RUTCOR
RESEARCH
REPORT

Discrete Higher Order Convex Functions and Their Applications

András Prékopa[a]

RRR 39-2000, August, 2000

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 University, 640 Bartholomew Road, Piscataway, NJ 08854-8003, prekopa@rutcor.rutgers.edu

RUTCOR RESEARCH REPORT
RRR 39-2000, AUGUST, 2000

DISCRETE HIGHER ORDER CONVEX FUNCTIONS AND THEIR APPLICATIONS

András Prékopa

Abstract. In this paper we present an overview about the recently developed theory of discrete moment problems, i.e., moment problems where the supports of the random variables involved are discrete. We look for the minimum or maximum of a linear functional acting on an unknown probability distribution subject to a finite number of moment constraints. Using linear programming methodology, we present structural theorems, in both the univariate and multivariate cases, for the dual feasible bases and show how the relevant problems can be solved by suitable adaptations of the dual method. The condition on the objective function is a kind of higher order convexity, expressed in terms of divided differences. A variant of the above, the discrete binomial moment problem, as well as generalization for discrete variable Chebyshev systems are also discussed. Finally, we present novel applications to valuations of finiancial instruments.

1 Introduction

Discrete moment problems came to prominence by the discovery (Samuels and Studden (1989), Prékopa (1988,1990a,b) that the sharp Bonferroni bounds can be obtained as optimum values of discrete moment problems.

The first sharp bound for the probability of the union of n events was obtained by Dawson and Sankoff (1967). They assumed the knowledge of the first two binomial moments of the number of events which occur. Kwerel (1975a,b) reformulated and extended the problem using linear programming methodology, where the first three binomial moments of the occurencies are supposed to be known. The general problems, where there is no limitation regarding the numbers of binomial moments have been formulated and studied by Prékopa (1988,1990a).

The multiveriate discrete moment problem has been introduced and studied by Prékopa (1992,1999).

By the use of the methodology of discrete moment problems a number of other known probability bounds could be derived (see the above cited papers by Samuels and Studden,Prékopa, furthermore the paper by Boros and Prékopa (1989)).

While Samuels and Studden follow the guidelines of the general moment problems in their discussion, Prékopa uses linear programming methodology, which enables him to come up with simple algorithmic solution to the problem wherever the size of the problem prevents us to present the bounds in closed forms.

The simplest discrete moment problem, where power moments are used, is closely connected with divided differences, higher order convex functions (defined in terms of divided differences) and Lagrange interpolation.

Let $f(z),\ z \in Z = \{z_0, \ldots, z_n\}$ be a discrete variable function, where $z_0 < \ldots < z_n$. Its first order divided differences are designated and defined as:

$$[z_{i_1}, z_{i_2}; f] = \frac{f(z_{i_2}) - f(z_{i_1})}{z_{i_2} - z_{i_1}},\ i_1 \neq i_2. \tag{1}$$

The kth order divided difference is designated and defined as:

$$[z_{i_1}, z_{i_2}, \ldots, z_{i_{k+1}}; f] = \frac{[z_{i_2}, \ldots, z_{i_{k+1}}; f] - [z_{i_1}, \ldots, z_{i_k}; f]}{z_{i_{k+1}} - z_{i_1}}, \tag{2}$$

where $z_{i_1}, z_{i_2}, \ldots, z_{i_{k+1}}$ are $k+1$ distinct elements of Z. Let $[z; f] = f(z)$, by definition.

The function f is said to be (strictly) convex of order k on Z, if all of its kth order divided differences are (positive) nonnegative.

A sufficient condition for that is the following: f is defined in $[z_0, z_n]$ and has (positive) nonnegative kth order derivatives in (z_0, z_n).

First order convexity means that the function is nondecreasing, second order convexity means that the function, obtained by connecting all neighbors of the points $(z_i, f(z_i)), i = 0, 1, \ldots, n$ by straight lines, is convex in the classical sense.

We have the determinental formula

$$[z_{i_1}, z_{i_2}, \ldots, z_{i_{k+1}}; f] = \frac{\begin{vmatrix} 1 & 1 & \ldots & 1 \\ z_{i_1} & z_{i_2} & \ldots & z_{i_{k+1}} \\ \vdots & \vdots & & \vdots \\ z_{i_1}^{k-1} & z_{i_2}^{k-1} & \ldots & z_{i_{k+1}}^{k-1} \\ f(z_{i_1}) & f(z_{i_2}) & \ldots & f(z_{i_{k+1}}) \end{vmatrix}}{\begin{vmatrix} 1 & 1 & \ldots & 1 \\ z_{i_1} & z_{i_2} & \ldots & z_{i_{k+1}} \\ \vdots & \vdots & & \vdots \\ z_{i_1}^{k-1} & z_{i_2}^{k-1} & \ldots & z_{i_{k+1}}^{k-1} \\ z_{i_1}^{k} & z_{i_2}^{k} & \ldots & z_{i_{k+1}}^{k} \end{vmatrix}}. \tag{3}$$

