RUTCOR
RESEARCH
REPORT

The Use of Discrete Moment Bounds in Probabilistic Constrained Stochastic Programming Models

András Prékopa [a]

RRR 27-98, September 1998

RUTCOR
Rutgers Center for
Operations Research
Rutgers University
640 Bartholomew Road
Piscataway, New Jersey
08854-8003
Telephone: 732-445-3804
Telefax: 732-445-5472
Email: rrr@rutcor.rutgers.edu
http://rutcor.rutgers.edu/~rrr

[a]RUTCOR, Rutgers Center for Operations Research Rutgers University 640 Bartholomew Road Piscataway, NJ 08854-8003. E-mail: prekopa@rutcor.rutgers.edu

RUTCOR RESEARCH REPORT
RRR 27-98, SEPTEMBER 1998

THE USE OF DISCRETE MOMENT BOUNDS IN PROBABILISTIC CONSTRAINED STOCHASTIC PROGRAMMING MODELS

András Prékopa

Abstract. In the past few years efficient methods have been developed for bounding probabilities and expectations concerning univariate and multivariate random variables based on the knowledge of some of their moments. Closed form as well as algorithmic lower and upper bounds of this type are now available. The lower and upper bounds are frequently close enough even if the number of utilized moments is relatively small. This paper shows how the probability bounds can be incorporated in probabilistic constrained stochastic programming models in order to obtain approximate solutions for them in a relatively simple way.

Acknowledgements: Work supported by OTKA, Hungarian National Fund for Scientific Research, project no. T014102.

1 Introduction

The underlying deterministic problems that are the starting points of the most important static stochastic programming model formulations are:

$$\begin{aligned} &\min c^T x \\ &\text{subject to} \\ &Tx = \xi \\ &Ax = b \\ &x \geq 0 \end{aligned} \tag{1.1}$$

and

$$\begin{aligned} &\min c^T x \\ &\text{subject to} \\ &Tx \geq \xi \\ &Ax = b \\ &x \geq 0, \end{aligned} \tag{1.2}$$

where A is an $m \times n$, T is an $r \times n$ matrix; c, x are n-component vectors, b and ξ are m-and r-component vectors, respectively.

Based on problems (1.1) and (1.2) we formulate the recourse type stochastic programming problem as:

$$\begin{aligned} &\min\{c^T x + \mathbb{E}[q(\xi - Tx)]\} \\ &\text{subject to} \\ &Ax = b \\ &x \geq 0, \end{aligned} \tag{1.3}$$

where the function $q(z),\ z \in \mathbb{R}^n$ is usually convex, and it is called the penalty function.

If the underlying problem is (1.1), then we assume that $q(0) = 0$ while if it is problem (1.2), then we assume that $q(z) = 0$ for $z \leq 0$. We frequently choose

$$q(z) = \sum_{i=1}^{n} (q_i^+[z_i]_+ + q_i^-[-z_i]_+), \tag{1.4}$$

where $q_i^+ \geq 0, q_i^- \geq 0, i = 1, \ldots, r$ and $q_i^- = 0, i = 1, \ldots, r$, if the underlying problem is (1.2). If we use the function (1.4), then problem (1.3) is called the simple recourse problem. It was first studied by Dantzig (1955), Beale (1955). For a detailed discussion see Wets (1966).

A more general form of the penalty function is the following:

$$q(z) = \sum_{i=1}^{n} q_i(z_i), \tag{1.5}$$

where $q_1(z_1), \ldots, q_r(z_r)$ are convex or higher order univariate convex functions. A univariate function $f(z)$ is said to be convex of order k if its k-th order devided differences are nonnegative on the set where the function is defined (see Popoviciu (1945)). Second order convexity means convexity of the function. If the penalty function has the form (1.5), then

$$\mathbb{E}[q(\xi - Tx)] = \sum_{i=1}^{r} \mathbb{E}[q_i(\xi_i - T_i x)]. \tag{1.6}$$

In practice we frequently encounter situations, where the probability distributions of the random variables $\xi_1, \ldots, \xi_r$ are unknown but some of their moments are known. If $\xi_1, \ldots, \xi_r$ are discrete with finite supports, and some of their moments are known, then each term in (1.6) can be bounded from both sides by the discrete moment technique. To obtain bounds for the expectation of $f(\xi)$, where ξ is a discrete random variable with finite support and known moments $\mu_k = \mathbb{E}[\xi^k], k = 1, \ldots, m$, the condition is that the function f should be a convex function of order $m + 1$. In this case we can apply the discrete moment bounding technique described in Prékopa (1990b).

If $m = 1$, then the upper bound for $\mathbb{E}\left[f(\xi)\right]$, provided by the discrete moment problem, is the same as the Edmundson-Madansky bound but the lower bound is better than Jensen's bound. In fact, any of the discrete moment bounds corresponds to an extremal distribution. If $m = 1$, then the supports of the extremal distributions have two elements, taken from the support of ξ. In case of the upper bound these are the smallest and largest possible values of ξ and the corresponding expectation is the Edmundson-Madansky bound. In case of the lower bound the extremal distribution corresponds to two consecutive possible values of ξ and the corresponding expectation is greater than or equal to $f\left(\mathbb{E}[\xi]\right)$. For more details see Prékopa (1995).

Sometimes the distributions of the random variables $\xi_1, \ldots, \xi_r$ are known but it is inconvenient to compute the exact values of the penalty function. In such cases the application of the bounding technique based on the discrete moment problems may also be advantageous.

The second stochastic programming problem is the probabilistic constrained problem that we formulate using problem (1.2) as:

$$\begin{gathered}
\min c^T x \\
\text{subject to} \\
P(Tx \geq \xi) \geq p \\
Ax = b \\
x \geq 0,
\end{gathered} \tag{1.7}$$

where p is some fixed probability, in practice near 1. One could define p as a decision variable and include a function of it in the objective function as an additive term. However for the sake of simplicity we will not do so.

The combination of the two model constructions, the hybrid model, can be obtained from (1.3) or (1.7) in such a way that we include the probabilistic constraint $P(Tx \geq \xi) \geq p$, among the constraints, and the penalty term $\mathbb{E}[q(\xi - Tx)]$ in the objective function. A simplified form of problem (1.7), where the joint probability constraint is replaced by individual constraints, i.e., one constraint for each component of Tx and ξ, was first studied by Charnes, Cooper and Symonds (1958). Miller and Wagner (1965) introduced first joint probability constraint assuming independence for the components of ξ. The general case was introduced and first studied by Prékopa (1970,1973).

