I^2SDS
The Institute for Integrating Statistics in Decision Sciences

Technical Report TR-2009-4
April 20, 2009

On The Relaxation of Probabilistically Constrained Stochastic Programming Problems with Random Right-hand Sides

Miguel Lejeune
Department of Decision Sciences
The George Washington University

A. Prepoka
RUTCOR
Rutgers University

On the Relaxation of Probabilistically Constrained Stochastic Programming Problems with Random Right-Hand Sides

M.A. LEJEUNE[a] A. PRÉKOPA[b]

[a] George Washington University, Washington, DC; mlejeune@gwu.edu
[b] RUTCOR - Rutgers University Center for Operations Research, Piscataway, NJ; prekopa@rutcor.rutgers.edu

We consider probabilistically constrained stochastic programming problems, in which the random variables are located in the right-hand sides of the stochastic constraints. The objective function is linear, and its optimization is subject to a set of linear constraints as well as a joint probabilistic constraint. Problems of that kind arise in many different contexts, and are particularly difficult to solve for random variables with continuous joint distributions, because the calculation of the cumulative distribution function and its gradient values involves numerical integration and/or simulation in higher dimensional spaces. In this paper, we first describe various relaxations to the problem in which the joint probabilistic constraint is replaced by individual or lower dimensional cumulative distribution functions, and analyze the computational tractability of the relaxed problems. We solve the original formulation and the proposed relaxations, study their computational efficiency, and evaluate the tightness and constraining power of the relaxations. Finally, we derive additional models in which the joint probabilistic constraint is approximated by constraints involving conditional expectations, analyze their convexity properties, and study the relationship of the various bounds that come up in the relaxations. The computational analysis is made in connection with applications to reservoir system design, coffee blending and power systems.

Key words: probabilistically constrained problem, reliability level, relaxation, stochastic optimization

1 Introduction

The following program

$$\begin{aligned} \min\ & c^T x \\ s.to\ & Tx \geq \xi \\ & Ax \geq b \\ & x \geq 0, \end{aligned},$$

where A in an $[m \times n]$-dimensional matrix, T in an $[r \times n]$-dimensional matrix, c and x are n-dimensional vectors, b and ξ are m- and r-dimensional vectors, respectively, is frequently the underlying program for static stochastic programming model formulations. One of them has the name programming under probabilistic constraints and is given by

$$\begin{aligned} \min\ & c^T x \\ s.to\ & \mathbb{P}(Tx \geq \xi) \geq p \\ & Ax \geq b \\ & x \geq 0, \end{aligned} \tag{1}$$

where p is a prescribed probability level, usually close to 1. If $F(z) \leq \mathbb{P}(\xi \leq z)$ is the cumulative distribution function of the random vector ξ, then the probabilistic constraint can be written in the equivalent form: $F(Tx) \geq p$.

The probabilistic constraint enforces the joint fulfillment of a system of linear inequalities with random right-hand side variables on or above a prescribed probability level *p*. Problems of that kind arise in many different contexts: reservoir management (Prékopa, Szántai, 1978), power management (Prékopa et al. 1980), chemical (Henrion et al., 2001a) and distillation processes (Henrion et al., 2001b), monitoring of pollution level (Gren, 2008), supply chain management (Lejeune, Ruszczyński, 2007, Lejeune, 2008), military operations (Kress et al., 2008), etc. We refer the reader to Prékopa (1995, 2003), Sen (1996) and Henrion (2004) for a more detailed review of application fields.

Facing prediction uncertainty and having to design large scale planning strategies, it is very difficult for human decision makers to develop effective, coordinated solutions in real time. A possible consequence of this is the implementation of very conservative decisions leading to unnecessary delays and costs. In that respect, probabilistically constrained problems are very helpful to design ahead-of-time efficient strategies in the presence of uncertainty. Such problems are extremely difficult to solve especially if the random variables have continuous joint distributions, because the calculation of the cumulative distribution function and of its gradient values involves numerical integration and/or simulation in higher dimensional spaces. The development of computationally tractable solution methods is therefore very important.

Programming under probabilistic constraints, also called *chance constrained programming*, was introduced by Charnes et al. (1958), who considered a set (i $=1,\ldots,r$) of individual probabilistic constraints imposed on each particular stochastic inequality:

$$\begin{aligned} &\min c^T x \\ &s.to\ \mathbb{P}(T_i x \geq \xi_i) \geq p_i, i = 1,...,r \\ &\quad Ax \geq b \\ &\quad x \geq 0. \end{aligned}$$

The use of individual probabilistic constraints $i = 1,...,r$ is relatively easy to handle, but the robustness of such formulations is questionable: individual probabilistic constraints are appropriate if the system is composed of components that do not affect each other. However, in most situations, probabilistic constraints, taken individually, do not give an accurate representation of the considered system.

Miller and Wagner (1965) proposed a formulation for joint probabilistic constraints where the random variables $\xi_i, i = 1,...,r$ are assumed to be independent. Their problem is the following:

$$\begin{aligned} &\min c^T x \\ &s.to\ \prod_{i=1}^{r} \mathbb{P}(T_i x \geq \xi_i) \geq p \\ &\quad Ax \geq b \\ &\quad x \geq 0. \end{aligned} \tag{2}$$

The probabilistic constraint has the equivalent form: $\prod_{i=1}^{r} F_i(T_i x) \geq p$, where F_i is the cumulative distribution function of $\xi_i, i = 1,...,r$.

The general case was introduced by Prékopa (1970, 1973, 1995). The problem in its most general form can be stated as follows:

$$\begin{aligned} &\min c^T x \\ &s.to\ \mathbb{P}(\mathrm{g}_i(x,\xi) \geq 0, i = 1,...,r) \geq p \\ &\quad Ax \geq b \\ &\quad x \geq 0, \end{aligned} \tag{3}$$

where $\mathrm{g}_i(x,y) \geq 0, i = 1,...,r$ are functions of *(x,y)* and ξ is a random vector with components that are not necessarily independent. In the special case where $g_i(x,y) = T_i x - y_i, i = 1,...,r$, we obtain problem (1) that however does not allow the transformation into (2). In fact, the stochastic independence of $\xi_i, i = 1,...,r$ is not assumed.

The first question to settle in connection with problem (3) is: under what conditions is the set of feasible solutions convex? A general theorem that answers this question is due to Prékopa (1971, 1973, 1995) and is stated below.

Theorem 1: If $g_i(x, y) \geq 0,\ i = 1,...,r$ are concave functions of $x \in R^n, y \in R^q$ and $\xi \in R^q$ is a continuously distributed random vector with logarithmically concave (logconcave) probability density function, then

$$\mathbb{P}(g_i(x,\xi) \geq 0, i = 1,...,r) \geq p$$

is a logarithmically concave function of x.

Theorem 1 implies that if ξ has continuous distribution with logarithmically concave probability density function, then the set of feasible solutions of problem (3) is convex. Note that a logarithmically concave function is also quasi-concave. For more convexity theorems concerning probabilistically constrained stochastic programming problems, the reader is referred to Prékopa (1995).

The collection of logconcave probability density functions includes the multivariate non-degenerate normal distribution, the uniform (over a convex set) distribution, and, under some conditions on the defining parameters, the multivariate Dirichlet, Wishart and gamma distributions. Prékopa (2001) has proved that the standard r-variate normal cumulative distribution function $\Phi(z_1,...,z_n; R)$, where R is the correlation matrix, is concave on the set $\left\{ z \middle| z_i \geq \sqrt{r-1}, i = 1,...,r \right\}$. There is no general result for discrete distributions that would parallel Prékopa's results (1973), except for the univariate case Prékopa (1995).

In the next section, we focus on the probabilistically constrained problem (3), present some of its alternative formulations, and describe several bounding schemes. The introduction of the bounding schemes into the stochastic problem results in the computation of the joint probability distribution function values of several lower dimensional random vectors instead of that of a single higher dimensional random vector. In Section 3, we characterize the proposed relaxations for stochastic problems with joint probabilistic constraints, solve the associated optimization models for several problem instances taken from the stochastic programming literature, and assess the computational tractability and the constraining power of the respective relaxations. In Section 4, we analyze the convexity of non-linear constraints enforcing a certain type of reliability level, study the correspondence between conditional expectation constraints and probabilistic constraints, and carry out computational experiments. A real-life power management problem illustrates the usefulness of the derived results and the substantial gain in reliability that can be obtained from an appropriate formulation, and solution, of the corresponding probabilistic problem. Section 5 contains concluding remarks.

2 Formulation

In this section, we introduce alternative formulations for the probabilistic program (3). The first one, relying on the *p*-efficiency concept, is applicable to multivariate random variables having discrete probability distributions (Prékopa, 1990,), while the other formulations can be applied to random variables regardless of the type (discrete or continuous) of probability distributions.

2.1 *p*-efficiency approximation for discretely distributed random variables

In the case of discretely distributed random variables, problem (3) can be transformed into a disjunctive program using the concept of *p*-efficiency (Prékopa, 1990). Let $p \in (0,1)$, and F be the probability distribution function of an *r*-dimensional integer-valued random variable $\xi \in \mathbb{Z}_+^r$. A point *z*, $z \in R^r$, is called a *p*-efficient point of the probability distribution function *F* if:

$$\begin{cases} F(z) \geq p, \\ \text{there is no } z' \leq z, z' \neq z \text{ such that } F(z') \geq p. \end{cases}$$

It has been shown (Dentcheva et al., 2000) that for any probability distribution on $\mathbb{Z}_+^r$, the set of *p*-efficient points is non-empty and finite. Consequently, problem (3) can be substituted by the following disjunctive program:

$$\begin{aligned} &\min c^T x \\ &s.to \quad Tx \geq z \\ &\qquad\quad Ax \geq b \\ &\qquad\quad z \in \mathrm{K}^p \\ &\qquad\quad x \geq 0, \end{aligned} \tag{4}$$

where $\{z^{(1)},\ldots,z^{(J)}\}$ is the set of *p*-efficient points and $\mathrm{K}^p = \bigcup_{i=1}^{J} (z^{(i)} + R_+^s)$. The disjunctive constraint means that $Tx \geq z^{(i)}$ must hold for at least one *i*: it replaces the joint probabilistic constraint

$\mathbb{P}(T_i x \geq \xi_i, i = 1,...,r) \geq p$. The p-efficient points are multi-dimensional vectors that may be found prior to the optimization of (4) with respect to x, or may be enumerated in the course of the optimization procedure.