A theorem (see, e.g., Prékopa (1995)) asserts that if all kth order divided differences of f, corresponding to consecutive points, are (positive) nonnegative, then all kth oder divided fifferences of f are (positive) nonnegative. This fact is also a simple consequence of a general theorem of Fekete (1912). Let $I \subset \{0, 1, \ldots, n\}$ with $|I| = m+1 \quad (m \le n)$. The Lagrange polynomial corresponding to the set of points $\{z_i, i \in I\}$ is defined as

$$L_I(z) = \sum_{h \in I} f(z_h) L_{I,h}(z), \tag{4}$$

where

$$L_{I,h}(z) = \prod_{i \in I \setminus \{h\}} \frac{z - z_h}{z_i - z_h}, \quad h \in I \tag{5}$$

is the hth fundamental polynomial. We will need $L_I(z)$ at the points in Z but it can be defined for all real z values. The polynomial can be written in Newton's form:

$$L_I(z) = \sum_{h=0}^{m} [z_j,\, j \in I^{h-1}; f] \prod_{i \in I^{h-1}} (z - z_i), \tag{6}$$

where I^h is the set of the first $h+1$ points in I. and $\prod_{i\in I^{-1}}(z-z_i)=1$ by definition. An important formula in Lagrange interpolation is the following:

$$f(z)-L_I(z)=[z_j, j\in I,\ z; f]\prod_{i\in I}(z-z_i). \tag{7}$$

In the s-variate case first we define $Z_j=\{z_{j0},\ldots,z_{jn_j}\},\ z_{j0}<\ldots<z_{jn_j},\ j=1,\ldots,s$. Let $Z=Z_1\times\cdots\times Z_s$ and consider a function $f(z),\ z\in Z$. The divided difference correponding to a subset

$$Z_{I_1\ldots I_s}=\{z_{1i}, i\in I_1\}\times\ldots\times\{z_{si}, i\in I_s\}=Z_{1I_1}\times\ldots\times=Z_{sI_s}$$

of the set Z can be defined in an iterative manner in such a way that first we take k_1-th order divided difference of f with respect to z_1, where $k_1=|I_1|-1$, then the k_2-th order divided difference of that with respect to z_2, where $k_2=|I_2|-1$, etc. This can be executed in a mixed manner, the result will always be the same.

Let $[z_{1i}, i\in I_1;\ldots;z_{si}, i\in I_s; f]$ designate this divided difference and call it of order $(k_1,\ldots,k_s)$. The sum $k_1+\ldots+k_s$ will be called the total order of the divided difference.

The set on which the above divided difference is definedis the Cartesian product of sets on the real line. Let us term such sets *rectangular*. Divided differences on non-rectangular sets have also been defined in the literature (see, e.g., Karlin, Micchelli and Rinott (1986)). These require, however, smooth functions while ours are defined on discrete sets.

A Lagarange interpolation polynomial corresponding to the points in $\{z_{1i}, i\in I_1\}\times\ldots\times\{z_{si}, i\in I_s\}$ is defined by the equation

$$\begin{aligned} &L_{I_1\cdots I_s}(z_1,\ldots,z_s)\\ &=\sum_{i_1\in I_1}\cdots\sum_{i_s\in I_s} f(z_{1i_1},\ldots,z_{si_s})L_{I_1i_1}(z_1)\cdots L_{I_si_s}(z_s), \end{aligned} \tag{8}$$

where

$$L_{I_ji_j}(z_j)=\prod_{h\in I_j\setminus\{i_j\}}\frac{z_j-z_{ji_j}}{z_{jh}-z_{ji_j}},\quad j=1,\ldots,s. \tag{9}$$

The polynomial (8) coincides with the function f at every point of the set $Z_{I_1\ldots I_s}$ and is of degree $m_1\ldots m_s$.

Newton's form of the Lagarange polynomial (8) can be given as follows.

$I_j^{(k_j)}$ designate the set of the first k_j+1 elements of $I_j, 0\le k_j\le m_j, j=1,\ldots,s$. Then the required form is

$$L_{I_1\ldots I_s}(z_1,\ldots,z_s)$$
$$= \sum_{k_1=0}^{m_1}\cdots\sum_{k_s=0}^{m_s}[z_{1h}, h\in I_1^{(k_1)};\ldots;z_{sh}, h\in I_s^{(k_s)};f] \tag{10}$$
$$\prod_{j=1}^{s}\prod_{h\in I_j^{(k_j-1)}}(z_j-z_{jh})$$

2 The Univariate Discrete Power Moment Problem

The problem is defined as the following LP:

$$\begin{aligned} &\text{Min (Max)} \sum_{i=0}^{n} f(z_i)x_i \\ &\text{subject to} \\ &\sum_{i=0}^{n} z_i^k x_i = \mu_k, \quad k=0,1,\ldots,m \\ &x_i \ge 0, \quad i=0,1,\ldots,n. \end{aligned} \tag{11}$$