The recourse and the hybrid models are of great interest, both from the theoretical and practical points of view. In this paper, however, we concentrate on the use of probability bounds in probabilistic constrained stochastic programming problems.

2 Bounding Schemes

In this section we summarize those bounding techniques that we apply in the next sections. All our bounding techniques are based on linear programming. We present univariate and multivariate binomial moment problems and the Boolean probability bounding scheme.

The univariate binomial moment problem is formulated for a finite number of events $A_1, \ldots, A_r$, defined in an arbitrary probability space. Our intention is to create sharp lower and upper bounds for the probability of a Boolean function of these events. Here we restrict ourselves to bounding $A_1 \cap \ldots \cap A_r$ and $A_1 \cup \ldots \cup A_r$.

Let ξ designate the number of events out of $A_1, \ldots, A_r$, which occur. It can be shown (see, e.g., Prékopa (1995), p.182-183) that

$$\mathbb{E}\Big[\binom{\xi}{k}\Big] = S_k, \quad k = 1, \ldots, r, \tag{2.1}$$

where

$$S_k = \sum_{1 \leq i_1 < \ldots < i_k \leq r} P(A_{i_1} \cap \ldots \cap A_{i_k}), \quad k = 1, \ldots, r.$$

In view of (2.1) we call S_k the k-th order binomial moment of the random variable ξ. We define $S_0 = 1$.

The binomial moment problem is the linear programming problem

$$\min \quad (\max) \quad \textstyle\sum_{i=0}^{r} f_i v_i \tag{2.2}$$

$$\text{subject to}$$
$$\sum_{i=0}^{r} \binom{i}{k} v_i = S_k, \quad k = 0, \ldots, m$$
$$v_i \geq 0, \quad i = 0, \ldots, r,$$

where $f_0, \ldots, f_n$ are constants and $m < r$. If

$$f_i = \begin{cases} 1, & \text{if } i \geq 1, \\ 0, & \text{if } i = 0, \end{cases} \tag{2.3}$$

then the optimal values of the problems (2.2) provide us with lower and upper bounds for $P(A_1 \cup \ldots \cup A_r)$. If

$$f_i = \begin{cases} 1, & \text{if } i = r, \\ 0, & \text{if } i < r, \end{cases} \tag{2.4}$$

then we obtain bounds for $P(A_1 \cap \ldots \cap A_r)$.

In view of the equation $P(A_1 \cap \ldots \cap A_r) = 1 - P(\bar{A}_1 \cap \ldots \cap \bar{A}_r)$, where $\bar{A}_i$ is the complementary event of A_i, a connection between the bounds obtained by the use of the functions (2.3) and (2.4) can be established. This connection tells us that the sharp lower (upper) bound for $P(A_1 \cap \ldots \cap A_r)$, using the minimization (maximization) problem (2.2) and the function (2.3) is the same as $1-$[the sharp upper (lower) bound for $P(\bar{A}_1 \cup \ldots \cup \bar{A}_r)$] using the maximization (minimization) problem (2.2) (replacing $\bar{S}_k$ for S_k) and the function (2.4).

The multivariate binomial moment problem is formulated for the sequences of events:

$$\begin{array}{c} A_{11}, \ldots, A_{1r_1}, \\ \ldots\ldots\ldots\ldots \\ A_{s1}, \ldots, A_{sr_s}. \end{array} \tag{2.5}$$

We define $\xi_1, \ldots, \xi_s$ as the numbers of events that occur in the rows $1, \ldots, s$, respectively, and introduce the notation

$$S_{\alpha_1 \ldots, \alpha_s} = \mathbb{E}\Big[\binom{\xi_1}{\alpha_1} \cdots \binom{\xi_s}{\alpha_s}\Big], \tag{2.6}$$

where $\alpha_1 \geq, \ldots, \alpha_s \geq 0$ are integers.

We formulate two types of multivariate binomial moment problems depending on the range, where $\alpha_1, \ldots \alpha_s$ are allowed to vary. The first problem is

$$\min \quad (\max) \quad \sum_{i_1=0}^{r_1} \cdots \sum_{i_s=0}^{r_s} f_{i_1 \ldots i_s} v_{i_1 \ldots i_s} \tag{2.7}$$
$$\text{subject to}$$
$$\sum_{i_1=0}^{r_1} \cdots \sum_{i_s=0}^{r_s} \binom{i_1}{\alpha_1} \cdots \binom{i_s}{\alpha_s} v_{i_1 \ldots i_s} = S_{\alpha_1 \ldots \alpha_s}$$
$$\text{for } 0 \leq \alpha_j \leq m, \quad j = 1, \ldots, s, v_{i_1 \ldots i_s} \geq 0, \text{ for all } i_1 \ldots i_s.$$

The second problem differs from (2.7) only in the condition regarding $\alpha_1, \ldots, \alpha_s$. In the second problem we assume that $\alpha_i \geq 0, \quad i = 1, \ldots, s, \ \alpha_1 + \cdots + \alpha_s \leq m$.

The optimum values of the linear programming problems (2.7) provide us with sharp lower and upper bounds for the probabilities of various Boolean functions of the above sequences of events. If, e.g.,

$$f_{i_1 \ldots i_s} = \begin{cases} 1, & \text{if } i_j \geq 1, j = 1, \ldots, s \\ 0 & \text{otherwise}, \end{cases} \tag{2.8}$$

then we obtain bounds for $P\Big(\bigcap_{i=1}^{s} \bigcup_{j=1}^{r_i} A_{ij}\Big)$.

The s event sequences (2.5) can be formed out of one event sequence $A_1, \ldots, A_r$, in order to improve on the bounds. In fact, in this case we use not only the sums $S_1, \ldots, S_m$ of the individual and joint probabilities of the events $A_1, \ldots, A_r$ but more detailed information represented by the joint binomial moments. In this case problem (2.2) (problem (2.7)) can be regarded as an aggregated (disaggregated) variant of problem (2.7) (problem (2.2)). A complete disaggregation, where each of the events $A_1, \ldots, A_r$ forms a separate sequence, takes us to the Boolean probability bounding scheme (see Boole (1854), Hailperin (1965)).

The Boolean problem can be formulated as follows. Let $A_1, \ldots, A_r$ be arbitrary events. We subdivide the sample space Ω into 2^r parts by taking

$$B_J = \Big(\cap_{i \in J} A_i\Big)\Big(\cap_{j \in \bar{J}} \bar{A}_j\Big), \tag{2.9}$$

where $J \subset \{1, \ldots, r\}, \bar{J} = \{1, \ldots, r\} - J$ and $\bar{A}_j = \Omega - A_j, \quad j = 1, \ldots, r$.

