilon_{\max}$ denotes maximum of the vector $\epsilon = (\varepsilon_1, \ldots, \varepsilon_m)$
and $[1/N^{1/\eta}] = (1/N^{1/\eta}, \ldots, 1/N^{1/\eta})$ is the vector of length m.

THEN for any prescribed $\epsilon \in (0,1)^m$ there always exists N large enough so that minimization (4) generates optimal solutions x_N which also satisfy the chance constraints (2) with the given ϵ.

Moreover, bounds on the optimal value ψ_ϵ of (2) based on the optimal value φ_N of (4) and vice versa can be constructed:

$$\begin{array}{rcccl} \varphi_{1/\varepsilon^{\eta}_{\max}(x_N)} - \beta_{\epsilon(x_N)}(x_{\epsilon(x_N)}) & \leq & \psi_{\epsilon(x_N)} & \leq & \varphi_N - \alpha_N(x_N), \\ \psi_{\epsilon(x_N)} + \alpha_N(x_N) & \leq & \varphi_N & \leq & \psi_{[1/N^{1/\eta}]} + \beta_{[1/N^{1/\eta}]}(x_{[1/N^{1/\eta}]}), \end{array}$$

with

$$\lim_{N\to+\infty} \alpha_N(x_N) = \lim_{N\to+\infty} \varepsilon_j(x_N) = \lim_{\varepsilon_{\max}\to 0_+} \beta_\epsilon(x_\epsilon) = 0$$

for any sequences of optimal solutions x_N and x_ϵ.

P r o o f. We denote

$$\delta_N = \sum_{j=1}^{m} \mathbb{E}[p_j(x_N, \omega)]$$

for some sequence x_N of optimal solutions of the problem (4). Our assumptions and general properties of the penalty function method, see [3, Theorem 9.2.2], ensure that for any sequence x_N of optimal solutions $\delta_N \to 0_+$ and also $\alpha_N(x_N) = N\delta_N \to 0$ as $N \to \infty$. Then by Chebyshev inequality

$$\begin{aligned} & P\Big(p_j(x_N,\omega) > 0\Big) = \\ = \;& P\Big(0 < p_j(x_N,\omega) \leq \sqrt{\delta_N}\Big) + P\Big(p_j(x_N,\omega) > \sqrt{\delta_N}\Big) \\ \leq \;& G_j(x_N, \sqrt{\delta_N}) - G_j(x_N, 0) + \frac{1}{\sqrt{\delta_N}} \mathbb{E}[p_j(x_N,\omega)] \\ \leq \;& G_j(x_N, \sqrt{\delta_N}) - G_j(x_N, 0) + \sqrt{\delta_N} \to 0, \text{ as } N \to \infty, \; j = 1, \ldots, m. \end{aligned}$$

Here for a fixed x, $G_j(x, \cdot)$ denotes the distribution function of $p_j(x,\omega)$ defined by

$$G_j(x,y) = P\Big(p_j(x,\omega) \leq y\Big), \; j = 1, \ldots, m.$$

Assumption (iii) implies that for every vector $\epsilon > 0$ (with small components) there exists some $x_\epsilon \in X$ such that

$$P\big(g_{ji}(x_\epsilon, \omega) \leq 0, \; i = 1, \ldots, k_j\big) \geq 1 - \varepsilon_j, \; j = 1, \ldots, m.$$

Then for any $\epsilon > 0$ the following relations hold

$$\begin{aligned}
&\sum_{j=1}^{m} \mathbb{E}[p_j(x_\epsilon,\omega)] = \\
&= \sum_{j=1}^{m} \int_\Omega |p_j(x_\epsilon,\omega)| I_{(p_j(x_\epsilon,\omega)>0)} P(d\omega) \\
&\le \sum_{j=1}^{m} \int_\Omega C(\omega) I_{(p_j(x_\epsilon,\omega)>0)} P(d\omega) \\
&\le \left(\int_\Omega C^{1+\kappa}(\omega) P(d\omega) \right)^{1/(1+\kappa)} \cdot \sum_{j=1}^{m} \left(\int_\Omega I_{(p_j(x_\epsilon,\omega)>0)} P(d\omega) \right)^{\kappa/(1+\kappa)} \\
&\le c \cdot \sum_{j=1}^{m} P\Big(p_j(x_\epsilon,\omega) > 0\Big)^{\kappa/(1+\kappa)} \\
&\le c \cdot m \cdot \varepsilon_{max}^{\kappa/(1+\kappa)},
\end{aligned}$$

where $c := \left(\int_\Omega C^{1+\kappa}(\omega) P(d\omega) \right)^{1/(1+\kappa)}$, which is finite due to the assumption (ii). Accordingly, for $\varepsilon_{max} \to 0_+$

$$0 \le \sum_{j=1}^{m} \mathbb{E}[p_j(x_\epsilon,\omega)] \le c \cdot m \cdot \varepsilon_{max}^{\kappa/(1+\kappa)} \to 0,$$