For small dimensional problems, the easiest way is to enumerate all p-efficient points $z^{(i)}$ (Prékopa et al., 1998, Beraldi, Ruszczyński, 2002, Boros et al., 2003) and to process the associated problems. However, for larger problems, the brute force approach that consists of enumerating all p-efficient points can be overly time- and resource-consuming; it is then better suited to convexify the problem (4) as follows:

$$\begin{aligned} \min \quad & c^T x \\ \text{s.to} \quad & Ax \geq b \\ & Tx \geq \sum_{i=1}^{J} \lambda_i z^{(i)} \\ & \sum_{i=1}^{J} \lambda_i = 1 \\ & \lambda_1 \geq 0, i = 1,...,I \\ & x \geq 0, \end{aligned} \tag{5}$$

Problem (5) imposes Tx to be greater than or equal to a convex combination of p-efficient points; its optimal value is a lower bound (for a minimization problem) on the optimal value of (4). Since the enumeration of all p-efficient points is difficult, the solution of (5) is sometimes carried out using a column generation method. In some cases, that method (Prékopa et al., 1998, Dentcheva et al., 2000) has been shown to be very efficient.

2.2 Bounding scheme for multivariate random variables

The approximation schemes presented in this section are applicable regardless of the type (i.e., discrete or continuous) of the probability distribution of the random variables. In case of continuous random variables, the computation of the joint probability distribution function requires multivariate integration that is computationally intensive, in general. (Szántai, 1988, Deák, 1988, Genz, 1992, Genz, Bretz, 2002).

2.2.1 Boole's lower bound on the intersection of events

An alternative formulation for (1) can be obtained using Boole's lower bound relative to the intersection of r events $A_1, \ldots, A_r$:

$$\mathbb{P}(A_i \cap \ldots \cap A_r) \geq S_1 - (r-1). \tag{6}$$

Using (6) for the events $A_i : \xi_i \leq z_i, i = 1, \ldots, r$, and imposing probabilistic constraint on the right-hand side, we obtain, from (1), the following problem:

$$\begin{aligned} &\min c^T x \\ &s.to \ \sum_{i=1}^{r} F_i(T_i x) - (r-1) \geq p \\ &\quad Ax \geq b \\ &\quad x \geq 0. \end{aligned} \tag{7}$$

The nonlinear constraint in (7) can be rewritten as: $1 - p \geq \sum_{i=1}^{r} \left(1 - F_i(T_i x)\right)$. The introduction of the auxiliary positive decision variables $p_i, \ldots, p_r$ such that $1 - p \geq \sum_{i=1}^{r} (1 - p_i)$ and of the support constraints $F_i\left(T_i x\right) \geq p_i, i = 1, \ldots, r$ allows for the reformulation of (7) as an optimization problem containing a finite number r of individual probabilistic constraints (Prékopa, 1995, 1999), whose deterministic equivalent is:

$$\begin{aligned} &\min c^T x \\ &s.to \ \ T_i x \geq F_i^{-1}(p_i), i = 1, \ldots, r \\ &\quad 1 - p \geq \sum_{i=1}^{r} (1 - p_i) \\ &\quad Ax \geq b \\ &\quad 0 \leq p_i \leq 1, i = 1, \ldots, r \\ &\quad x \geq 0, \end{aligned} \tag{8}$$

where $p_i, \ldots, p_r$ are also decision variables, and p is a fixed probability level. Thus, if x is a feasible solution for (8), it is also a feasible solution for problem (1).

2.2.2 Binomial moments bounding scheme

The formulation presented in this section is taken from (Prékopa, 1999), and it is based on the binomial moment problem of the same author formulated for a finite number of events $A_1, \ldots, A_r$, defined in a specified probability space. We use sharp lower and upper bounds for the probability of the following Boolean functions of events: $A_1 \cap \ldots \cap A_r$, $A_1 \cup \ldots \cup A_r$. Denoting by ξ the number of events that occur out of $A_1, \ldots, A_r$, it is well known that

$$E\left[\binom{\xi}{k}\right] = S_k, k = 1, \ldots, r \qquad \text{where} \qquad S_k = \sum_{1 \leq i_1 < \ldots < i_k \leq r} P\left(A_{i_1}, \ldots, A_{i_r}\right).$$

In view of the above equation, S_k is called the k^{th} binomial moment of the random variable ξ.

The binomial moment problem is defined as the following linear programming problem:

$$\begin{aligned}
&\min(\max) && \sum_{i=0}^{r} f_i v_i \\
&s.\ to && \sum_{i=0}^{r} \binom{i}{k} v_i = S_k, k = 0,\dots,m \qquad (9)\\
& && v_i \geq 0, i = 0,\dots,r,
\end{aligned}$$

where $f_0,\dots,f_r$ are some constants and $m < r$. If

$$f_i = \begin{cases} 0, & \text{if } i = 0, \\ 1, & \text{if } i > 0, \end{cases} \qquad (10)$$

then the optimal values of (9) provide lower (upper) bounds for $\mathbb{P}(A_1 \cup \dots \cup A_r)$. If

$$f_i = \begin{cases} 0, & \text{if } i < r, \\ 1, & \text{if } i = r, \end{cases} \qquad (11)$$

then the optimal values of (9) provide lower (upper) bounds for $\mathbb{P}(A_1 \cap \dots \cap A_r)$. Since

$$\mathbb{P}\left(A_1 \cap \dots \cap A_r\right) = 1 - \mathbb{P}\left(A_1^C \cup \dots \cup A_r^C\right),$$

with A_i^C being the complementary event of A_i, the sharp lower (upper) bound for $\mathbb{P}(A_1 \cap \dots \cap A_r)$ is the same as 1 - the sharp upper (lower) bound for $\mathbb{P}(A_1^C \cup \dots \cup A_r^C)$.

Let $F_{i_1,\dots,i_k}$ designate the joint probability distribution function of the random variables $\xi_{i_1},\dots,\xi_{i_k}$. Using the results above, we formulate the following linear programs

$$\begin{aligned}
&\min(\max) && v_r \\
&s.to && v_0 + v_1 + v_2 + \dots + v_r = 1 \\
& && v_1 + 2v_2 + 3v_3 + \dots + rv_r = \sum_{i=1}^{r} F_i\left(T_i x\right) \\
& && v_2 + \binom{3}{2} v_3 + \dots + \binom{r}{2} v_r = \sum_{1 \leq i_1 < i_2 \leq r} F_{i_1,i_2}\left(T_{i_1} x, T_{i_2} x\right) \\
& && \dots \\
& && v_m + \binom{m+1}{m} v_{m+1} + \dots + \binom{r}{m} v_r = \sum_{1 \leq i_1 < \dots < i_m \leq r} F_{i_1,\dots,i_r}\left(T_{i_1} x, \dots, T_{i_m} x\right) \\
& && v_0, v_1, \dots, v_r \geq 0,
\end{aligned}$$

where the optimal values provide lower and upper bounds for $F\left(T_i x, \dots, T_r x\right)$.

Using this bounding scheme and introducing it in problem (3) for replacing the joint probabilistic constraints, we obtain the following approximation problems for the minimization problem in (1):

$$
\begin{aligned}
\min \quad & c^T x + \alpha v_r \\
s.to \quad & v_0 + v_1 + v_2 + \ldots + v_r = 1 \\
& v_1 + 2v_2 + 3v_3 + \ldots + r v_r = \sum_{i=1}^{r} F_i(T_i x) \\
& v_2 + \binom{3}{2} v_3 + \ldots + \binom{r}{2} v_r = \sum_{1 \le i_1 < i_2 \le r} F_{i_1, i_2}(T_{i_1} x, T_{i_2} x) \\
& \ldots \\
& v_m + \binom{m+1}{m} v_{m+1} + \ldots + \binom{r}{m} v_r = \sum_{1 \le i_1 < \ldots < i_m \le r} F_{i_1, \ldots, i_r}(T_{i_1} x, \ldots, T_{i_m} x) \\
& v_0, v_1, \ldots, v_r \ge 0 \\
& v_r \ge p \\
& x \ge 0,
\end{aligned}
\tag{12}
$$

and

$$
\begin{aligned}
\min \quad & c^T x - \alpha v_r \\
s.to \quad & v_0 + v_1 + v_2 + \ldots + v_r = 1 \\
& v_1 + 2v_2 + 3v_3 + \ldots + r v_r = \sum_{i=1}^{r} F_i(T_i x) \\
& v_2 + \binom{3}{2} v_3 + \ldots + \binom{r}{2} v_r = \sum_{1 \le i_1 < i_2 \le r} F_{i_1, i_2}(T_{i_1} x, T_{i_2} x) \\
& \ldots \\
& v_m + \binom{m+1}{m} v_{m+1} + \ldots + \binom{r}{m} v_r = \sum_{1 \le i_1 < \ldots < i_m \le r} F_{i_1, \ldots, i_r}(T_{i_1} x, \ldots, T_{i_m} x) \\
& v_0, v_1, \ldots, v_r \ge 0 \\
& v_r \ge p \\
& x \ge 0,
\end{aligned}
\tag{13}
$$

where α is an arbitrary positive number.