Let $v_J = P(B_J)$ and $p_I = P(\cap_{i \in I} A_i), J, I \subset \{1, \ldots, r\}$. For the case of $I = \emptyset$ we define $p_I = 1$. Further, we introduce the incidence matrix $H = (h_{IJ})$, where

$$h_{IJ} = \begin{cases} 1, & \text{if } I \subset J, \\ 0 & \text{otherwise}. \end{cases} \tag{2.10}$$

We have the equations

$$\sum_{J \subset \{1, \ldots, r\}} h_{IJ} v_J = p_I, \quad I \subset \{1, \ldots, r\}. \tag{2.11}$$

The linear programming problems

$$\begin{gathered} \min \quad (\max) \quad \textstyle\sum_{J \subset \{1, \ldots, r\}} f_J v_J \\ \text{subject to} \\ \sum_{J \subset \{1, \ldots, r\}} h_{IJ} v_J = p_I, \quad I \subset \{1, \ldots, r\}, |I| \leq m, \\ v_J \geq 0 \text{ all} \quad J \subset \{1, \ldots, r\}, \end{gathered} \tag{2.12}$$

are called the Boolean probability bounding problems. The f_J values are constants. If

$$f_J = \begin{cases} 1, & \text{if } J \neq \emptyset, \\ 0 & \text{otherwise}, \end{cases} \tag{2.13}$$

then the problems (2.12) provide us with lower and upper bounds for the probability $P(A_1 \cup \ldots \cup A_r)$. If

$$f_J = \begin{cases} 1, & \text{if } J = \{1, \ldots, r\}, \\ 0 & \text{otherwise}, \end{cases} \tag{2.14}$$

then problems (2.12) provide us with bounds for $P(A_1 \cap \ldots \cap A_r)$. These bounds are sharp in the sense that no better bounds can be given based on the knowledge of p_I, $|I| \leq m$.

If f_J is defined by (2.13) and f_i by (2.3), then problem (2.2) can be obtained from problem (2.12) by the following aggregation. We add those equations in (2.12) for which $|I| = k, (0 \leq k \leq m)$ and then introduce the new variables:

$$v_i = \sum_{|J|=i} v_J, \quad i = 0, \ldots, m.$$

If a linear programming problem is a minimization (maximization) problem, then the objective function value corresponding to any dual feasible basis is a lower (upper) bound for the optimum value. Thus, a variety of bounds can be obtained by the use of the above LP's, provided that we are able to construct dual feasible bases. For univariate problems of type (2.2), we have a good overview of the dual feasible bases for a variety of objective functions. This is provided by the basis structure theorems (see Prékopa (1988, 1990a, 1990b)). However, there are no general dual feasible basis structure theorems in the multivariate case. Still a variety of dual feasible bases can be obtained in connection with problem (2.7) (see Prékopa (1992, 1998a)).

Given a dual feasible basis for any of the above problems, we can find the corresponding bound in closed form provided that we can obtain a closed form for the inverse of the basis. Otherwise we can present algorithmic bounds, i.e., obtain the bounds by executing the dual method of linear programming.

Many probability bounds that have not been obtained in the above described linear programming framework can be obtained as objective function values corresponding to some dual feasible bases in some of the above LP's. Examples are described in Kwerel (1975), Prékopa (1988, 1990a) and Boros and Prékopa (1989). In Section 3 we show how Hunter's upper bound (Hunter (1976)) can be assigned to a dual feasible basis in the Boolean probability bounding scheme.

Hunter's bound is an upper bound for $P(A_1 \cup \ldots \cup A_n)$. We take the complete undirected graph of vertices $1, \ldots, r$, assign the weight $p_{ij} = P(A_i \cap A_j)$, $i \neq j$ to the edge (i, j), pick a spanning tree with edge set T and then obtain the bound as

$$P(A_1 \cup \ldots \cup A_r) \leq S_1 - \sum_{(i,j)\in T} p_{ij}. \tag{2.15}$$

The best bound corresponds to the heaviest spanning tree T.

3 The Use of Binomial Moment Bounds in Probabilistic Constrained Stochastic Programming Problems

In this section we present stochastic programming problems, where the probability bounding schemes are incorporated into the problems as subproblems. Our aim is to use the joint probability distribution functions of several lower dimensional random vectors, rather than that of one higher dimensional random vector. There is a way to use the probability bounds, combined with simulation, in the original probabilistic constraint (see Szántai (1988)). While this is quite efficient, the number of stochastic inequalities in the probabilistic constraint cannot be very large if we want to solve the problem in reasonable time. We hope that the methodology that we propose in this paper will significantly improve on the possibilities to solve probabilistic constrained problems. Let us introduce the notation:

$$F_{i_1\cdots i_k}(z_{i_1},\ldots,z_{i_k}) = P(\xi_{i_1} \leq z_{i_1},\ldots,\xi_{i_k} \leq z_{i_k}), \quad 1 \leq i_1 < \ldots < i_k \leq r$$

.

If A_i designates the event $\xi_i \leq z_i$, $i = 1,\ldots,r$, then we have the equalities:

$$S_k = \sum_{1\leq i_1<\ldots<i_k\leq r} F_{i_1\ldots i_k}(z_{i_1},\ldots z_{i_k}), \quad 1 \leq k \leq r.$$

The simplest bound that we can use is Boole's lower bound for the intersection of r events:

$$P(A_1 \cap \ldots \cap A_r) \geq S_1 - (r-1).$$

We may use it in problem (1.7) to replace the probabilistic constraint

$$P(Tx \geq \xi) \geq p \tag{3.1}$$

by the constraint

$$\sum_{i=1}^{r} F_i(T_i x) - (r-1) \geq p. \tag{3.2}$$

The advantage of constraint (3.2) is that in many cases the function on the left hand side in (3.2) is concave. We only need to ensure that the function $F_i(z_i)$ is concave in the range where $T_i x$ takes its values, $i = 1, \ldots, r$. On the other hand, since Boole's lower bound is not very tight, the use of the constraint (3.2) may significantly increase the optimum value of problem (1.7).