and also $\beta_\epsilon(x_\epsilon) \to 0$. If we set

$$\varepsilon_j(x_N) = P\Big(p_j(x_N,\omega) > 0\Big),\ j = 1,\dots,m,$$

then the optimal solution x_N of the expected value problem is feasible for the chance constrained program with $\epsilon(x_N) = (\varepsilon_1(x_N),\dots,\varepsilon_m(x_N))$, because the following relations hold

$$\begin{aligned}
&P\big(g_{ji}(x_N,\omega) \le 0,\ i = 1,\dots,k_j\big) \ge 1 - \varepsilon_j(x_N) \\
&\qquad \Longleftrightarrow P\Big(p_j(x_N,\omega) > 0\Big) \le \varepsilon_j(x_N).
\end{aligned}$$

Hence, we get the inequality

$$\begin{aligned}
\varphi_N &= f(x_N) + N \cdot \sum_{j=1}^{m} \mathbb{E}[p_j(x_N,\omega)] \\
&\ge f(x_{\varepsilon(x_N)}) + N \cdot \sum_{j=1}^{m} \mathbb{E}[p_j(x_N,\omega)] \\
&= \psi_{\varepsilon(x_N)} + \alpha_N(x_N).
\end{aligned}$$

Finally,

$$\begin{aligned}\psi_\epsilon &= \left(\psi_\epsilon + \varepsilon_{\max}^{-\eta} \sum_{j=1}^m \mathbb{E}\big[p_j(x_\epsilon, \omega)\big]\right) - \varepsilon_{\max}^{-\eta} \sum_{j=1}^m \mathbb{E}[p_j(x_\varepsilon, \omega)] \\ &\geq \varphi_{\varepsilon_{\max}^{-\eta}} - \varepsilon_{\max}^{-\eta} \sum_{j=1}^m \mathbb{E}\big[p_j(x_\epsilon, \omega)\big] \\ &= \varphi_{\varepsilon_{\max}^{-\eta}} - \beta_\epsilon(x_\epsilon).\end{aligned}$$

This completes the proof. □

Note that the theorem does not make any statement on the convergence of optimal solutions but it relates optimal values for certain values of the levels and the penalty parameter. We will investigate the behaviour of the optimal solutions in the numerical study.

Remark 2.2. The assumption (iv) ensures that the probability function

$$P\big(g_{ji}(x, \omega) \leq 0,\ i = 1, \ldots, k_j\big)$$

is continuous in the decision vector, which can be easy seen if we realize that the only point of discontinuity of the function is $g_{ji}(x, \omega) = 0,\ i = 1, \ldots, k_j$ for any x.

The bounds (5) and the terms $\alpha_N(x)$, $\epsilon(x)$ and $\beta_\varepsilon(x)$ depend on the choice of the penalty function ϑ. Notice, however, that when we want to evaluate one of the bounds in (5), we must be prepared to face some problems. We are able to compute $\alpha_N(x_N)$, $\epsilon(x_N)$, hence the upper bound for the optimal value $\psi_{\epsilon(x_N)}$ of the chance constrained program (2) with probability levels $\epsilon(x_N)$. But we are not able to compute $\beta_{\epsilon(x_N)}(x_{\epsilon(x_N)})$ without having the solution $x_{\epsilon(x_N)}$ which we do not want to find or even may not be able to find.

3. SAMPLE APPROXIMATIONS USING MONTE-CARLO TECHNIQUES

In this part, we will address the rates of convergence for the chance constrained problems and the problems with expectation type objectives which cover the penalty type objectives. Usually, the sample approximation of the chance constrained problems leads only to the feasible solutions of the original problem. Moreover, the sample reformulation results in a large mixed-integer optimization problem, see below. Hence, it may be interesting to investigate the ability to generate the feasible solutions of the original chance constrained problem using the penalty function problems, where no additional integer variables are necessary. Our approach is is summarized in Table 1.

For the case when the set of feasible solutions, the objective function and the constraints are convex, stronger results on the sample approximations are valid, cf. [7]. The results below generalize those of [1, 13, 16] for the case with several chance constraints and they are valid without assuming convexity of any parts of the problems. We will draw our attention to the case when the set of feasible solutions is finite, i. e. $|X| < \infty$, and to the bounded infinite X.

			Program with a random factor	
			↙ ↘	
1.	Stochastic programming formulation	Chance constrained problem (CCP)	$\overset{a.s.}{\Longleftrightarrow}$	Penalty function problem (PFP)
2.	Sample approximation	↓		↓
		SA CCP		SA PFP
3.	Solution validation	↓		↓
		Reliability		Reliability

Tab. 1. Formulation and approximation schema.