2.2.3 Bounding scheme using Slepian's inequality

Slepian's inequality (Slepian, 1962, Bawa, 1973) can be stated as follows. If R and R' are two correlation matrices such that $R \ge R'$, then, for any $z = (z_1, \ldots, z_r) \in R^r$, we have

$$\Phi(z; R) \ge \Phi(z; R'), \tag{14}$$

where Φ is the r-variate standard normal probability distribution function. If $(\xi_1, \ldots, \xi_r)$ is a non-degenerate, normally distributed random vector with $E[\xi_i] = \mu_i, Var[\xi_i] = \sigma_i^2, i = 1, \ldots, r$ and correlation matrix R, then the joint probability distribution function of $\xi_1, \ldots, \xi_r$ is

$$\Phi\left(\frac{z_i - \mu_i}{\sigma_i}, i = 1, \ldots, r; R \right). \tag{15}$$

If $R \ge R'$, then we have the inequality

$$\Phi\left(\frac{z_i - \mu_i}{\sigma_i}, i = 1, \ldots, r; R \right) \ge \Phi\left(\frac{z_i - \mu_i}{\sigma_i}, i = 1, \ldots, r; R' \right). \tag{16}$$

If we impose a probabilistic constraint on the right-hand side expression in (17), rather than on the left-hand side expression, and substitute it for the probabilistic constraint in (1), then the set of feasible solution becomes smaller and the optimum value larger. Hence, the new optimum value provides us with a lower bound on the optimum value of problem (1).

Denoting by $z^{'*} = c^T x^{'*}$ ($z^* = c^T x^*$) the optimal value of (1) associated with *R' (R)*, and by $x_i^{'*}$ (x_i^*) the optimal values of the decision variables, the following valid inequality

$$z^* \le z'^* \tag{18}$$

and disjunctive cuts

$$\begin{gathered} x_i^* \le x_i^{'*} + \delta_i M, i = 1,\dots,r \\ \sum_{i=1}^{r} \delta_i \le r-1, \delta_i \in \{0,1\}, i = 1,\dots,r, \end{gathered} \tag{19}$$

where M is a large positive number, and δ is an r-dimensional binary vector, can be introduced in (1). In particular, if $R \ge 0$ and $R' = I$, then (20) becomes:

$$\Phi\left(\frac{T_i x - \mu_i}{\sigma_i}, i = 1,\dots,r; R\right) \ge \prod_{i=1}^{r} \Phi\left(\frac{T_i x - \mu_i}{\sigma_i}\right),$$

which is equivalent to

$$\ln\Phi\left(\frac{T_i x - \mu_i}{\sigma_i}, i = 1,\dots,r; R\right) \ge \sum_{i=1}^{r} \ln\Phi\left(\frac{T_i x - \mu_i}{\sigma_i}\right). \tag{21}$$

Note that in (22) the functions on both sides are concave, by Theorem 1. Slepian's inequality is especially useful if the correlations between the different random variables are all ≥ 0 or ≤ 0, because the use of independent random variables can then provide bounds. In the next section, we further evaluate the respective computational tractability and constraining power of the formulations presented in this section.

3 Evaluation of Approximation Approaches

3.1 Complexity and Solution Method

In this section, we briefly describe the numerical solution method used to solve the probabilistic problem (1) and some possible relaxations (8), (12) and (13). The solution of the stochastic problem (1) containing joint probabilistic constraints with dependent random variables is very challenging; the complexity of the task rapidly increases with the dimension of the random variable. With respect to the computational tractability of the approximation problems for (1), we first observe that the convexity of problem (8) can be proved under mild conditions. In fact, if $F_i(z_i)$ is concave for $z_i \geq z_{i0}$, then $F_i^{-1}(p_i)$ is a convex function of p_i for $p_i \geq F_i^{-1}(z_{i0})$. The above condition holds for many cumulative distribution functions, e.g., in case of the univariate normal distribution, we have $z_0 = 0$. However, the logconcavity property does not carry over for sums, the function in the left-hand side of the constraint $\sum_{i=1}^{r} F_i(T_i x) - (r-1) \geq p$ is not guaranteed to be concave, which in turn implies that the program (7) is not necessarily convex. The same remark applies for problems (12) and (13). Moreover, (12) and (13) contain equality constraints involving nonlinear functions.

The solutions of the nonlinear optimization problems described above are solved with the open-source solver Ipopt (Wächter, Biegler, 2006) which implements an interior point line search filter method used to solve problems of the form

$$\begin{aligned} &\min f(x) \\ &\quad g^L \leq g(x) \leq g^U \\ &\quad x^L \leq x \leq x^U, \end{aligned}$$

where $x \in R^n$ are the decision variables, possibly with lower and upper bounds x^L and x^U, f is the objective functions and g are the general nonlinear constraints with lower and upper bounds $g(x^L)$ and $g(x^U)$. The functions $f(x)$ and $g(x)$ can be linear or nonlinear, convex or concave, but must be twice differentiable.

The models are coded using the AMPL modeling language. In the computational tests, the random variables are assumed to be normally distributed. Therefore, the solution of the joint probabilistic problem and of its approximations necessitates the computation of univariate and multivariate normal probabilities, and their incorporation in the optimization procedure. The computations of the normal probabilities are carried out using Genz' method (Genz, 1992). Besides the calculation of those probabilities, the interior point approach also requires the computations at each iteration of the first and second derivatives of the normal distributions which have closed-form formulations. We use the specialized AMPL external function constructed by Bonami and Lejeune (2007) to compute the normal probabilities and their first and second derivatives whose values can be "transmitted" to the Ipopt solver.

3.2 Test Sets

3.2.1 Reservoir Management – Problem I

The objective is to design a reservoir system, in which reservoirs are used to hedge against the possibility of flooding that may occur as a result of random stream of water. The probabilistic constraints impose limits on the probability that the water rises above the downstream reservoir capacity. The capacities x_j of the reservoirs j are the decision variables. They are limited from above, and are to be designed in such a way that they can retain the flood by a prescribed, large probability preventing from it the downstream locations. We consider a problem, in which there are two possible sites, on a main river, where reservoirs can be built, and a tributary comes in between the two sites. Assume that once in a year water amounts ξ_1, ξ_2 are to be retained by the two reservoirs. Let x_1 and x_2 designate upstream and downstream reservoir capacities, respectively. Let further c_1 and c_2 be the cost per unit of reservoir capacity. For references, see Prékopa (1995).

The probabilistically constrained problem for the design of a reservoir system containing two reservoirs takes the following form:

$$
\begin{aligned}
\min\ & c_1 x_1 + c_2 x_2 \\
s.to\ & \mathbb{P}\begin{pmatrix} x_1 + x_2 \geq \xi_1 + \xi_2 \\ x_2 \geq \xi_2 \end{pmatrix} \geq p \\
& 0 \leq x_1 \leq V_1 \\
& 0 \leq x_2 \leq V_2 .
\end{aligned}
$$

The random variables ξ_1, ξ_2 are assumed to be normally distributed: $\xi_i \sim N(\mu_i, \sigma_i^2), i = 1,2$, $\mu_1 = 1, \sigma_1 = 0.1,\ \mu_2 = 2, \sigma_2 = 0.2$, and the covariance $\operatorname{cov}(\xi_1, \xi_2)$ between ξ_1 and ξ_2 is known. The random variable $\theta_1 = \xi_1 + \xi_2$ has mean $E[\theta_1] = 3$ and variance $Var[\theta_1] = Var[\xi_1] + Var[\xi_2] + 2\operatorname{cov}(\xi_1, \xi_2)$, where $\operatorname{cov}(\xi_1, \xi_2) = \rho \cdot \sigma_1 \cdot \sigma_2$, and ρ denotes the correlation coefficient between ξ_1 and ξ_2.

Table 1 displays the considered problem instances, i.e., the different values taken by the coefficients of the objective function, the correlations between the random variables, the reliability level p, and the maximal quantity that the reservoirs can contain, while Table 2 reports the value of the objective function and the values of the decision variables for the various problem formulations previously discussed.

Table 1: Parameter settings

Instance #	c_1	c_2	p	σ	cov	$Var[\theta_1]$	V_1	V_2
1	2	1	0.9	0	0	0.05	0.8	2.5
2	2	1	0.9	-0.8	-0.016	0.018	0.8	2.5
3	2	1	0.9	0.8	0.016	0.082	0.8	2.5
4	2	1	0.9	0.10	0.002	0.054	0.8	2.5
5	1	2	0.9	0	0	0.05	0.8	2.5
6	1	2	0.9	-0.8	-0.016	0.018	0.8	2.5
7	1	2	0.9	0.8	0.016	0.082	0.8	2.5
8	1	2	0.9	0.12	0.0023	0.055	0.8	2.5
9	1	2	0.99	0	0	0.05	2	5
10	1	2	0.99	-0.8	-0.016	0.018	2	5
11	1	2	0.99	0.8	0.016	0.082	2	5
12	1	2	0.99	0	0	0.05	0.8	3
13	1	2	0.99	-0.8	-0.016	0.018	0.8	3
14	1	2	0.99	0.8	0.016	0.082	0.8	3

It appears that the problems (1) and (8) result in almost the same optimal value z^* of the objective function in many parameter settings. However, in instances 4 and 8 in which the available resources are very tight to reach the prescribed reliability level, the approximation of form (8) turns out to be infeasible, while the original problem (1) is actually feasible. This indicates that (8) does not always provide a tight enough approximation.