It is preferable, instead of (3.2), to use the constraints

$$\begin{gathered} P(T_i x \geq \xi_i) \geq p_i, \quad i = 1, \ldots, r \\ \sum_{i=1}^{r}(1 - p_i) \leq 1 - p. \end{gathered} \tag{3.3}$$

Relations (3.3) imply that $P(Tx \geq \xi) \geq p$. In fact,

$$\begin{aligned} P(Tx \geq \xi) &= P(T_i x \geq \xi_i,\, i = 1, \ldots, r) \\ &= 1 - P(\{T_1 x < \xi_1\} \cup \ldots \cup \{T_r x < \xi_r\}) \\ &\geq 1 - \sum_{i=1}^{r} P(T_i x < \xi_i) = 1 - \sum_{i=1}^{r}(1 - P(T_i x \geq \xi_i)) \\ &\geq 1 - \sum_{i=1}^{r}(1 - p_i) \geq p. \end{aligned} \tag{3.4}$$

In (3.3) the $p_1, \ldots, p_r$ are variables while p is a constant, $0 < p < 1$. The constraints (3.3) ensure that $p \leq p_i \leq 1$. Then problem (1.7) is a convex problem provided that $F_i(z)$ is concave for $z \geq F_i^{-1}(p)$, $i = 1, \ldots, r$ and the other constraints of the problem guarantee that $T_i x \geq F_i^{-1}(p)$, $i = 1, \ldots, r$.

Now we look for such replacement of the probabilistic constraint in problem (1.7) that use joint distributions of the random variables $\xi_1, \ldots, \xi_r$. The optimum values of the linear programming problems

$$\begin{gathered} \min \quad (\max) \quad v_r \\ \text{subject to} \\ v_0 + v_1 + v_2 + v_3 + \cdots + v_r = 1 \\ v_1 + 2v_2 + 3v_3 + \cdots + rv_r = \sum_{i=1}^{r} F_i(z_i) \\ v_2 + \binom{3}{2} v_3 + \cdots + \binom{r}{2} v_r = \sum_{1 \leq i_1 < i_2 \leq r} F_{i_1 i_2}(z_{i_1}, z_{i_2}) \\ \cdots\cdots\cdots\cdots \\ v_m + \binom{m+1}{m} v_{m+1} + \cdots + \binom{r}{m} = \sum_{1 \leq i_1 < \ldots < i_m \leq r} F_{i_1 \ldots i_m}(z_{i_1}, \ldots, z_{i_m}) \\ v_0 \geq 0, v_1 \geq 0, v_2 \geq 0, \ldots, v_r \geq 0 \end{gathered} \tag{3.5}$$

provide us with lower and upper bounds for $F(z_1, \ldots, z_n)$. Using these , we can formulate two approximate problems for problem (1.7). The first one uses the minimum problem (3.5) and is the following:

$$\min\{c^T x + \alpha v_r\} \tag{3.6}$$

$$\text{subject to}$$
$$v_0 + v_1 + v_2 + v_3 + \cdots + v_r = 1$$
$$v_1 + 2v_2 + 3v_3 + \cdots + rv_r = \sum_{i=1}^{r} F_i(T_i x)$$
$$v_2 + \binom{3}{2}v_3 + \cdots + \binom{r}{2}v_r = \sum_{1\le i_1 < i_2 \le r} F_{i_1 i_2}(T_{i_1}x, T_{i_2}x)$$
$$\cdots\cdots\cdots\cdots$$
$$v_m + \binom{m+1}{m}v_{m+1} + \cdots + \binom{r}{m} = \sum_{1\le i_1 < \ldots < i_m \le r} F_{i_1 \ldots i_m}(T_{i_1}x, \ldots, T_{i_m}x)$$
$$Ax = b$$
$$v_0 \ge 0, v_1 \ge 0, v_2 \ge 0, \ldots, v_{r-1} \ge 0, v_r \ge p, x \ge 0,$$

where α is an arbitrary nonnegative number. The second one uses the maximum problem (3.5) and is the following

$$\max\{-c^T x + \alpha v_r\} \tag{3.7}$$
$$\text{subject to}$$
$$v_0 + v_1 + v_2 + v_3 + \cdots + v_r = 1$$
$$v_1 + 2v_2 + 3v_3 + \cdots + rv_r = \sum_{i=1}^{r} F_i(T_i)x$$
$$v_2 + \binom{3}{2}v_3 + \cdots + \binom{r}{2}v_r = \sum_{1\le i_1 < i_2 \le r} F_{i_1 i_2}(T_{i_1}x, T_{i_2}x)$$
$$\cdots\cdots\cdots\cdots$$
$$v_m + \binom{m+1}{m}v_{m+1} + \cdots + \binom{r}{m} = \sum_{1\le i_1 < \ldots < i_m \le r} F_{i_1, \ldots i_m}(T_{i_1}x, \ldots T_{i_m}x)$$
$$Ax = b$$
$$v_0 \ge 0, v_1 \ge 0, v_2 \ge 0, \ldots, v_{r-1} \ge 0, v_r \ge p, x \ge 0,$$

where α is some fixed positive number.

We show that under some circumstances a relaxed version of problem (3.7) is a convex programming problem. We look at the case m=2 and relax problem (3.7) in such a way that we replace $\le$ for $=$ in the second constraint. Then the probability bounding subproblem takes the form:

$$\max v_r \tag{3.8}$$
$$\text{subject to}$$
$$v_0 + v_1 + v_2 + v_3 + \cdots + v_r = 1$$
$$v_1 + 2v_2 + 3v_3 + \cdots + rv_r \le S_1$$
$$v_2 + \binom{3}{2}v_3 + \cdots + \binom{r}{2}v_r = S_2$$
$$v_1 \ge 0, v_2 \ge 0, \ldots, v_r \ge 0.$$

The optimum value of this problem is equal to $S_2/\binom{r}{2}$. This is a concave function of x provided that all functions $F_{i_1 i_2}(z_{i_1}, z_{i_2})$ are concave. If it holds, then (3.7) is a convex problem. In Section 5 we show that the above distribution functions are in fact concave, at least in some cases.

4 The Use of the Boolean Probability Bounding Scheme in Probabilistic Constrained Stochastic Programming Problem

The Boolean probability bounding scheme provides us with better bounds than the binomial moment problems. If we use it to obtain a lower bound for the probability in the probabilistic constraint, then we formulate the problem:

$$\begin{gathered}\min\{c^T x + \alpha v_N\} \\ \text{subject to} \\ \sum_{J \subset N} h_{IJ} v_J = F_I(T_I x), \quad I \subset N, |I| \leq m \\ Ax = b \\ v_J \geq 0, J \subset N, \quad J \neq N, \quad v_N \geq p,\, x \geq 0,\end{gathered} \tag{4.1}$$

where α is a fixed nonnegative number, $N = \{1, \ldots, r\}$, $I \subset N$, $T_I = (T_i,\, i \in I)$, $z_I = (z_I,\, i \in I)$, $F_I(z_I) = P(\xi_i \leq z_i,\, i \in I)$.