In this section, we will refer to the problem (2) as the original problem. We denote the probability functions using the equivalence (3)

$$q_j(x) = P\big(p_j(x,\omega) > 0\big). \tag{5}$$

Then the multiple chance constrained problem (2) can be rewritten as

$$\begin{aligned} \psi_\epsilon \quad = \quad & \min_{x\in X} f(x), \\ \text{s.t.} & \\ & q_1(x) \leq \varepsilon_1, \\ & \vdots \\ & q_m(x) \leq \varepsilon_m. \end{aligned} \tag{6}$$

Let $\omega^1, \dots, \omega^S$ be an independent Monte Carlo sample of the random vector ω. Then, the sample version of the function q_j is defined to be

$$\hat{q}_j^S(x) = S^{-1} \sum_{s=1}^{S} I_{(0,\infty)}\big(p_j(x,\omega^s)\big), \tag{7}$$

where $I_{(\cdot)}$ denotes the indicator function. Finally, the sample version of the multiple jointly chance constrained problem (6) is defined as

$$\begin{aligned} \hat{\psi}_\gamma^S \quad = \quad & \min_{x\in X} f(x), \\ \text{s.t.} & \\ & \hat{q}_1^S(x) \leq \gamma_1, \\ & \vdots \\ & \hat{q}_m^S(x) \leq \gamma_m, \end{aligned} \tag{8}$$

where the levels γ_j are allowed to be different from the original levels ε_j. Let the set X be compact and $g_{ji}(\cdot,\omega^s)$ be continuous for all triplets (i,j,s). The sample approximation of the chance constrained problem can be reformulated as a large mixed-integer nonlinear program

$$\begin{array}{rrcl} & \min_{(x,u)\in X\times\{0,1\}^{mS}} f(x) & & \\ \text{s.t.} & & & \\ & g_{1i}(x,\omega^s) - M(1-u_{1s}) & \leq & 0,\ i=1,\dots,k_1,\ s=1,\dots,S \\ & & \vdots & \\ & g_{mi}(x,\omega^s) - M(1-u_{ms}) & \leq & 0,\ i=1,\dots,k_m,\ s=1,\dots,S, \\ & \frac{1}{S}\sum_{s=1}^S u_{1s} & \geq & 1-\varepsilon_1, \\ & & \vdots & \\ & \frac{1}{S}\sum_{s=1}^S u_{ms} & \geq & 1-\varepsilon_m, \\ & u_{1s},\dots,u_{ms} & \in & \{0,1\},\ s=1,\dots,S, \end{array} \tag{9}$$

where we set $M=\max_{j=1,\dots,m}\max_{i=1,\dots,k_j}\max_{s=1,\dots,S}\sup_{x\in X} g_{ji}(x,\omega^s)$. Due to the increasing number of binary variables u_{ms}, it may be very difficult to solve the problem (9) even using special solvers for the mixed-integer problems.

Finally, note that the main results on the sample average approximation (SAA) techniques for the expectation type stochastic programs with a finite or bounded set of feasible solutions can be found in [23].

3.1. Lower bound for the chance constrained problem

We will assume that it holds $\gamma_j > \varepsilon_j$ for all j, i. e. that the levels of the sample approximated problem are less restrictive. We derive the rate of convergence of the probability that the feasible solution of the original problem is feasible for the sample approximated problem. Hence, the optimal value of the sample approximated problems is lower bound for the optimal value of the original problem with some probability.

For a fixed $\overline{x}\in X$, the probability of the event $p_j(\overline{x},\omega^n)>0$ is $q_j(\overline{x})$. If the $\overline{x}$ is feasible for the original chance constrained problem, we get $q_j(\overline{x})\leq\varepsilon_j$, $j=1,\dots,m$. Using Bonferroni inequality

$$P(\cap_{j=1}^m A_j) \geq 1-\sum_{j=1}^m \big(1-P(A_j)\big)$$

for the events $A_j=\{p_j(\overline{x},\omega)>0\}$ and the inequality based on the Chernoff inequality for the cumulative distribution function of the binomial distribution, see [15, 16]

$$1-P\big(\hat{q}_j^S(\overline{x})\leq\gamma_j\big)\leq\exp\big\{-S(\gamma_j-\varepsilon_j)^2/(2\varepsilon_j)\big\},$$

we obtain

$$\begin{aligned} P\big(\hat{q}_1^S(\overline{x})\leq\gamma_1,\dots,\hat{q}_m^S(\overline{x})\leq\gamma_m\big) &\geq 1-\sum_{j=1}^m \exp\big\{-S(\gamma_j-\varepsilon_j)^2/(2\varepsilon_j)\big\} \\ &\geq 1-m\exp\big\{-S/2\min_{j\in\{1,\dots,m\}}(\gamma_j-\varepsilon_j)^2/\varepsilon_j\big\}. \end{aligned} \tag{10}$$

This means, that we can choose the sample size S to obtain that the feasible solution $\overline{x}$ is also feasible for the sample approximation with a probability at least $1-\delta$, i.e.

$$S \geq \frac{2}{\min_{j\in\{1,\dots,m\}}(\gamma_j-\varepsilon_j)^2/\varepsilon_j} \ln\frac{m}{\delta}, \tag{11}$$

which corresponds to the result of [1] for $m=1$. Previous analysis also implies, that the probability $P(\hat{\psi}_\gamma^S \leq \psi_\epsilon)$ increases exponentially fast with increasing sample size S.