Table 2: Optimal values for each parameter setting and formulation

	Problem (1) with joint probability constraints			Problem (8) with intersection of events approach		
#	z^*	x_1^*	x_2^*	z^*	x_1^*	x_2^*
1	4.088	0.794	2.500	4.089	0.795	2.500
2	3.853	0.677	2.500	3.854	0.677	2.500
3	INFEASIBLE					
4	4.096	0.798	2.500	INFEASIBLE		
5	5.789	0.800	2.494	5.790	0.800	2.495
6	5.585	0.800	2.393	5.586	0.800	2.393
7	INFEASIBLE					
8	5.796	0.800	2.498	INFEASIBLE		
9	6.090	1.052	2.519	6.091	1.052	2.520
10	5.858	0.856	2.501	5.858	0.856	2.501
11	6.218	1.193	2.513	6.250	1.189	2.530
12	6.243	0.800	2.721	6.243	0.800	2.721
13	5.870	0.800	2.535	5.870	0.800	2.535
14	6.532	0.800	2.866	6.533	0.800	2.866

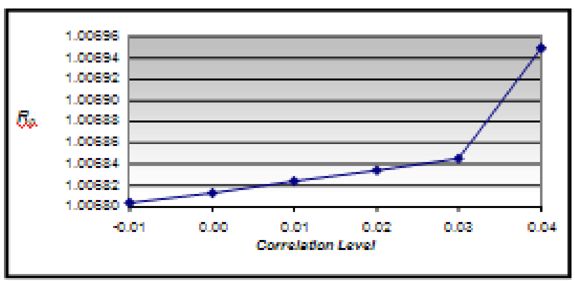
Considering now independent random variables, it appears that the solution of problem (2) does not present major difference in terms of computational tractability. The optimal value of the objective function in (2) is roughly the same as this of the programs analyzed in Table 2. However, unlike program (8) that is widely applicable, problem (2) assumes the random variables to be independent of each other.

Table 3: Optimal values with (2) for experiments with independent random variables

	Problem (2)				
#	z^*	x_1^*	x_2^*	p_1	p_2
1	4.088	0.794	2.500	0.994	0.906
4	5.789	0.800	2.494	0.993	0.906
7	6.091	1.052	2.519	0.995	0.995

Figure 1 represents the ratio $R_p = p_{IE(p)}/p$ of the actual reliability level $p_{IE(p)}$ *obtained* with the intersection of events problem to the *enforced* reliability level p as a function of the correlation level between the random variables. Clearly, the higher the value of the correlation among random variates, the larger the value of the ratio and the less tight the intersection of events approximation is.

Figure 1: Impact of correlation and enforced reliability level

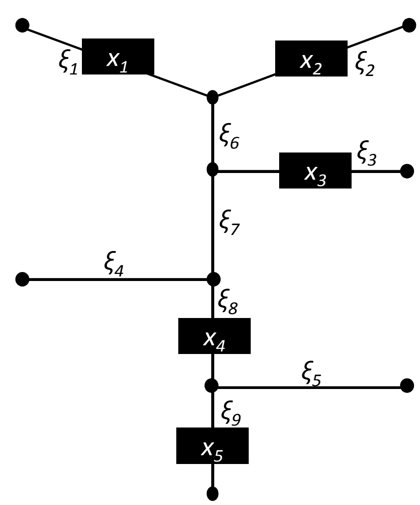


3.2.2 Reservoir Management – Problem II

The second and more complex reservoir management problem is based on the river network represented in Figure 2 and was first studied by Prékopa and Szántai (1978). As above, the goal is to design a reservoir system such that the likelihood of a flooding is below a high prescribed probability level.

The water stream is modeled as a five-dimensional random variable $\xi = (\xi_1,\ldots,\xi_5)$ following a normal distribution, while the water containment system will be composed of 5 reservoirs whose capacities $x_1,\ldots,x_5$ are upper-bounded decision variables.

Figure 2: River network



The auxiliary variables $\xi_6,\ldots,\xi_9$, depending on both stochastic and decision variables, are given by:

$$\begin{cases} \xi_6 \le \xi_1 - \min(\xi_1, x_1) + \xi_2 - \min(\xi_2, x_2) \\ \xi_7 \le \xi_3 - \min(\xi_3, x_3) + \xi_6 \\ \xi_8 \le \xi_4 + \xi_7 \\ \xi_9 \le \xi_8 - \min(\xi_8, x_4) + \xi_5. \end{cases}$$

The probabilistically constrained problem for the design of this problem reads:

$$\begin{aligned} &\min\ 0.4x_1 + 0.5x_2 + 0.6x_3 + 1.2x_4 + 1.8x_5 \\ &s.to\ \mathbb{P}\left(\xi_9 \le x_5\right) \ge p \\ &\qquad 0 \le x_1 \le 1 \\ &\qquad 0 \le x_2 \le 1 \\ &\qquad 0 \le x_3 \le 1 \\ &\qquad 0 \le x_4 \le 2 \\ &\qquad 0 \le x_5 \le 3. \end{aligned}$$

At first sight, this problem looks straightforward with a single individual probabilistic constraint constraining the happening of a flooding to be below the fixed probability level p. However, the typology of the river network is such that the inequality $\xi_9 \le x_5$ subject to the probabilistic constraint holds if the nine inequalities below hold jointly. Thus, the probabilistic constraint in the above problem is actually a joint probabilistic constraint that is formulated as follows:

$$\mathbb{P}\begin{pmatrix} \xi_5 & \le & x_5 \\ \xi_4 + \xi_5 & \le & x_4 + x_5 \\ \xi_1 + \xi_4 + \xi_5 & \le & x_1 + x_4 + x_5 \\ \xi_2 + \xi_4 + \xi_5 & \le & x_2 + x_4 + x_5 \\ \xi_3 + \xi_4 + \xi_5 & \le & x_3 + x_4 + x_5 \\ \xi_1 + \xi_2 + \xi_4 + \xi_5 & \le & x_1 + x_2 + x_4 + x_5 \\ \xi_1 + \xi_3 + \xi_4 + \xi_5 & \le & x_1 + x_3 + x_4 + x_5 \\ \xi_2 + \xi_3 + \xi_4 + \xi_5 & \le & x_2 + x_3 + x_4 + x_5 \\ \xi_1 + \xi_2 + \xi_3 + \xi_4 + \xi_5 & \le & x_1 + x_2 + x_3 + x_4 + x_5 \end{pmatrix}. \tag{23}$$

The means and variances of each component of ξ are reported in Table 4.

Table 4: Mean and variance of water flows

ξ_i	$E[\xi_i]$	$Var[\xi_i]$
ξ_1	0.8	0.2
ξ_2	1.5	0.3
ξ_3	1.2	0.6
ξ_4	0.5	0.4
ξ_5	0.7	0.3

We will consider the following three correlation matrices $R^{(i)}$, $i=1,2,3$:

$$R^{(1)} = \begin{pmatrix} 1 & 0 & 0.6 & 0.4 & 0 \\ 0 & 1 & 0.5 & 0.3 & 0.3 \\ 0.6 & 0.5 & 1 & 0.7 & 0.6 \\ 0.4 & 0.3 & 0.7 & 1 & 0.4 \\ 0 & 0.3 & 0.6 & 0.4 & 1 \end{pmatrix},\ R^{(2)} = \begin{pmatrix} 1 & -0.5 & 0 & 0.3 & -0.5 \\ -0.5 & 1 & -0.8 & 0 & 0.2 \\ 0 & -0.8 & 1 & 0 & 0.3 \\ 0.3 & 0 & & 1 & 0 \\ -0.5 & 0.2 & 0.3 & 0 & 1 \end{pmatrix},\ \text{and } R^{(3)} \text{ is an identity matrix.}$$

Considering the Bonferroni's inequality which gives the approximation of form (8) for the stochastic problem (1) subject to a joint probabilistic constraint joint, Chen et al. (2007) recommend the formulation

$$\begin{aligned} &\min c^T x \\ &s.to\ \ T_i x \geq F_i^{-1}(p_i), i=1,\dots,r \\ &\qquad A x \geq b \\ &\qquad x \geq 0 \end{aligned}, \tag{24}$$

where p_i are fixed parameters whose values are set equal to

$$p_i = 1 - \frac{(1-p)}{r}, i=1,\dots,r \tag{25}$$

and not decision variables. We recall that r is the number of dimensions in the multivariate random variable ξ. Problem (24) is a linear program, and the rationale for doing so, according to Chen et al. (2007), is that problem (8) is non-convex and possibly intractable if the p_i, $i=1,\dots,r$ are decision variables.

Table 5 reports, for each problem instance, the optimal value, the optimal solution obtained with the original formulation (3), the intersection of events approach (8), and the above approximation (25) that reformulates the stochastic programming problem as a linear programming one, and the "individual probability level" of each constraint. To each inequality (i=1,...,9) subject to the joint probability (23) condition, we compute the individual probability level as follows: $p_i = \mathbb{P}\left(\xi \leq x^*\right), i=1,\dots,9$, where x^* is the optimal solution for the problem with the considered formulation. In the first inequality (i=1), $\xi = \xi_5$ and $x^* = x_5^* = 1.396$ in the optimal solution of the joint probabilistic constraint formulation for the first problem instance. The acronym LP refers to approximation (25). Table 5 does not display the optimal values of x_2 and x_3, since they are equal to 1 in each instance and with each formulation.