In a similar way we can replace the constraining function in the probabilistic constraint by its upper bound, if we use the optimum value of the maximization problem in (2.12) with objective function (2.14). Then we formulate the problem:

$$\begin{gathered}\max\{-c^T x + \alpha v_N\} \\ \text{subject to} \\ \sum_{J \subset N} h_{IJ} v_J = F_I(T_I x), \quad I \subset N, |I| \leq m \\ Ax = b \\ v_J \geq 0, J \subset N, \quad J \neq N, \quad v_N \geq p,\, x \geq 0,\end{gathered} \tag{4.2}$$

where α is a fixed positive number.

Let $x_{opt}, x^l_{aggr}, x^l_{Boole}, x^u_{aggr}, x^u_{Boole}$ designate the x parts of the optimal solutions to the problems (1.7), (3.6), (4.1) and the maximization counterparts of (3.6), (4.1), respectively. Then, if α is small, we have the approximate inequalities:

$$c^T x^u_{aggr} \leq c^T x^u_{Boole} \leq c^T x_{opt} \leq c^T x^l_{Boole} \leq c^T x^l_{aggr} \tag{4.3}$$

The first two inequalities hold exactly if α is chosen equal to 0. We proceed with the discussion by establishing a connection between Hunter's (1976) bound and the Boolean probability bounding scheme. We choose the coefficient vector of the objective function equal to that given by (2.13). However, we use a modified version of the Boolean problem. We remove the variable corresponding to $J = \emptyset$ and the constraint corresponding to $I = \emptyset$.

Theorem 4.1 *Hunter's upper bound is the objective function value corresponding to a dual feasible basis of the following LP:*

$$\max \sum_{\emptyset \neq J \subset N} v_J \tag{4.4}$$
$$\textit{subject to}$$
$$\sum_{\emptyset \neq J \subset N} h_{IJ} v_J = p_I, \quad \emptyset \neq I \subset N, |I| \leq 2$$
$$v_J \geq 0, \ \textit{all } J.$$

The dual feasible basis can be chosen as the collection of those columns of the matrix H of the equality constraints in (4.4) which are labelled by $\{1\}, \ldots, \{r\}$ and those J which are vertex sets in any paths that exist in T.

Proof of Theorem 4.1: The number of equality constraints in problem (4.4) is $r + \binom{r}{2}$, the same as the collection of vectors we choose to create a suitable dual feasible basis. The proof has three steps. Let B designate the matrix of the selected columns.

(1) The collection of the selected vectors are linearly independent. To show this let us subdivide the rows of the matrix (h_{IJ}) into two blocks, according as $|I| = 1$ or $|I| = 2$. Assume that in H the first r columns and rows are labeled by $\{1\}, \ldots, \{r\}$, respectively. Then the first r columns of B are the first r of the $r + \binom{r}{2}$ - component unit vectors. Thus, to prove that B is non-singular it is enough to prove that if we remove the first r rows and columns from B, we obtain a nonsingular matrix. To show this we first arrange the rows in the second block of H in such a way that first come those labeled by pairs $I = \{i, j\}$ which are at the same time edges of T; then come those labeled by $I = \{k, l\}$ for which k and l are connected by a path of length 2 etc.

After that we arrange the columns in such a way that first come the same $r - 1$ pairs, in the same order, that label the first $r - 1$ rows; the further ordering of the columns follows the ordering of the remaining rows: the tth column should have that label J which consists of the nodes sitting in the path connecting the pair of nodes that constitute the label of the tth row I. The obtained matrix is an upper triangular matrix with all 1's in the main diagonal. Hence, it is nonsingular.

(2) If $y = (y_1, \ldots, y_r, y_{12}, y_{13}, \ldots, y_{r-1r})^T$ designates the solution of the linear equation

$$y^T B = \mathbf{1}^T, \tag{4.5}$$

where $\mathbf{1}$ is the $r + \binom{r}{2}$-component vector with all 1's as components, then

$$\begin{array}{ll} y_i = 1, & \text{for } i = 1, \ldots, r \\ y_{ij} = -1, & \text{for } i = \{i, j\} \in T \\ y_{ij} = 0, & \text{for } i = \{i, j\} \notin T. \end{array}$$

We only have to check that y satisfies (4.5). The first r equations in (4.5) that correspond to the first r columns in B, are satisfied trivially. If we pick a column of B, the label of which is J and $|J| = k$, then there are $k-1$ edges in that path the vertex set of which is J. These $k-1$ edges are all in T and they determine $k-1$ row labels I such that the corresponding rows (and no other rows) have a 1 in that column. Thus, equation (4.5) is satisfied for that column.

(3) For any column $h_{\circ J}$ of H which is not in B

$$y^T h_{\circ J} \geq 1. \tag{4.6}$$

Now, J is not the vertex set of a path. This implies that if $|J| = k$, then there are at most $k-1$ edges in T such that the vertices incident to them are in J. Thus, (4.6) holds and the proof of the theorem is complete. $\square$.

Theorem 4.1 can be used in problem (4.1) in the following way. We write up problem (4.4) for the events: $\xi_i > T_i x$, $i = 1, \ldots, r$. Then the optimum value of problem (4.4) provides us with an upper bound for the probability of the union, or what is the same, $1 - P(Tx \geq \xi)$. Thus, $1-$(optimum value of the problem (4.4)) is a lower bound for $P(Tx \geq \xi)$. Hence, if we remove the variable v_J for $J = \emptyset$ and the constraint corresponding to $I = \emptyset$ from the probability bounding subproblem in (4.1), then we may replace v_N in problem (4.1) by

$$1 - \sum_{\emptyset \neq J \subset N} v_J. \tag{4.7}$$

Since the prescribed probability level p (in the probabilistic constraint) is usually large and the binomial as well as the Boolean probability bounding schemes perform quite well in such cases, we can expect similar good performance from the approximate problems. If m is increased in the bounding schemes, then the lower and upper bounds become closer to each other and to the true probability value.

When we solve problem (4.1), then we may do it in such a way that at any iteration, with respect to the variables in x, we fully solve the problem with respect to the v variables or we handle the two groups of variables simultaneously. In the first case we always have a good initial dual feasible basis, provided by Theorem 4.1, and the dual method can be recommended to solve the problems for the v variables. In the second case it is advantageous to use the dual method simoltaneously for all variables, for the same reason.

If the objective function in problem (4.1) is extended by penalty terms, stemming from the simple recourse problem (i.e., we are dealing with the hybrid problem), then again the dual method is very suitable for the solution of the overall problem. This is because the simple recourse problem has an efficient solution based on the dual method (see Prékopa (1990c,1995), Ruff-Fiedler, Prékopa, Fábián (1998)). Hansen, Jaumard and Nguesté (1998) developed an efficient column generation metod for the solution of problem (4.4) that can also be utilized here.