3.2. Feasibility for the chance constrained problem

We derive the rate of convergence of the probability that the set of feasible solutions of the sample approximated problem is contained in the feasibility set of the original problem.

3.2.1. Finite $|X|$

First, we will draw our attention to the case when the set of feasible solutions is finite, i.e. $|X|<\infty$, which appears in the bounded integer programs. We will assume that it holds $\gamma_j<\varepsilon_j$ for all j, i.e. that the levels of the sample approximated problem are more restrictive.

We define the random variable $Y_{sj}=I_{(p_j(x,\omega^s)\leq 0)}$, i.e. $Y_{js}=1$ if $p_j(x,\omega^s)\leq 0$ and 0 otherwise. Let

$$\begin{aligned}
X_{\gamma_j}^S &= \Big\{x\in X:\ \frac{1}{S}\sum_{s=1}^S Y_{js}\geq 1-\gamma_j\Big\},\\
X_{\varepsilon_j} &= \big\{x\in X:\ P(p_j(x,\omega)\leq 0)\geq 1-\varepsilon_j\big\},\\
X_\gamma^S &= \bigcap_{j=1}^m X_{\gamma_j}^S,\\
X_\epsilon &= \bigcap_{j=1}^m X_{\varepsilon_j}.
\end{aligned}$$

Then, for $x\in X\setminus X_{\varepsilon_j}$ we obtain $\mathbb{E}[Y_{js}]=P(p_j(x,\omega)\leq 0)<1-\varepsilon_j$, which we can use to get an estimate for the probability

$$\begin{aligned}
P(x\in X_{\gamma_j}^S) &= P\bigg(\frac{1}{S}\sum_{s=1}^S Y_{js}\geq 1-\gamma_j\bigg)\\
&\leq P\bigg(\sum_{s=1}^S (Y_{js}-\mathbb{E}[Y_{js}])\geq S(\varepsilon_j-\gamma_j)\bigg)\\
&\leq \exp\big\{-2S(\varepsilon_j-\gamma_j)^2\big\},
\end{aligned}\tag{12}$$

where we used Hoeffding's inequality, cf. [8]. We use this estimate to get an upper bound for the probability that there exists a feasible solution of the sample approximated problem which is infeasible for the original problem.

$$\begin{aligned}
1 - P(X_\gamma^S \subseteq X_\epsilon) &= P\big(\exists_{\tilde{j}\in\{1,\dots,m\}} \exists_{x\in X_\gamma^S} :\ P(p_{\tilde{j}}(x,\omega) \le 0) < 1 - \varepsilon_{\tilde{j}}\big) \\
&\le \sum_{j=1}^{m} \sum_{x\in X\setminus X_{\varepsilon_j}} P(x \in X_{\gamma_j}^S) \\
&\le |X \setminus X_\epsilon| \sum_{j=1}^{m} \exp\Big\{ -2S(\varepsilon_j - \gamma_j)^2 \Big\} \\
&\le m|X \setminus X_\epsilon| \exp\Big\{ -2S \min_{j\in\{1,\dots,m\}} (\varepsilon_j - \gamma_j)^2 \Big\}.
\end{aligned}$$

Using previous upper bound it is possible to estimate the sample size S such that the feasible solutions of the sample approximated problems are feasible for the original problem with a high probability $1-\delta$, i. e.

$$S \ge \frac{1}{2\min_{j\in\{1,\dots,m\}}(\gamma_j - \varepsilon_j)^2} \ln \frac{m|X \setminus X_\epsilon|}{\delta}. \tag{13}$$

If we set $m = 1$, we get the same inequality as [13].

3.2.2. Bounded $|X|$

Below we will consider the case when the set of feasible solutions X is bounded but infinite in general. Again, let $\gamma_j < \varepsilon_j$ for all j. However, we will need the following additional assumption which states Lipschitz continuity of the penalized constraints, i. e.

$$|p_j(x,\omega) - p_j(x',\omega)| \le L_j \|x - x'\|,\ \forall\, x, x' \in X,\ \forall\, \omega \in \Omega,\ \forall\, j,$$

for some $L_j > 0$. Let $D = \sup\{\|x - x'\|_\infty :\ x, x' \in X\}$ be the diameter of X. In this case, it is necessary to consider the constraints which are satisfied strictly, i. e. with some deviation τ:

$$\begin{aligned}
X_{\gamma_j,\tau}^S &= \Big\{ x \in X :\ \frac{1}{S}\sum_{s=1}^{S} I_{(p_j(x,\omega^s)+\tau\le 0)} \ge 1 - \gamma_j \Big\} \\
X_{\gamma,\tau}^S &= \bigcap_{j=1}^{m} X_{\gamma_j,\tau}^S.
\end{aligned}$$