Table 5: Comparison of optimal solutions

Formulation	R	p	Probability Level for Individual Constraint									Optimal Solution			Optimal Value
			p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9	x_1^*	x_4^*	x_5^*	z^*
O	$R^{(1)}$	0.8	0.990	0.998	0.998	0.970	0.937	0.959	0.919	0.828	0.811	0.800	1.720	1.396	5.995
IE	$R^{(1)}$	0.8	1	1	1	1	1	1	0.994	1	0.993	0.964	0.966	1.720	6.997
LP	$R^{(1)}$	0.8	0.998	1	1	0.998	0.995	0.998	0.996	0.965	0.970	0.99	2	2.48	8.368
O	$R^{(1)}$	0.9	1	1	1	0.995	0.976	0.995	0.976	0.910	0.918	0.998	1.885	1.524	6.869
IE	$R^{(1)}$	0.9	1	1	1	1	1	1	0.994	1	0.993	0.964	0.966	2.210	7.878
LP	$R^{(1)}$	0.9	1	1	1	1	0.997	0.999	0.996	0.978	0.978	1	2	2.85	9.036
O	$R^{(2)}$	0.8	0.987	0.996	0.999	0.952	0.941	0.973	0.951	0.865	0.900	0.906	1.351	1.371	5.551
IE	$R^{(2)}$	0.8	0.969	1	1	0.999	1	0.984	0.972	0.995	0.982	0.933	0.965	1.260	5.875
LP	$R^{(2)}$	0.8	0.990	1	1	0.997	0.991	0.998	0.989	0.978	0.978	0.76	2	1.40	6.320
O	$R^{(2)}$	0.9	1	1	1	0.987	0.976	0.992	0.976	0.941	0.952	0.833	1.239	1.830	6.214
IE	$R^{(2)}$	0.9	0.986	1	1	1	1	0.995	0.986	0.999	0.992	0.966	0.985	1.360	6.229
LP	$R^{(2)}$	0.9	0.999	1	1	1	0.995	1	0.994	0.989	0.989	0.75	2	1.60	6.689
O	$R^{(3)}$	0.8	0.993	0.994	0.999	0.950	0.946	0.970	0.965	0.817	0.967	1	1.226	1.431	5.547
IE	$R^{(3)}$	0.8	0.970	1	1	1	1	0.991	0.985	0.995	0.990	0.923	0.946	1.260	5.965
LP	$R^{(3)}$	0.8	0.999	1	1	0.999	0.997	0.999	0.997	0.978	0.978	0.85	2	1.58	6.686
O	$R^{(3)}$	0.9	0.988	0.999	1	0.988	0.981	0.993	0.988	0.910	0.938	1	1.650	1.374	5.953
IE	$R^{(3)}$	0.9	0.987	1	1	1	1	0.998	0.994	0.999	0.996	0.959	0.973	1.260	6.346
LP	$R^{(3)}$	0.9	1	1	1	1	0.999	1	0.999	0.999	0.999	0.85	2	1.81	7.105

A few comments are in order. First, as explained above, problem (8) is convex for a wide class of probability distributions, i.e., those whose probability density functions is logconcave (uniform (over a convex set) distribution, and, under some mild conditions, the multivariate Dirichlet, Wishart, normal and gamma distributions). Second, the above results (Table 5) shows that problem (8) is computationally tractable. It can be solved to optimality for each problem instance. Third, the results shown in Table 4 clearly indicate that the optimal solutions provided by the LP approximation (25) are much more conservative than those obtained with the IE approach and result in much higher costs. On average, over the six problem instances, it can be seen that:

- the optimal value of the objective function obtained with the LP approach (25) is 10.47% higher than that of the IE approach (8);
- the optimal value of the objective function obtained with the LP approach (25) (resp., the IE approach (8)) is 22.09% (resp., 8.61%) higher than that of the original problem (1).

Additionally, Table 1 shows that approximations are very conservative for the correlation matrix $R^{(1)}$ in which all the correlation coefficients are positive (or 0). Finally, with the original formulation (1), we note that the optimal objective function value is smaller (regardless of the value of *p*) for the correlation matrix $R^{(2)}$ than it is with the correlation matrix $R^{(3)}$ (which assumes that the random water streams are uncorrelated): $z^*\left(p;R^{(2)}\right) > z^*\left(p;R^{(3)}\right), p = 0.8, 0.9.$ However, the conclusion is reverse with the two approximation methods: $z^*\left(p;R^{(2)}\right) < z^*\left(p;R^{(3)}\right), p = 0.8, 0.9.$

3.2.3 Coffee blending

The coffee blending problem has a linear objective function which is minimized subject to a set of linear constraints related to the limited availability of the coffee types, the fulfillment of quality requirements, and a joint probability constraint that imposes that the demand for coffee be satisfied with probability *p,* representing the reliability level of the production system. The reader is referred to Szántai (1988) and Prékopa (1995) for more detailed explanations. Denoting by *D* the feasible set determined by the linear constraints, and setting $\xi_1 = \sum_{k=1}^{8} x_{k1}, \xi_2 = \sum_{k=1}^{8} x_{k2}$, and $\xi_3 = \sum_{k=1}^{8} x_{k3}$, the program is formulated as:

$$\begin{aligned} &\min c^T x \\ &s.to\ x \in D \\ &\quad \mathbb{P}\left(\sum_{k=1}^{8} x_{kl} \geq \xi_l, l = 1,2,3\right) \geq p \\ &\quad x_{kl} \geq 0, k = 1,\ldots,8, l = 1,\ldots,3, \end{aligned} \tag{26}$$

with $\xi_i \sim N(\mu_i, \sigma_i^2), 1 = 1,2,3$, with $\mu_1 = 3, \sigma_1 = 0.5, \mu_2 = 40, \sigma_2 = 5, \mu_3 = 20, \sigma_3 = 3$. We will consider the following three matrices $R^{(i)}$*, i=1,2,3* defining the correlation between random variables:

$$R^{(1)} = \begin{pmatrix} 1 & 0.1 & 0.1 \\ 0.1 & 1 & 0.9 \\ 0.1 & 0.9 & 1 \end{pmatrix}, \quad R^{(2)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{and} \quad R^{(3)} = \begin{pmatrix} 1 & 0.1 & 0.1 \\ 0.1 & 1 & -0.9 \\ 0.1 & -0.9 & 1 \end{pmatrix}.$$

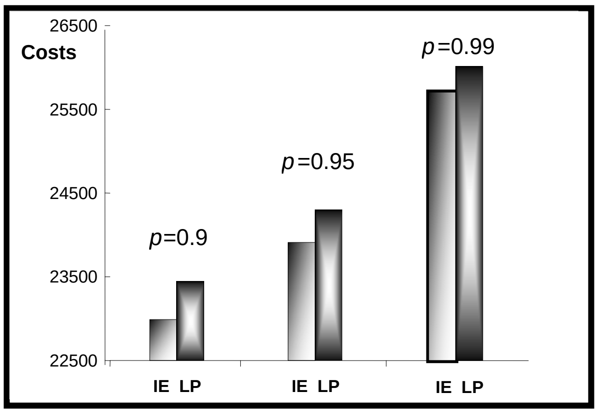
We provide in Table 6 the optimal objective value and the values of the random variables p_1, p_2, p_3 (i.e., the individual probability level) for various values of the enforced overall reliability level *p* in (8). Our solution method finds the optimal solution for all problem instances regardless of the enforced reliability level (*p* = 0.9, 0.95, 0.99), and justifies the computing resources allocated for the computation of the first- and second-order derivative of the normal probabilities. Table 6 reports the optimal value and individual probability levels with formulation (8) for the considered problem instances.

Table 6: Optimal values for problems with intersection of events bound

	Problem (8) with intersection of events			
p	p_1	p_2	p_3	z^*
0.9	0.992	0.925	0.983	22988
0.95	0.996	0.963	0.991	23910.4
0.99	0.999	0.993	0.998	25704.1

Figure 3 compares the optimal objective function values with the two approximation methods (25) and (8), and shows that the optimal value of (25) is larger, for each instance, than that of (8).

Figure 3: Comparison of approximation methods



Results for the lower (13) bound of problem (1), obtained using the binomial moment bounding scheme, are given in Table 7. All results have been obtained by setting the parameter α in (13) equal to 15. None of the problem instances formulated with (13) could be solved to optimality.

Regardless of the reliability level considered, it can be seen that the optimal objective values obtained with programs (1) and (13) for any of the three correlation matrices are lower than the optimal value obtained with program (8). It is logical that the optimal solution of (13) be lower than these of program (8), since (13) provides a lower bound on the objective value, while (8) approximates (1) through the enforcement of requirements that are at least as demanding as those of (1). The low magnitude of the gap between the optimal values of (1) on the one side and that of (8) on the other side indicates that (8) provides here a tight approximation of (1).

Table 7: Optimal values for the programs involving a joint probabilistic constraint and the binomial moment bounding scheme

p	Correlation level	Problem (13) with binomial moment bounding scheme: Lower Bound	Problem (1) with joint probability constraints	Formulation (2) for independent random components
0.9	R_1	22458.4	22564.0	
0.9	R_2	22087.3	22949.4	22949.4
0.9	R_3	22242.2	22961.6	
0.95	R_1	23487.9	23603.6	
0.95	R_2	22598	23866.6	23866.6
0.95	R_3	23274.6	23885.2	
0.99	R_1	25420	25500.6	
0.99	R_2	25063.8	25702.0	25702.0
0.99	R_3	25203.5	25680.6	

Considering the correlation matrices $R^{(i)}, i = 1,2,3$, and denoting by $s_{ij}^{(k)}$ the *(i,j)*th element of $R^{(k)}$, one can see that $s_{ij}^{(1)} \geq s_{ij}^{(k)}, i, j = 1,...,r, k = 2,3$, therefore allowing the reliance on Slepian's inequality

$$\mathbb{P}\left(\frac{\xi_i - \mu_i}{\sigma_i} \leq z_i, i = 1,...,r; R^{(1)} \right) \geq \mathbb{P}\left(\frac{\xi_i - \mu_i}{\sigma_i} \leq z_i, i = 1,...,r; R^{(k)} \right), k = 2,3 \ .$$

For any reliability level enforced, it can be seen in Table 7 that the optimal values obtained when considering the correlation matrix $R^{(1)}$ are lower than those obtained when considering $R^{(2)}$ and $R^{(3)}$. The solution of the stochastic program associated with correlation matrix $R^{(1)}$ can moreover be solved with the introduction of the valid inequality (21).