5 Convexity of the Approximate Problem

The convexity of a large class of probabilistic constrained stochastic programming problems was established by Prékopa, by the use of logconcave measures. (For detailed description of the relevant results see Prékopa (1995).) If the random vector ξ has logconcave probability distribution, then the constraining function in the constraint $P(Tx \geq \xi) \geq p$ is logconcave, which imlies that the set of x vectors satisfying this constraint is convex.

The difficulty in using joint probabilistic constraint is computational. In the course of an algorithmic solution of the problem we have to compute function and gradient values of $P(Tx \geq \xi)$, which is computationally intensive. In order to alleviate this difficulty we formulated the problems in Sections 3 and 4. In those problems we encounter two kinds of difficulties. The first one is that on the right hand sides of the problems of Section 3 we have sums of probability distribution functions. If ξ has logconcave probability density function, then the distribution of ξ and all its marginal distributions are logconcave (Prékopa (1973)). However, the logconcavity property does not carry over for sums, thus we cannot guarantee that the functions standing on the left hand sides of the problems are logconcave or quasi-concave. We can overcome this difficulty in some special cases. The second difficulty in problems of Sections 3 and 4 is that we have equality constraints involving nonlinear functions. We offer some analysis in this respect, too. A theorem concerning the multivariate normal distribution follows.

Theorem 5.1 *The standard multivariate normal probability distribution function* $\Phi(z_1, \ldots, z_r; R)$ *is concave in the set* $\{z|\, z_i \geq \sqrt{r-1},\ i = 1, \ldots, r\}$.

Proof: We present the proof for the case of $r = 2$. The proof for the general case is presented in Prékopa (1998b). In order to simplify the notation the variables of the distribution function are designated by x and y. The function is then $\Phi(x, y; \rho)$. If $\rho = 1$, then we have the equation

$$\Phi(x, y; \rho) = \Phi(\min(x, y)),$$

where Φ is the univariate standard normal distribution function. If $\rho = -1$, then

$$\Phi(x, y; \rho) = \Phi(x) + \Phi(y) - 1.$$

Since $\Phi(z)$ is concave for $z \geq 0$, in both cases $\Phi(x, y; \rho)$ is concave in the set $\{(x, y)|\, x \geq 0, y \geq 0\}$.

Let $|\rho| < 1$. Since $\Phi(x, y; \rho)$ is a continuous function, it is enough to prove that if $x_1, x_2, y_1, y_2 \geq 1$, then

$$\Phi(\frac{1}{2}x_1 + \frac{1}{2}x_2, \frac{1}{2}y_1 + \frac{1}{2}y_2; \rho) \geq \frac{1}{2}\Phi(x_1, y_1; \rho) + \frac{1}{2}\Phi(x_2, y_2; \rho). \tag{5.1}$$

In addition, it is enough to prove (5.1) for the case of $1 \leq x_1 \leq x_2$, $1 \leq y_1 \leq y_2$, for the following reason. If (5.1) holds for this case, then as a special case of it we obtain that

$\Phi(x, y; \rho)$ is concave in each variable in the set $\{(x, y) | \, x \geq 1, y \geq 1\}$. (A stronger result in this respect, presented in Prékopa (1970), is the following: $\Phi(x, y; \rho)$ is concave in each variable for $x \geq 0, \; y \geq 0$, if $\rho \geq 0$ and for $x \geq \alpha, \; y \geq \alpha$ for $\rho < 0$, where

$$\alpha = \sqrt{\frac{\varphi(1)}{2\Phi(1) + \varphi(1)}} = 0.3546$$

and φ is the univariate standard normal probability density function.) The second observation is that if $1 \leq x_1 \leq x_2$ and $1 \leq y_2 \leq y_1$, then since $\Phi(x, y; \rho)$ is a bivariate probability distribution function, we have the inequality:

$$\Phi(x_2, y_1; \rho) + \Phi(x_1, y_2; \rho) - \Phi(x_1, y_1; \rho) - \Phi(x_2, y_2; \rho) \geq 0. \tag{5.2}$$

Hence we can derive (5.1) also for these points $(x_1, y_1), (x_2, y_2)$:

$$\begin{aligned} & \Phi(\frac{1}{2}x_1 + \frac{1}{2}x_2, \frac{1}{2}y_1 + \frac{1}{2}y_2; \rho) \geq \\ \geq \; & \frac{1}{4}\Big(\Phi(x_1, y_1; \rho) + \Phi(x_1, y_2; \rho) + \Phi(x_2, y_1; \rho) + \Phi(x_2, y_2; \rho)\Big) \\ \geq \; & \frac{1}{2}\Phi(x_1, y_1; \rho) + \frac{1}{2}\Phi(x_2, y_2; \rho). \end{aligned} \tag{5.3}$$

Let ξ_1, ξ_2 be two random variables such that

$$P(\xi_1 \leq x, \xi_2 \leq y) = \Phi(x, y; \rho).$$

We may represent ξ_1, ξ_2 in the form:

$$\begin{aligned} \xi_1 &= \eta_1 \\ \xi_2 &= \rho\eta_1 + \sqrt{1 - \rho^2}\,\eta_2, \end{aligned} \tag{5.4}$$

where η_1, η_2 are independent, $N(0, 1)$-distributed random variables. The hyperplanes:

$$\begin{aligned} & x = 1 \\ & \rho x + \sqrt{1 - \rho^2}\, y = 1 \end{aligned}$$

have unit-length normal vectors $(1, 0), (\rho, \sqrt{1 - \rho^2})$, respectively, and are tangent to the unit circle

$$\{(x, y) \,|\, x^2 + y^2 = 1\}.$$

Using the joint density function of η_1, η_2 we may write $\Phi(x, y; \rho)$ in the form:

$$\Phi(x, y; \rho) \;=\; \iint\limits_{\substack{u \leq x \\ \rho u + \sqrt{1-\rho^2} v \leq y}} \frac{1}{2\pi} e^{-\frac{1}{2}(u^2 + v^2)} du\, dv. \tag{5.5}$$

We want to derive (5.1) for $x_2 \geq x_1 \geq 1, y_2 \geq y_1 \geq 1$. Let $x_3 = \frac{1}{2}(x_1 + x_2)$, $y_3 = \frac{1}{2}(y_1 + y_2)$,

$$D(x, y) = \{(u, v)|,\ u \leq x, \rho u + \sqrt{1 - \rho^2}v \leq y\}$$

$$H_1 = D(x_3, y_3) \setminus D(x_1, y_1), \quad H_2 = D(x_2, y_2) \setminus D(x_3, y_3).$$