According to the proof of [13, Theorem 10], for $\lambda_j \in (0, \varepsilon_j - \gamma_j)$ there exist finite sets $Z_j^\tau \subseteq X$ with

$$|Z_j^\tau| \le \lceil 1/\lambda_j \rceil \, \lceil 2L_j D/\tau \rceil^n$$

where $\lceil\cdot\rceil$ denotes the upper integer part, and for any $x \in X^S_{\gamma,\tau}$ and any j there exists $z \in Z^\tau_j$ such that $\|z - x\|_\infty \le \tau/L_j$. Using the finite sets Z^τ_j we can define

$$\begin{aligned}
Z^{\tau,S}_{\gamma_j} &= \Big\{x \in Z^\tau_j :\ \frac{1}{S}\sum_{s=1}^{S} I_{(p_j(x,\omega^s)\le 0)} \ge 1-\gamma_j\Big\},\\
Z^\tau_{\varepsilon_j-\lambda_j} &= \big\{x \in Z^\tau_j :\ P(p_j(x,\omega)\le 0) \ge 1-\varepsilon_j+\lambda_j\big\},\\
Z^{\tau,S}_{\gamma} &= \bigcap_{j=1}^{m} Z^{\tau,S}_{\gamma_j},\\
Z^\tau_{\epsilon-\lambda} &= \bigcap_{j=1}^{m} Z^\tau_{\varepsilon_j-\lambda_j}.
\end{aligned}$$

Moreover, for all j it holds that $Z^{\tau,S}_{\gamma_j} \subseteq Z^\tau_{\varepsilon_j-\lambda_j}$ implies $X^{\tau,S}_{\gamma_j} \subseteq X_{\epsilon_j}$. For the previous finite sets, the inequality (12) is valid, i.e. we obtain

$$\begin{aligned}
1 - P(Z^{\tau,S}_{\gamma} \subseteq Z^\tau_{\epsilon-\lambda}) \ \le\ & m\left\lceil \frac{1}{\min_{j\in\{1,\dots,m\}}\lambda_j}\right\rceil \left\lceil \frac{2L_{\max}D}{\tau}\right\rceil^n\\
& \exp\Big\{-2S\min_{j\in\{1,\dots,m\}}(\varepsilon_j-\gamma_j-\lambda_j)^2\Big\},
\end{aligned}$$

where $L_{\max} = \max_j L_j$. Since $Z^{\tau,S}_{\gamma} \subseteq Z^\tau_{\epsilon-\lambda}$ implies $X^{\tau,S}_{\gamma} \subseteq X_\epsilon$, we get the inequality for the probabilities

$$P(X^S_{\gamma,\tau} \subseteq X_\epsilon) \ge P(Z^{\tau,S}_{\gamma} \subseteq Z^\tau_{\epsilon-\lambda}).$$

Using the bound it is possible to estimate the sample size S such that the feasible solutions of the sample approximated problems are feasible for the original problem with a high probability $1-\delta$, i.e.

$$\begin{aligned}
S \ \ge\ & \frac{1}{2\min_{j\in\{1,\dots,m\}}(\varepsilon_j-\gamma_j-\lambda_j)^2}\\
& \left(\ln\frac{m}{\delta} + \ln\left\lceil\frac{1}{\min_{j\in\{1,\dots,m\}}\lambda_j}\right\rceil + n\ln\left\lceil\frac{2L_{\max}D}{\tau}\right\rceil\right).
\end{aligned}$$

If we choose $\lambda_j = (\varepsilon_j-\gamma_j)/2$, we obtain

$$\begin{aligned}
S \ \ge\ & \frac{2}{\min_{j\in\{1,\dots,m\}}(\varepsilon_j-\gamma_j)^2}\\
& \left(\ln\frac{m}{\delta} + \ln\left\lceil\frac{2}{\min_{j\in\{1,\dots,m\}}(\varepsilon_j-\gamma_j)}\right\rceil + n\ln\left\lceil\frac{2L_{\max}D}{\tau}\right\rceil\right).
\end{aligned}$$

Setting $m = 1$ we obtain the same estimate as [13].

4. MIXED-INTEGER VAR AND PENALTY FUNCTION PROBLEMS

In this section, we compare the penalty function approach with the chance constrained problems on a mixed-integer portfolio problem of a small investor. We consider 13 most liquid assets which are traded on the main market (SPAD) on Prague Stock Exchange. Weekly returns from the period 6th February 2009 to 10th February 2010 are used to estimate the means and the variance matrix. Suppose that the small investor trades assets on the "mini-SPAD" market. This market enables to trade "mini-lots" (standardized number of assets) with favoured transaction costs.