3.2.4 Power management: STABIL problem

The last problem considered is the STABIL problem (Prékopa et al., 1980) involving the construction of a plan for the Hungarian electrical energy sector in the seventies. It has a linear objective function minimizing the profit function multiplying by –1, while satisfying 106 deterministic constraints (manpower balance, investment features, foreign trade balance, balance of the state budget, finance, and electricity demand satisfaction), as well as the joint probabilistic constraint described below.

The problem is formulated by:

$$\begin{aligned} &\min c^T x \\ &s.to\ a_i x \geq b_i, i = 1,...,106 \\ &\quad \mathbb{P}\begin{pmatrix} -25x_{25} \geq \xi_1 \\ -16.67x_{26} \geq \xi_2 \\ 0.8696x_{24} + x_{40} \geq \xi_3 \\ 0.9(x_1 + x_2 + x_3 + x_4) - 0.115x_{24} \geq \xi_4 \end{pmatrix} \geq p \\ &\quad x \geq 0, \end{aligned} \tag{27}$$

where $c^T x$ is given by x_{35} - x_{36}, x_{35} and x_{36} representing respectively the increase in the wage bill and the enterprise profit before taxation, and $\xi_i, i=1,...,4$ are normally distributed random variables with the following means and standard deviations:

$$\mu_1 = -48313, \sigma_1 = 483, \quad \mu_2 = -426, \sigma_2 = 4, \quad \mu_3 = 16000, \sigma_3 = 160, \quad \mu_4 = 14950, \sigma_4 = 19 \; .$$

The joint distribution of the random variables is normal, and the following two correlation matrices are considered:

$$R^{(1)} = \begin{pmatrix} 1 & -0.8 & 0.4 & 0.4 \\ -0.8 & 1 & 0.1 & 0.1 \\ 0.4 & 0.1 & 1 & 0.9 \\ 0.4 & 0.1 & 0.9 & 1 \end{pmatrix} \quad , \quad R^{(2)} = \begin{pmatrix} 1 & -0.7 & 0.3 & 0.3 \\ -0.7 & 1 & 0.1 & 0.1 \\ 0.3 & 0.1 & 1 & 0.9 \\ 0.3 & 0.1 & 0.9 & 1 \end{pmatrix}$$

The first two components of (27) restrain the planned deficit of foreign trade (in $US and roubles) to be below a certain level, while the last two components express the relationships between the electrical sector and the other sectors of the Hungarian economy.

Below, the solution of the program associated with formulation (8) based on the intersection of events containing a set of individual constraints is reported and discussed. The two reliability levels *p = 0.9 , 0.95* are considered and the optimal values of the corresponding optimization problems are respectively equal to -4370.3 and -4369.1. In Table 8, we report the optimal reliability levels associated with each individual stochastic constraint $\mathbb{P}\left(\sum_j t_{ij} x_j \geq \xi_i\right) \geq p_i, i=1,...,r$ corresponding to each of the inequalities $\sum_j t_{ij} x_j \geq \xi_i, i=1,...,r$ in the joint probabilistic constraint (27).

Table 8: Reliability level for each stochastic inequality (8)

	p=0.9	*p=0.95*
i	p_i	p_i
1	98.55%	98.74%
2	99.99%	99.99%
3	100%	100%
4	91.33%	96.26%

As for the coffee blending problem, we observe that the optimal individual reliability levels differ very much, ranging from 91.33% to 100% and from 96.26% to 100% when the overall enforced reliability levels are, respectively, 90% and 95%. Enforcing identical (25) individual reliability levels would be detrimental.

Table 9 reports the best solution found for the original problem (1) containing a joint probability constraint and for the relaxation (13) based on the binomial moment bounding scheme.

Table 9: Optimal values

Setting	Problem (1) with joint probability constraints	Problem (13) with binomial moment bounding scheme: Lower Bound
$R^{(1)}$, $p=0.9$	-4370.3	-4384.9
$R^{(1)}$, $p=0.95$	-4369.4	-4384.91
$R^{(2)}$, $p=0.9$	- 4370.6	-4384.92
$R^{(2)}$, $p=0.95$	-4369.9	-4382.6

From a managerial perspective, it is very important to note that the optimal value of the probabilistically constrained problem is the same as that of the underlying deterministic problem ($R^{(1)}$, $p=0.9$). However, their optimal solutions differ significantly: the optimal solution of the underlying deterministic problem has a reliability level of 10% while that of the probabilistically constrained problem gives a reliability level of 90%! It turns out that the appropriate formulation of the problem above allows reaching a nine times higher reliability level without having to support additional costs.

3.3 Impact of correlation on approximation tightness

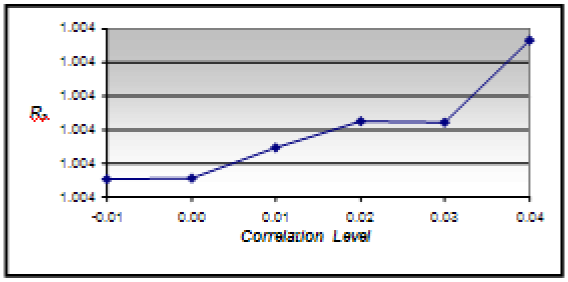
In this section, we study the effect about the correlation structure among random variables on the tightness of the approximation approaches for the probabilistically constrained problem of form (1). More precisely, we study the impact of the correlation level between the variates of the multidimensional random vector on the following metrics:

- the ratio $R_z = z_j^* / z_O^*$ of the optimal value of the approximation problem j (i.e., j =IE (8) or LP (25)) to that of the original problem (1). We examine the impact of the correlation structure and the enforced probability level p on the value taken by R_z ;
- the highest reliability level ($\overline{p}_j$) that can be *enforced* with the approximation approach j given the correlation structure.

The notation z_j^* is the optimal value obtained with formulation *j* when the enforced reliability level is *p*. The acronyms O, IE, and LP respectively refer to the problem with joint probabilistic constraint (1), the intersection of events problem (28), and the linear programming approximation (25).

The results displayed in Figure 4 are based on the first reservoir problem. The graph shows that the ratio R_z of the optimal value of the approximation problem (8) to that of the original problem is an increasing function of the level of positive correlation between random variables.

Figure 4: Impact of correlation on ratio of optimal values and maximum reliability level



The same empirical conclusions can be drawn from the second reservoir problem. Table 10 shows that the relative gap measured by R_z between the optimal value of the approximations and that of the original problem increases with the correlation level. We recall that $R^{(2)}$contains positive and negative correlation components, $R^{(3)}$represents the case of uncorrelated variables, and that $R^{(1)}$only contains positive correlation components. This implies that the relative cost (i.e., as measured by the value of the objective function) of using an approximation method instead of the original formulation increases with the correlation level. Table 10 also indicates that the relative difference between the optimal objective value of the approximation problems on one hand and that of the original problem on the other one decreases as *p* increases.

Table 10: Impact of correlation

Formulation	P	Ratio of optimal value R_z		
		$R^{(2)}$	$R^{(3)}$	$R^{(1)}$
IE	0.8	103%	108%	117%
LP	0.8	111%	121%	139%
IE	0.9	101%	107%	115%
LP	0.9	107%	120%	132%

Regarding the feasibility of the approximation problems, we observe that a feasible solution for the second reservoir problem can be found for a very high reliability level (i.e., 0.9999) for the three formulations (O, IE, LP) when we consider the correlation matrices $R^{(2)}$ and $R^{(3)}$. However, in the case of $R^{(1)}$, which assumes a positive correlation among all components of the random vector,

- a reliability level of 0.995 can be enforced with the original formulation, while,
- the largest probability level for which a feasible solution can be found with the approximations IE and LP is respectively equal to 0.982 and 0.924.

The above empirical experiences attest the fact that the approximation approaches become increasingly conservative and costly when the positive correlation among random variables increases. The approximation approach can even become so conservative that the corresponding optimization problem becomes infeasible, while the original problem is actually feasible (see also first reservoir problem, instances 4 and 8 in Table 2).

4 Convexity analysis of constraints enforcing reliability level

4.1 Constraints involving conditional expectations

Until now, we have studied several approximations of the probabilistic constraint in (1). There is another possibility, however, to replace the probabilistic constraint by constraints involving conditional expectations. A model of this type was first proposed by Prékopa (1973). This means that problem (1) is replaced by the following one:

$$\begin{aligned}
&\min c^T x \\
&s.to\ \mathrm{E}\left[\xi_i - T_i x \mid \xi_i - T_i x > 0\right] \le d_i, i = 1,\dots,r \\
&\quad Ax \ge b \\
&\quad x \ge 0.
\end{aligned} \tag{29}$$

It is proved in the aforementioned paper that if the univariate random variable ξ has continuous distribution and logconcave probability density function, then the function

$$g(t) = \mathrm{E}\left[\xi - t \middle| \xi - t > 0\right] \tag{30}$$

is decreasing; the function can also be written as

$$g(t)=\frac{\int_t^{\infty}(1-F(x))dx}{1-F(t)} \quad ,$$

where F is the probability distribution function of ξ. If we assume that each random variable $\xi_i, i=1,...,r$ has that property and designate by $g_i(t)$ the function corresponding to ξ_i, then problem (29) can be reformulated in the following manner:

$$\begin{aligned} &\min c^T x \\ &s.to\ T_i x \geq g_i^{-1}(d_i), i=1,...,r \\ &\quad Ax \geq b \\ &\quad x \geq 0. \end{aligned} \tag{31}$$

Problem (31) is a linear programming problem and can be solved in a standard way.