Using this notation (5.1) is equivalent to

$$\iint\limits_{H_1} \frac{1}{2\pi} e^{-\frac{1}{2}(u^2+v^2)} du\, dv \quad \geq \quad \iint\limits_{H_2} \frac{1}{2\pi} e^{-\frac{1}{2}(u^2+v^2)} du\, dv \tag{5.6}$$

Introducing polar coordinates $u = w \cos\psi,\ v = w \sin\psi$, the determinant of the Jacobian is w and the integrand in (5.6) equals

$$\frac{1}{2\pi} w e^{-\frac{w^2}{2}},\ w \geq 0, -\pi \leq \psi < \pi. \tag{5.7}$$

The function (5.7) is increasing if $0 \leq w \leq 1$ and is decreasing if $w \geq 1$, for any fixed ψ. The sets H_1, H_2 are disjoint and do not intersect the interior of the unit circle. Hence, the function (5.7) is decreasing in the sets along any ray that start at the origin.

Let l_1 and l_2 be the lengths of the intersections of the ray ψ =const. with the sets H_1 and H_2, respectively. These intersections are simultaneously empty or have positive lenghts. Figure 1 shows that we always have the inequality $l_1 \geq l_2$. More precisely, if ψ_1 is the smallest and ψ_2 is the largest of the angles of the lines connecting the origin and the points $(x_i, \frac{y_i - \rho x_i}{\sqrt{1-\rho^2}})$, $i = 1, 2, 3$, then $l_1 = l_2$ if $\psi \leq \psi_1$ or $\psi \geq \psi_2$, and $l_1 > l_2$ if $\psi_1 < \psi < \psi_2$. This implies that the integral of the function (5.7) along the intersection of the ray $w \geq 0, \psi$ =const. and H_1 is greater than or equal to the integral of the function along the intersection of the ray and H_2. Integrating with respect to ψ, the inequality follows. □

The proof shows that $\Phi(x, y; \rho)$ is also strictly concave in $\{(x, y)|\, x \geq 1, y \geq 1\}$.

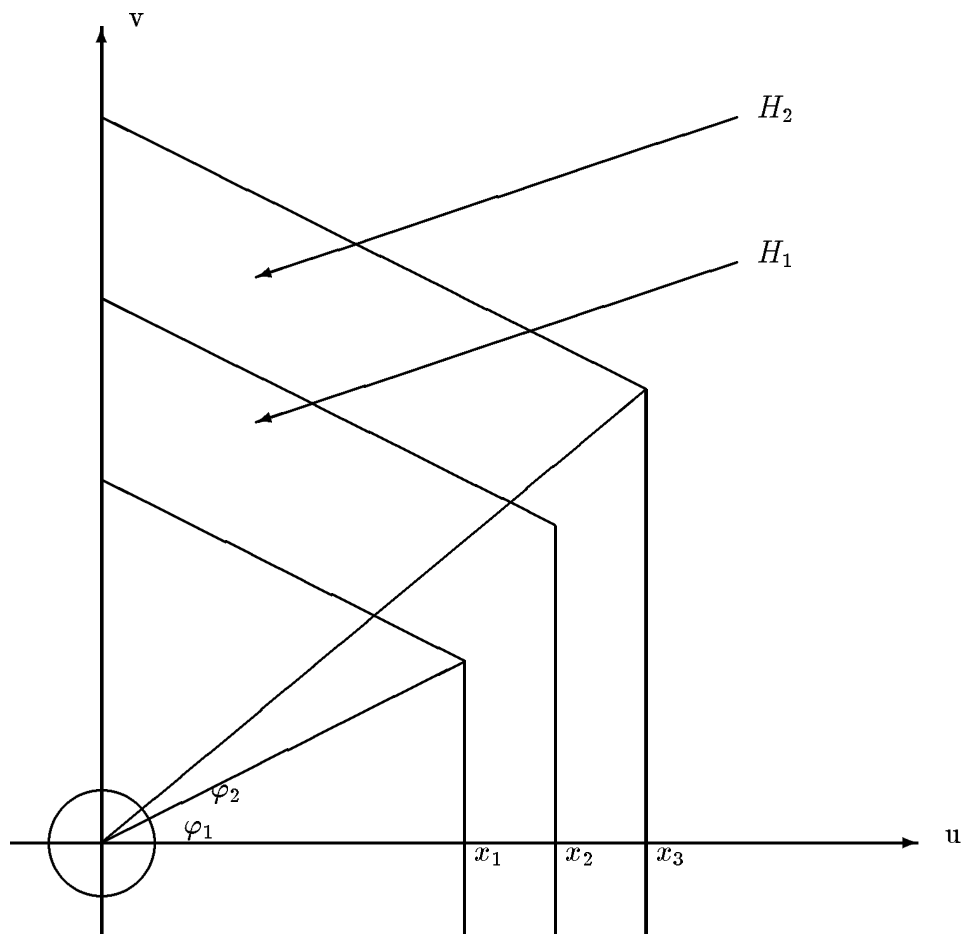

Figure 5.1

We intend to use the problems of Sections 3 and 4 for the case of small r values. In these cases the limitation $z_i \geq \sqrt{r-1}$, $i = 1, \dots, r$ is not very restrictive. In fact, if the random variables are independent, then the value of their joint distribution at $z_i = \sqrt{r-1}$, $i = 1, \dots, r$ is $a_r = \Phi^r(\sqrt{r-1})$ and $a_2 = 0.71, a_3 = 0.78, a_4 = 0.84, a_5 = 0.89, a_6 = 0.93$. For large probabilities the binomial moment probability bounds are reasonably good if $m = 3$, which means that at most 3-variate probability distributions appear in the problems. On the other hand, numerical results show that the Boolean probability bounding schemes are quite accurate in case of $m = 3$. We have obtained concavity results for other probability distributions too, such as the multivariate Dirichlet, gamma and uniform distributions.

Despite of these favorable results, the problems formulated in Sections 3 and 4 are, in general, not convex problems, even if concave probability distribution functions are used. The reason is that in problem (3.5) the objective function in the minimization (maximization) problem is not a decreasing (increasing) function of the right hand side values.