We denote Q_i the quotation of the "mini-lot" of security i, f_i the fixed transaction costs (not depending on the investment amount), c_i the proportional transaction costs (depending on the investment amount), R_i the random return of the security i, x_i the number of "mini-lots", y_i binary variables which indicate, whether the security i is bought or not. Then, the random loss function depending on our decisions and the random returns has the following form

$$-\sum_{i=1}^{n}(R_i - c_i)Q_i x_i + \sum_{i=1}^{n} f_i y_i.$$

The chance constrained portfolio problem can be formulated as follows

$$\begin{aligned} &\min_{(r,x,y)\in\mathbb{R}\times X} r \\ &P\Big(-\sum_{i=1}^{n}(R_i - c_i)Q_i x_i + \sum_{i=1}^{n} f_i y_i \le r\Big) \ge 1-\varepsilon, \end{aligned} \tag{14}$$

which is in fact minimization of Value at Risk (VaR). Corresponding penalty function problem using the penalty $\vartheta^{1,1}$ is

$$\min_{(r,x,y)\in\mathbb{R}\times X} r + N\cdot\mathbb{E}\Big[-\sum_{i=1}^{n}(R_i - c_i)Q_i x_i + \sum_{i=1}^{n} f_i y_i - r\Big]^{+}. \tag{15}$$

Setting $N = 1/(1-\varepsilon)$ we minimize Conditional Value at Risk (CVaR) exactly, see [21]. Similar problem with CVaR and transaction costs was considered by [2, 4].

The set of feasible solutions contains a budget constraint and the restrictions on the minimal and the maximal number of "mini-lots" which can be bought, i. e.

$$\begin{aligned} X = \ &\big\{x \in \mathbb{N}^n \times \{0,1\}^n \\ &B_l \le \textstyle\sum_{i=1}^{n}(1+c_i)Q_i x_i + \sum_{i=1}^{n} f_i y_i \le B_u, \\ &l_i y_i \le x_i \le u_i y_i,\ i = 1,\ldots,n\big\}, \end{aligned}$$

where B_l and B_u are the lower and the upper bound on the capital available for the portfolio investment, $l_i > 0$ and $u_i > 0$ are the lower and the upper number of units for each security i.

4.1. Estimated sample sizes

In our case, the cardinality of the integer part of the set of feasible solutions is bounded, i. e. $|X| \le 116^{13} \cdot 2^{13}$. Moreover, if the support of the distribution of the returns is bounded, than the free variable t can be restricted to the closed interval which is bounded by the worst loss and by the best profit which can occur for our loss function

ε	γ	δ	S
0.1	0.2	0.01	93
0.05	0.1	0.01	185
0.01	0.02	0.01	9 211
0.1	0.2	0.001	139
0.05	0.1	0.001	277
0.01	0.02	0.001	13 816

Tab. 2. Sample sizes – lower bound.

ε	γ	δ	S
0.1	0.05	0.01	143 142
0.05	0.025	0.01	574 783
0.01	0.005	0.01	1 468 253 656
0.1	0.05	0.001	144 984
0.05	0.025	0.001	582 152
0.01	0.005	0.001	1 486 674 337

Tab. 3. Sample sizes – feasibility.

considering the restrictions. Then we get the following estimate for the sample size which is necessary to generate a lower bound for the optimal value

$$S \geq \frac{2\varepsilon}{(\gamma-\varepsilon)^2} \ln \frac{1}{\delta},$$

and to generate a feasible solution

$$S \geq \frac{2}{(\varepsilon-\gamma)^2} \left(\ln \frac{1}{\delta} + 26 \ln 116 + 26 \ln 2 + \ln \left\lceil \frac{2}{(\varepsilon-\gamma)} \right\rceil + \ln \left\lceil \frac{2D}{\tau} \right\rceil \right),$$

which is based on the decomposition of the set of feasible solutions into the integer and real bounded part. In Tables 2 and 3, there are examples of the sample sizes for different combinations of the parameters $\gamma, \varepsilon, \delta$ where we have chosen $\tau = 10^{-6}$ and $D = 2 \cdot 10^6$ which is the difference between the worst loss and the best profit. The sample size which is necessary to generate the lower bound for the optimal value of the original problem is quite low and will be covered partly by the following numerical experiment, see Table 2. However, the samples, which are necessary to ensure that the set of feasible solutions of the sample approximated problem is contained in the feasibility set of the original problem, are quite large and rapidly increase with decreasing level ε, see Table 3.

4.2. Numerical comparison

We generated 100 samples for each sample size S, i. e. $100 \times S$ realizations, from the truncated normal distribution where the truncation points were set to -1 for all random returns. We used the modelling system GAMS and the solver CPLEX to solve the sample approximations of the chance constrained problems (14) and the penalty function problems (15) for different sample sizes S, levels γ and penalty parameters N. Basic descriptive statistics for the optimal values are contained in Tables 4, 5. Minimal and mean reliabilities of the obtained solutions and basic descriptive statistics for the optimal values can be found in Tables 4 and 5. As we can see from Table 5, the "Penalty term"

$$N \cdot \mathbb{E}\Big[- \sum_{i=1}^{n} (R_i - c_i) Q_i x_i + \sum_{i=1}^{n} f_i y_i - r \Big]^{+}$$

really decreases with increasing penalty parameter N and reduces violations of the constraint $(R_i - c_i)Q_i x_i + \sum_{i=1}^{n} f_i y_i - r \leq 0$ for each sample size.