Similarly, as the approximating problem (8) is created from problem (3), we can take $d_i,...,d_r$ as variables in (31) and impose on them the additional constraint

$$\sum_{i=1}^{r} d_i \mathbb{P}\left(\xi_i - T_i x > 0\right) \leq d_0 \quad ,$$

where d_0 is a constant. In that case, however, problem (31) is not necessarily convex. A sufficient condition of the convexity of the problem is that each function $g_i^{-1}(d_i), i=1,...,r$ is convex, or, what is the same, that each function $g_i\,(d_i), i=1,...,r$ is concave.

We now study the concavity of (30) when the random variable has a normal distribution. For ξ_i having a standard normal distribution with mean 0 and standard deviation 1, we have

$$f(t)=\varphi(t)=\frac{e^{-t^2/2}}{\sqrt{2\pi}} \quad \text{, and} \quad \varphi'(t)=-t\varphi(t)$$

$$F(t)=\Phi(t)=\int_{-\infty}^{t}\frac{e^{-x^2/2}}{\sqrt{2\pi}}dx \quad \text{, and} \quad F'(t)=\varphi(t).$$

The first derivative of *g* with respect to *t* is given by

$$\begin{aligned} g'(t) &= \frac{-(1-\Phi(t))^2+\varphi(t)\int_t^{\infty}(1-\Phi(x))\,dx}{(1-\Phi(t))^2} \\ &= -1+\frac{\varphi(t)}{1-\Phi(t)}g(t). \end{aligned} \tag{32}$$

Since $\varphi(t)$ is a logconcave function, it follows that *g'(t) < 0*. For the second derivative of *g(t)*, we obtain:

$$
\begin{aligned}
g''(t) &= \left(\frac{d}{dt}\frac{\varphi(t)}{1-\Phi(t)}\right)g(t) + \frac{\varphi(t)}{1-\Phi(t)}g'(t) \\
&= \frac{\varphi'(t)(1-\Phi(t)) + \varphi^2(t)}{(1-\Phi(t))^2}g(t) + \frac{\varphi(t)}{1-\Phi(t)}g'(t) \\
&= \frac{-t\varphi(t)(1-\Phi(t)) + \varphi^2(t)}{(1-\Phi(t))^2}g(t) + \frac{\varphi(t)}{1-\Phi(t)}\left(-1 + \frac{\varphi(t)}{1-\Phi(t)}g(t)\right) \\
&= \frac{-t\varphi(t)g(t)}{1-\Phi(t)} + \frac{\varphi^2(t)g(t)}{(1-\Phi(t))^2} - \frac{\varphi(t)}{1-\Phi(t)} + \frac{\varphi^2(t)g(t)}{(1-\Phi(t))^2} \\
&= \frac{2\varphi^2(t)g(t)}{(1-\Phi(t))^2} - \frac{\varphi(t)}{1-\Phi(t)}(tg(t)+1) \\
&= \underbrace{\frac{\varphi(t)}{1-\Phi(t)}}_{\geq 0}\left(\frac{2\varphi(t)g(t)}{1-\Phi(t)} - tg(t) - 1\right).
\end{aligned}
$$

To test the concavity of *g(t)*, we must therefore evaluate the sign of

$$2\varphi(t)g(t) - (t\, g(t)+1)(1-\Phi(t)). \tag{33}$$

The solution of

$$
\begin{aligned}
&\min \quad t \\
&s.to \quad 2\varphi(t)g(t) - (t\, g(t)+1)(1-\Phi(t)) \leq 0 \\
&\qquad\quad t \in R
\end{aligned} \tag{34}
$$

shows that the smallest value of t for which (33) is negative is t = 7.20369 for which the objective value of (34) is equal to -0.000249534; (33) is thus positive on

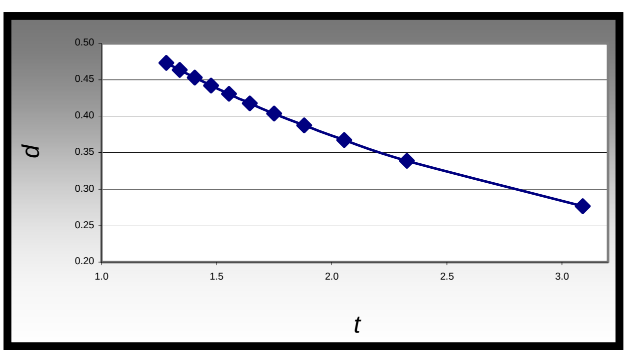
$$]-\infty, 7.20369[\ .$$

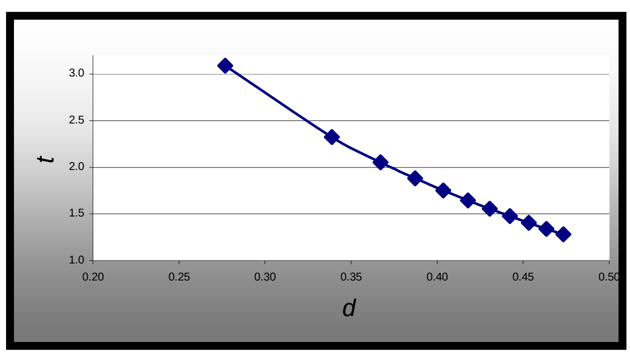
Finally, we note that

$$\mathbb{P}(\xi \leq 7.20369) = 0.999999999999705,$$

which means that *g(t)* (30) is convex except for extremely large values of t, and that the feasible set determined by (30) is convex except for extremely demanding reliability levels d_i characterized by extremely low values for d_i, requiring extremely high values of t for (30) to hold. This also implies that the inverse $g^{-1}(t)$ of *g(t)* is convex for most values of t (right-hand side in Figure 5).

Figure 5: Function g(t)





4.2 Correspondence between types of reliability levels

In this section, we proceed to a correspondence study between the reliability guaranteed by two types of stochastic constraints. Assuming that the random variable has a standard normal distribution, we study the correspondence between the service level p_i that requires the probabilistic constraint of form

$$\mathbb{P}(t_i \geq \xi_i) \geq p_i \quad ,$$

which is equivalent to

$$t_i \geq F_i^{-1}(p_i)$$

and the reliability level d_i that requires the constraint (in (29), where $d_i,\ldots,d_r$ are not decision variables)

$$\mathrm{E}\left[\xi_i - T_i x \mid \xi_i - T_i x > 0\right] \leq d_i \,.$$

The latter one is equivalent to

$$T_i x \geq g_i^{-1}(d_i)\,.$$

To determine which value of d_i corresponds to various probabilities (p_i = 0.9, 0.95, …), we solve the following non-linear problem:

$$\begin{aligned} &\min \; d_i \\ &s.to \; \frac{\int_{t^p}^{\infty}(1-F_i(x))\,dx}{1\text{-}F_i(t^p)} \leq d_i \\ &\quad d_i \geq 0, \end{aligned} \tag{35}$$

where $t^p = F_i^{-1}(p_i)$ is given and d_i is the decision variable.

The correspondence between the two types of reliability levels obtained by solving (35) for different values of p_i are given in Table 11.

Table 11 : Correspondence between p_i and d_i

p_i	0.9	0.91	0.92	0.93	0.94	0.95	0.96	0.97	0.98	0.99	0.999
d_i	0.4734	0.4636	0.4533	0.4423	0.4306	0.4179	0.4037	0.3873	0.3672	0.3389	0.2769

In view of the difficulty to solve probabilistically constrained problems, the results of the correspondence study are especially valuable. Two different types of constraint (i.e., joint probabilistic constraint and conditional expectation constraint) enforcing the same requirements, and allows the choice of the associated problem formulation that is most convenient (i.e., with respect to the convexity results derived above) to solve.

4.3 Application

After having shown in Section 4.1 that the enforcement of reliability levels of type d_i with normally distributed random variables result in non-linear convex problems, we consider the STABIL problem in which similar reliability levels are prescribed. More precisely, we solve the problem below that enforces the reliability level d_0

$$\begin{aligned}
&\min c^T x \\
&s.to\ \mathrm{E}\left[\xi_i - T_i x \middle| \xi_i - T_i x > 0\right] \le d_i, i = 1,\dots,r \\
&\qquad \sum_{i=1}^{r}\left(d_i \mathbb{P}\left(\xi_i - T_i x > 0\right)\right) \le d_0 \\
&\qquad Ax \ge b \\
&\qquad x \ge 0 \\
&\qquad d_i \in R, i = 1,\dots,r
\end{aligned} \tag{36}$$

with $d_i, i = 1,\dots,r$ being decision variables.
Substituting $d_i, i = 1,\dots,r$ by

$$\mathrm{E}\left[\xi_i - T_i x \middle| \xi_i - T_i x > 0\right], i = 1,\dots,r,$$

the constraint

$$\sum_{i=1}^{r}(d_i\ \mathbb{P}(\xi_i - T_i x > 0)) \le d_0$$

becomes

$$\sum_{i=1}^{r}\left(\mathrm{E}\left[\xi_i - T_i x \middle| \xi_i - T_i x > 0\right]\mathbb{P}(\xi_i - T_i x > 0)\right) \le d_0 \quad . \tag{37}$$

Since

$$\mathrm{E}\left[\xi_i - T_i x \middle| \xi_i - T_i x > 0\right] = \frac{\int_{T_i x}^{\infty}(z_i - T_i x) f(z_i)\, dz_i}{1 - F_i(T_i x)} \quad ,$$

and

$$\mathbb{P}\left(\xi_i - T_i x > 0\right) = 1 - F_i(T_i x) \ ,$$

constraint (37) can be rewritten as

$$\int_{T_i x}^{\infty}(z_i - T_i x) f(z_i)\, dz_i$$

and is thus an "integrated probabilistic constraint" (Klein Haneveld, 1986), in which

$$\eta_i = \left[z_i - T_i x\right]_+, i = 1,\dots,r$$

is the positive part of the expression $z_i - T_i x, i = 1,\dots,r$ and can be interpreted, in the STABIL context, as the amount of unserved power. It is thus clear that problem (36) restrains the expected amount of unserved power to be below a certain threshold d.