We may overcome this difficulty if we relax problem (3.5) by changing some of the equalities into inequalities. The background for that is provided by a property of the dual variables $y_0, \ldots, y_m$ corresponding to any dual feasible basis in problem (3.5): these variables have alternating signs starting with $-$ $(+)$ in case of the minimization (maximization) problem. In agreement with this, we may change every odd (even) numbered equality in (3.5) into $\leq$ $(\geq)$. Numerical experience will tell us how far the new probability bounds go from the true probability by these relaxations.

If m is small then we have formulas for the probability bounds. Still to keep the LP's rather than to work with the formulas seems to be advantageous, especially if we design algorithms for the solutions of the problems.

We close our discussion by looking at the maximization problem (3.5) in case of $m = 2$. If a_i designates the i-th column of the matrix of the equality constraints, $i = 0, \ldots, r$, then all dual feasible bases are of the form (see Prékopa (1995, Section 5.9)): $B = (a_i, a_{i+1}, a_r)$. The dual vector $y = (y_0, y_1, y_2)^T$, corresponding to B, can be computed from the equation

$$y^T B = (0, 0, 1)$$

and the result is

$$y_0 = \frac{i(i+1)}{(r-i)(r-i-1)}, \quad y_1 = \frac{-2i}{(r-i)(r-i-1)}, \quad y_2 = \frac{2}{(r-i)(r-i-1)}.$$

Thus, the objective function value corresponding to this basis is

$$\frac{i(i+1) + 2(S_2 - iS_1)}{(r-i)(r-i-1)}. \tag{5.8}$$

The smallest value (5.8), i.e., the best upper bound for $P(A_1 \cap \ldots \cap A_r)$, corresponds to

$$i_{\min} = \left\lceil \frac{(r-1)S_1 - 2S_2}{r - S_1} \right\rceil. \tag{5.9}$$

If $i \leq i_{\min}$, then the sequence (5.8) is decreasing and it is increasing if $i \geq i_{\min}$. Using this, we can write the approximate stochastic programming problem in the following form:

$$\begin{gathered} \min c^T x \\ \text{subject to} \\ Ax = b, x \geq 0 \\ \tfrac{1}{(r-i)(r-i-1)}\Big[i(i+1) + 2\textstyle\sum_{1 \leq i < j \leq r} F_{ij}(T_i x, T_j x) - \sum_{i=1}^{r} F_i(T_i x)\Big] \geq p, \\ i = 0, 1, \ldots, r-2. \end{gathered} \tag{5.10}$$

References

[1] Beale, E.M.L. (1955). On Minimizing a Convex Function Subject to Linear Inequalities. *Journal of the Royal Statistical Society, B* **47**, 173-184.

[2] Bonferroni, C.E. (1937). *Teoria Statistica delle Classi e Calcolo delle Probabilità.* Volume in onor di Riccardo Dalla Volta, Università di Firenze, 1-62.

[3] Boole, G. (1954). *Laws of Thought* (American reprint of 1854 edition). Dover, New York.

[4] Boros, E., and A. Prékopa (1989). Closed Form Two-Sided Bounds for Probabilities that Exactly r and at Least r out of n Events Occur. *Mathematics of Operations Research* **14**, 317-342.

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

[6] Dantzig, G.B. (1955). Linear Programming under Uncertainty. *Management Science* **1**, 197-206.

[7] Dawson, D., and A. Sankoff (1967). An Inequality for Probabilities. *Proceedings of the American Mathematical Society* **18**, 504-507.

[8] Hailperin, Th. (1965). Best Possible Inequalities for the Probability of a Logical Function of Events. *The American Mathematical Monthly* **72**, 343-359.

[9] Hansen, P., B. Jaumard, and G.-B.D. Nguesté (1998). Best Second Order Bounds for Two-Terminal Network Reliability with Dependent Edge Failures. To appear in *Discrete Applied Mathematics*, to appear.

[10] Hunter, D. (1976). Bounds for the Probability of a Union. *Journal of Applied Probability* **13**, 597-603.

[11] Kwerel, S.M. (1975). Most Stringent Bounds on Aggregated Probabilities of Partially Specified Dependent Probability Systems. *Journal of the American Statistical Association* **70**, 472-448.

[12] Miller, L.B., and H. Wagner (1965). Chance-Constrained Programming with Joint Constraints. *Operations Research* **13**, 930-945.

[13] Popoviciu, T. (1945). Les Fonctions Convexes. In: *Actualités Scientifiques et Industrielles* **992**, Hermann, Paris.

[14] Prékopa, A. (1970). On Probabilistic Constrained Programming. Proceedings of the Princeton Symposium on Mathematical Programming. Princeton University Press, 113-138.

[15] Prékopa, A. (1973). Contributions to the Theory of Stochastic Programming. *Mathematical Programming* **4**, 202-221.

[16] Prékopa, A. (1988). Boole-Bonferroni Inequalities and Linear Programming. *Operations Research* **36**, 145-162.

[17] Prékopa, A. (1990a). Sharp Bounds on Probabilities Using Linear Programming. *Operations Research* **38**, 227-239.

[18] Prékopa, A. (1990b). The Discrete Moment Problem and Linear Programming. *Discrete Applied Mathematics* **27**, 235-254.

[19] Prékopa, A. (1990c). Dual Method for the Solution of a One-Stage Stochastic Programming Problem with Random RHS Obeying a Discrete Probability Distribution. *ZOR-Methods and Models of Operations Research* **34**, 441-461.

[20] Prékopa, A. (1992). Inequalities on Expectations Based on the Knowledge of Multivariate Moments. In: *Stochastic Inequalities* (M.Shaked and Y.L. Tong, eds.), Institute of Mathematical Statistics, Lecture Notes - Monograph Series **22**, 309-331.

[21] Prékopa, A. (1995). Stochastic Programming. Kluwer Scientific Publishers.

[22] Prékopa, A. (1998a). Bounds on Probabilities and Expectations Using Multivariate Moments of Discrete Distributions. *Studia Sci. Math. Hungar.* **34**, 349-380.

[23] Prékopa, A. (1998b). On the Concavity of Multivariate Probability Distributions. RUTCOR Research Report, to appear.

[24] Ruff-Fiedler, O., A. Prékopa, and C.I. Fábián, (1998). On a Dual Method for a Specially Structured Linear Programming Problem. Mathematical Methods of Operations Research, to appear.

[25] Szántai, T. (1988). A computer Code for Solution of Probabilistic Constrained Stochastic Programming Problems. In: *Numerical Techniques for Stochastic Optimization* (Yu. Ermoliev and R.J.-B. Wets, eds.), Springer-Verlag, New York, 229-235.

[26] Wets, R.J.-B. (1966). Programming under Uncertainty: The Complete Problem. *Zeitschrift für Wahrsch.* **4**, 316-339.