		Reliabilities		Optimal values	
S	γ	min	mean	mean	st.dev
100	0.1	0.8844	0.9592	41784.66	7525.69
100	0.05	0.9054	0.9516	41821.60	7465.46
100	0.01	0.8939	0.9456	42312.34	7612.11
250	0.1	0.9546	0.9824	52429.77	9887.54
250	0.05	0.9545	0.9820	52431.23	9884.16
250	0.01	0.9555	0.9807	52626.23	9909.60
500	0.1	0.9744	0.9903	67824.32	15849.91
500	0.05	0.9744	0.9903	67824.32	15849.91
500	0.01	0.9726	0.9906	67942.02	15757.14
750	0.1	0.9849	0.9952	74655.08	19435.11
750	0.05	0.9849	0.9952	74652.82	19436.71
750	0.01	0.9866	0.9953	74679.40	19187.28
1000	0.1	0.9870	0.9966	93390.26	28293.28
1000	0.05	0.9870	0.9966	93414.25	28269.13
1000	0.01	0.9870	0.9966	93384.85	28264.63

Tab. 4. Chance constrained problems – optimal values.

To verify the reliability of the obtained optimal solutions, we used the independent samples of 10 000 realizations from the truncated normal distribution which was used to model the random returns. The columns "Reliability" contain relative number of realizations for which the chance constraint is fulfilled. As can be easy seen, the reliability of the obtained solutions increases with increasing levels γ and penalty parameters N for each sample size S. Both problems are also able to generate comparable solutions for the same sample sizes, see Tables 4 and 5. Furthermore, we can compare the descriptive

		Reliabilities		Optimal values		Penalty term	
S	N	min	mean	mean	st.dev	mean	st.dev
100	1	0.7622	0.8770	33403.52	4311.27	7731.06	2530.95
100	10	0.8967	0.9581	42830.41	7489.58	3.09	30.93
100	100	0.8967	0.9581	42902.79	7484.36	0.00	0.00
100	1000	0.8967	0.9581	42903.93	7474.20	0.00	0.00
250	1	0.8330	0.8888	37382.48	6017.16	9673.10	2115.72
250	10	0.9495	0.9788	52156.49	9360.82	2570.37	2255.62
250	100	0.9571	0.9841	53493.47	9862.21	0.00	0.00
250	1000	0.9571	0.9840	53458.34	9898.87	0.00	0.00
500	0	0.5408	0.5408	4202.39	596.15	4202.39	596.15
500	1	0.8716	0.9016	43537.39	8424.45	11865.88	3520.77
500	10	0.9723	0.9871	63886.92	13472.75	5109.98	2719.34
500	100	0.9813	0.9935	68995.38	15851.31	0.00	0.00
500	1000	0.9813	0.9934	68914.67	15748.83	0.00	0.00
750	1	0.8697	0.8990	44922.49	9914.34	13191.21	5100.68
750	10	0.9785	0.9878	68251.45	17167.97	7328.08	3405.62
750	100	0.9890	0.9957	75669.31	19379.62	0.00	0.00
750	1000	0.9890	0.9956	75541.31	19234.11	0.00	0.00
1000	1	0.8739	0.8976	51840.01	12169.11	16647.91	7013.17
1000	10	0.9753	0.9886	82550.78	23493.47	9591.71	5108.62
1000	100	0.9900	0.9966	94331.08	27977.78	0.00	0.00
1000	1000	0.9900	0.9966	94357.45	28209.17	0.00	0.00

Tab. 5. Penalty function problems – reliabilities.

statistics of the optimal values $\hat{\psi}^S_\gamma$, $\hat{\varphi}^S_N$. We observe that the variability of the values increases with the sample size. Thus, we pay for the increasing reliability of the optimal solutions by decreasing reliability of the optimal values when we increase the size of the sample. Finally, we can compare the used sample sizes with theoretically estimated sizes in Tables 2 and 3, which can be now seen as very conservative.

5. CONCLUSION

Reformulation of chance constrained programs by incorporating a suitably chosen penalty function into the objective helps to arrive at problems with expectation in the objective and a fixed set of feasible solutions. The obtained problems are much simpler to solve and analyze then the chance constrained programs, cf. [12]. The recommended form of the penalty function follows the basic ideas of penalty methods and its suitable properties follow by generalization of the results from [6, 11].

The numerical study shows that not only the sample approximated chance constrained problems but also the penalty function problems are able to generate solutions which are feasible for the original chance constrained problem with a high reliability.

ACKNOWLEDGEMENT

This work was supported by the Czech Science Foundation under the grant [P402/12/G097].