Table 12 displays the results for reliability levels *d* equal to 0.473, 0.418 and 0.339, which respectively correspond to a probability level *p* of not having a stockout equal to 0.9, 0.95 and 0.99 (Table 12).

Table 12 : Reliability level d for the STABIL problem

p	0.9	0.95	0.99
d	0.473	0.418	0.339
$\int_{T_1x}^{\infty} (z_1 - T_1x) f(z_1)\, dz_1$	0.158	0.139	0.112
$\int_{T_2x}^{\infty} (z_2 - T_2x) f(z_2)\, dz_2$	0.152	0.135	0.111
$\int_{T_3x}^{\infty} (z_3 - T_3x) f(z_3)\, dz_3$	0	0	0
$\int_{T_4x}^{\infty} (z_4 - T_4x) f(z_4)\, dz_4$	0.163	0.144	0.116
Objective value	-4373.080	-4372.840	-4372.441

5 Conclusion

We considered probabilistically constrained problems, in which the random variables are located in the right-hand sides of linear constraints. The objective function is linear, and its optimization is subject to a set of linear constraints and one joint probabilistic constraint. This latter one guarantees the joint fulfillment of a system of linear inequalities with random right-hand sides to be above a probability level.

We proposed several approximations for such probabilistic problems, and solved the corresponding optimization problems. We then discussed the characteristics of the relaxations, and evaluated their tightness and computational tractability on applications related to reservoir system design, coffee blending and power systems. We examined in which cases the proposed relaxations do not approximate closely the original joint probabilistic formulation. These results, supported by computational experiments, have a very insightful managerial content, in that they pinpoint under which circumstances the decisions stemming from these approximation models enforce stronger conditions than those initially required, and can be suitable for risk-averse decision making. We also gave insights into the conditions under which the approximation problems become infeasible, while the original one is actually feasible.

In the last part of this paper, we focused upon problems in which the probabilistic constraint is replaced by conditional expectation constraints or integrated chance constraints. Considering the standard normal distribution, we derived convexity results for the conditional expectation functions and for the feasibility set of the problem in which they appear. We studied the correspondence between the reliability levels that appear in two types of constraints mentioned above. The correspondence study is very valuable, since it identifies two very different problem formulations (i.e., one with joint probabilistic constraints and one with conditional expectation constraints) enforcing the same requirements, and allows the choice of the one that is the easiest to solve: the choice can be made with respect to the derived convexity properties of the various proposed formulations. Finally, a power real-life problem highlights the importance of correctly modeling reliability. More precisely, we show how a very substantial increase in the reliability level (9 times higher) can be obtained, without having to support any additional costs, by using the correct probabilistic formulation and relying on an appropriate solution technique.

References

Bawa V. S. 1973. On Chance Constrained Programming Problems with Joint Constraints. *Management Science* **10** (11), 1326-1331.

Beraldi P.,Ruszczyński A.2002.The Probabilistic Set Covering Problem.*OperationsResearch* **50**,956-967.

Bonami P., Lejeune M.A. 2007. Convex Integer Probabilistic Problems. *Working Paper*.

Boros E., Elbassioni K., Gurvich V., Khachiyan L., Makino K. 2003. An Intersection Inequality for Discrete Distributions and Related Generation Problems. In: *Lecture Notes in Computer Science* **2719**. Eds: Baeten J.C.M., Lenstra J.K., Parrow J., Woeginger G.J. Springer, Berlin-Heidelberg, 543-555.

Charnes A., Cooper W.W., Symonds G.H. 1958. Cost Horizons and Certainty Equivalents: An Approach to Stochastic Programming of Heating Oil. *Management Science* **4**, 235-263.

Chen W., Sim M., Sun J. 2007. A Robust Optimization Perspective of Stochastic Programming. *Operations Research* **55** (6), 1058-1071.

Deák I. 1988. Multidimensional Integration and Stochastic Programming. In: *Numerical Techniques for Stochastic Optimization.* Eds: Ermoliev Y., Wets R.J.-B. Springer-Verlag, Berlin, 187-200.

Dentcheva D., Prékopa A., Ruszczyński A. 2000. Concavity and Efficient Points of Discrete Distributions in Probabilistic Programming. *Mathematical Programming* **89**, 55-77.

Genz A. 1992. Numerical Computation of Multivariate Normal Probabilities. *Journal of Computational and Graphical Statistics* **1**, 141-149.

Genz A., Bretz F. 2002. Methods for the Computation of Multivariate *t*-Probabilities. *Journal of Computational and Graphical Statistics* **11**, 950-971.

Gren I.-M. 2008. Adaptation and Mitigation Strategies for Controlling Stochastic Water Pollution: An Application to the Baltic Sea. *Ecological Economics* **66** (2-3), 337-347.

Henrion R. 2004. Introduction to Chance-Constrained Programming. *Tutorial Paper for the Stochastic Programming Community Home Page*.

Henrion R., Li P., Möller A., Steinbach M., Wendt M., Wozny G. 2001a. Stochastic Optimization for Operating Chemical Processes under Uncertainty. In:*Optimization of Large Scale Systems*. Eds: Grötschel M., Krumke S., Rambau J. Online Optimization of Large Scale Systems. Springer, Berlin, 457-478.

Henrion R., Li P., Möller A., Wendt M., Wozny G. 2001b. Optimal Control of a Continuous Distillation Process under Probabilistic Constraints. In: *Optimization of Large Scale Systems*. Eds: Grötschel M., Krumke S., Rambau J. Springer. Berlin, 499-517.

Klein Haneveld W.K. 1986. Duality in Stochastic Linear and Dynamic Programming. *Lecture Notes in Economics and Mathematical Systems* 274. Springer-Verlag. New York, NY.

Kress M., Penn M., Polukarov M. 2008. The Minmax Multidimensional Knapsack Problem with Application to a Chance-Constrained Problem. *Naval Research Logistics* **54**, 656-666.

Lejeune M.A. 2008. A Unified Approach for Cycle Service Levels. *Working Paper*: http://www.optimization-online.org/DB_HTML/2008/11/2144.html.

Lejeune M.A., Ruszczyński A. 2007. An Efficient Trajectory Method for Probabilistic Inventory-Production-Distribution Problems. *Operations Research* **55** (3) 1-17.

Luedtke J., Ahmed S., Nemhauser G. 2009. An Integer Programming Approach for Linear Programs with Probabilistic Constraints. Forthcoming *in Mathematical Programming.*

Miller B.L., Wagner H.M. 1965. Chance Constrained Programming with Joint Constraints. *Operations Research* **13**, 930-945.

Prékopa A. 1970. On Probabilistic Constrained Programming. *Proceedings of the Princeton Symposium on Mathematical Programming*. Princeton University Press, 113-138.

Prékopa A. 1971. Logarithmic Concave Measures with Applications to Stochastic Programming. *Acta Sci. Math.* **32**, 301-316.

Prékopa A. 1973. Contributions to the Theory of Stochastic Programming. *Mathematical Programming* **4**, 202-221.

Prékopa A. 1990. Dual Method for a One-Stage Stochastic Programming with Random *rhs* Obeying a Discrete Probability Distribution. *Zeitschrift of Operations Research* **34**, 441-461.

Prékopa A. 1995. *Stochastic Programming*. Kluwer. Dordrecht-Boston.

Prékopa A. 1999. The Use of Discrete Moment Bounds in Probabilistic Constrained Stochastic Programming Models. *Annals of Operations Research* **85**, 21-38.

Prékopa A. 2001. On the Concavity of Multivariate Probability Distribution Functions. *Operations Research Letters* **29**, 1-4.

Prékopa A. 2003. Probabilistic Programming. In *Stochastic Programming: Handbook in Operations Research and Management Science 10.* Eds: Ruszczyński A. Shapiro A. Elsevier Science Ltd, 267-351.

Prékopa A., Ganczer S., Deák I., Patyi K. 1980. The STABIL Stochastic Programming Model and its Experimental Application to the Electrical Energy Sector of the Hungarian Economy. In *Stochastic Programming.* Ed: Dempster M.A. Academic Press. London, 369-385.

Prékopa A., Szántai T. 1978. Flood Control Reservoir System Design Using Stochastic Programming. *Mathematical Programming Study* **9**, 138-151.

Prékopa A., Vizvári B., Badics T. 1998. Programming under Probabilistic Constraint with Discrete Random Variable. In *New Trends in Mathematical Programming.* Eds: Gianessi F., Rapcsák S., Komlósi T. Kluwer Academic Publishers, 235-255.

Sen S. 1992. Relaxations for Probabilistically Constrained Programs with Discrete Random Variables. *Operations Research Letters* **11**, 81-86.

Sen S. 1996. Stochastic Programming: Computational Issues and Challenges. In: *Encyclopedia of Operations Research and Management Science*. Eds: Harris C.I., Gass S.I. Kluwer Academic Publishers.

Slepian D. 1962. On the One-Sided Barrier Problem for Gaussian Noise. *Bell System Technical Journal* **41**, 463-501.

Szántai T. 1988. A Computer Code for Solution of Probabilistic-Constrained Stochastic Programming Problems. In: *Numerical Techniques for Stochastic Optimization.* Eds: Ermoliev Y., Wets R.J.-B. Springer-Verlag. New York, 229-235.

Wächter A., Biegler L.T. 2006. On the Implementation of a Primal-Dual Interior Point Filter Line Search Algorithm for Large-Scale Nonlinear Programming. *Mathematical Programming* **106** (1), 25-57.