(Received June 27, 2011)

REFERENCES

[1] S. Ahmed and A. Shapiro: Solving chance-constrained stochastic programs via sampling and integer programming. In: Tutorials in Operations Research, (Z.-L. Chen and S. Raghavan, eds.), INFORMS 2008.

[2] E. Angelelli, R. Mansini, and M. G. Speranza: A comparison of MAD and CVaR models with real features. J. Banking Finance *32* (2008), 1188–1197.

[3] M. S. Bazara, H. D. Sherali, and C. M. Shetty: Nonlinear Programming: Theory and Algorithms. Wiley, Singapore 1993.

[4] M. Branda: Solving real-life portfolio problem using stochastic programming and Monte-Carlo techniques. In: Proc. Mathematical Methods in Economics 2010, (M. Houda, J. Friebelová, eds.), University of South Bohemia, České Budějovice 2010.

[5] M. Branda: Stochastic programming problems with generalized integrated chance constraints. Accepted to Optimization 2011.

[6] M. Branda and J. Dupačová: Approximations and contamination bounds for probabilistic programs. Accepted to Ann. Oper. Res. 2011 (Online first). See also SPEPS *13*, 2008.

[7] G. Calafiore and M. C. Campi: Uncertain convex programs: randomized solutions and confidence levels. Math. Programming, Ser. A *102* (2008), 25–46.

[8] A. DasGupta: Asymptotic Theory of Statistics and Probability. Springer, New York 1993.

[9] J. Dupačová and M. Kopa: Robustness in stochastic programs with risk constraints. Accepted to Ann. Oper. Res. 2011 (Online first).

[10] J. Dupačová, A. Gaivoronski, Z. Kos, and T. Szantai: Stochastic programming in water management: A case study and a comparison of solution techniques. Europ. J. Oper. Res. *52* (1991), 28–44.

[11] Y. M. Ermoliev, T. Y. Ermolieva, G. J. Macdonald, and V. I. Norkin: Stochastic optimization of insurance portfolios for managing exposure to catastrophic risks. Ann. Oper. Res. *99* (2000), 207–225.

[12] P. Lachout: Approximative solutions of stochastic optimization problems. Kybernetika *46* (2010), 3, 513-523.

[13] J. Luedtke and S. Ahmed: A sample approximation approach for optimization with probabilistic constraints. SIAM J. Optim. *19* (2008), 674–699.

[14] J. Nocedal and S. J. Wright: Numerical Optimization. Springer, New York 2000.

[15] B. Pagnoncelli, S. Ahmed, and A. Shapiro: Computational study of a chance constrained portfolio selection problem. Optimization Online 2008.

[16] B. Pagnoncelli, S. Ahmed, and A. Shapiro: Sample average approximation method for chance constrained programming: Theory and applications. J. Optim. Theory Appl. *142* (2009), 399–416.

[17] A. Prékopa: Contributions to the theory of stochastic programming. Math. Programming *4* (1973), 202–221.

[18] A. Prékopa: Dual method for the solution of a one-stage stochastic programming problem with random RHS obeying a discrete probability distribution. Math. Methods Oper. Res. *34* (1990), 441–461.

[19] A. Prékopa: Stochastic Programming. Kluwer, Dordrecht and Académiai Kiadó, Budapest 1995.

[20] A. Prékopa: Probabilistic programming. In: Stochastic Programming, (A. Ruszczynski and A. Shapiro,eds.), Handbook in Operations Research and Management Science, Vol. 10, Elsevier, Amsterdam 2003, pp. 267–352.

[21] R. T. Rockafellar and S. Uryasev: Conditional value-at-risk for general loss distributions. J. Banking Finance *26* (2002), 1443–1471.

[22] R. T. Rockafellar and R. Wets: Variational Analysis. Springer-Verlag, Berlin 2004.

[23] A. Shapiro: Monte Carlo sampling methods. In: Stochastic Programming, (A. Ruszczynski and A. Shapiro, eds.), Handbook in Operations Research and Management Science, Vol. 10, Elsevier, Amsterdam 2003, pp. 353–426.

[24] S. W. Wallace and W. T. Ziemba: Applications of stochastic programming. MPS-SIAM Book Series on Optimization *5* (2005), Society for Industrial and Applied Mathematics.

[25] E. Žampachová and M. Mrázek: Stochastic optimization in beam design and its reliability check. In: MENDEL 2010 – 16th Internat. Conference on Soft Computing, (R. Matoušek), ed.), Mendel Journal series, FME BUT, Brno 2010, pp. 405–410.

[26] E. Žampachová, P. Popela, and M. Mrázek: Optimum beam design via stochastic programming. Kybernetika *46* (2010), 3, 571–582.

Martin Branda, Charles University in Prague, Faculty of Mathematics and Physics, Department of Probability and Mathematical Statistics, Sokolovská 83, 186 75, Praha 8 and Institute of Information Theory and Automation, Academy of Sciences of the Czech Republic, Pod Vodárenskou věží 4, 182 08 Praha 8. Czech Republic.

e-mail: branda@karlin.mff.cuni.cz