Here known are $z_0,\ldots,z_n$; $f(z_0),\ldots,f(z_n)$ and $\mu_1,\ldots,\mu_m;\mu_0 = 1$, by definition. The unknowns are $x_0,x_1,\ldots,x_n$ which form a probability distribution with support $Z = \{z_0,z_1,\ldots,z_n\}$. If X is a random variable with support Z and we know $\mu_k = E(X^k), k = 1,\ldots,m$, than the optimum values of problem (11) provide us with sharp lower and upper bounds for $E[f(X)]$. The term "sharp" refers to the fact that knowing only $\mu_1,\ldots,\mu_m$, no better bounds can be given to $E[f(X)]$, then the optimum values of problems (11).

If we introduce the notations:

$$\boldsymbol{a}_i = \begin{pmatrix} 1 \\ z_i \\ \vdots \\ z_i^m \end{pmatrix}, \; f_i = f(z_i), i=0,\ldots,n; \; \mathbf{b} = \begin{pmatrix} 1 \\ \mu_1 \\ \mu_2 \\ \vdots \\ \mu_m \end{pmatrix},$$

then problem (11) can be written in the form:

$$\begin{aligned}
&\text{Min (Max)} \sum_{i=0}^{n} f_i x_i \\
&\text{subject to} \\
&\sum_{i=0}^{n} \boldsymbol{a}_i x_i = \mathbf{b} \\
&x_i \geq 0, \quad i = 0, \ldots, n.
\end{aligned} \tag{12}$$

The matrix $A = (\boldsymbol{a}_0, \ldots, \boldsymbol{a}_n)$ is an $(m+1) \times (n+1)$ Vandermonde matrix, hence every collection of $m+1$ vectors of A forms a basis in this LP.

The basis B is said to be feasible if

$$B^{-1}\mathbf{b} \geq 0 \tag{13}$$

and dual feasible in the minimization (maximization) problem if

$$f_i - f_B^T B^{-1} \boldsymbol{a}_i \ \geq (\leq)\, 0, \quad i = 0, \ldots, n, \tag{14}$$

where f_B is the vector of basic components of $f = (f_0, f_1, \ldots, f_n)^T$. An inequality in (14) holds with equality sign if i is the subscript of a basic vector. In what follows I_B designates the set of subscripts of the basic vectors.

For fixed B, consider the m-degree polynomial

$$L_{I_B}(z) = f_B^T B^{-1} \begin{pmatrix} 1 \\ z \\ \vdots \\ z^m \end{pmatrix}. \tag{15}$$

By the remark made above, we have the relations

$$L_{I_B}(z_i) = f_i = f(z_i), i \in I_B \tag{16}$$

and this implies that $L_{I_B}(z)$ is the m-degree Lagarange polynomial corresponding to the points $\{z_i, i \in I_B\}$. The application of formula (7) gives

$$\begin{aligned}
&f(z) - L_{I_B}(z) \\
&= [z_i, i \in I_B, z; f] \prod_{i \in I_B} (z - z_i).
\end{aligned} \tag{17}$$

Assume now that all $m+1$st order divided differences of f are positive. Then, from (17) we see that $f(z) - L_{I_B}(z) \neq 0$ for $z \notin \{z_i, i \in I_B\}$. Since the left hand side of (14) is the same as that of (17) for $z = z_i$, it follows that each inequality in (14) holds stricktly if $i \notin I_B$. This means that all bases in problems (11)-(12) are dual non-degenerate.

If we look again at (17), we see that the first factor on the right hand side is positive for every $z \notin \{z_i, i \in I_B\}$. Thus, the basis B is dual feasible in the minimization (maximization) problem iff

$$\prod_{i \in I_B} (z - z_i) > (<)\, 0 \quad for\ z \notin \{z_i, i \in I_B\}. \tag{18}$$

This immediately implies

Theorem 2.1. If the function f has positive divided differences of order $m+1$, then B is a dual feasible basis to problem (11)-(12) iff the subscript set I_B has the following structure:

	$m+1\,\text{even}$	$m+1\,\text{odd}$
min	$i, i+1, \ldots, j, j+1$	$0, i, i+1, \ldots, j, j+1$
max	$0, i, i+1, \ldots, j, j+1, n$	$i, i+1, \ldots, j, j+1, n$

If $B_1\,(B_2)$ is a dual feasible basis in the minimization (maximization) problem (11)-(12), then we have the inequalities

$$L_{B_1}(z) \le f(z) \le L_{B_2}(z) \quad z \in Z \tag{19}$$

and

$$E[L_{B_1}(X)] \le E[f(X)] \le E[L_{B_2}(X)]. \tag{20}$$

Since $L_{B_1}(z), L_{B_2}(z)$ are m-degree polynomials, (20) provides us with bounds for $E[f(X)]$, based on the knowledge of the moments $\mu_1, \ldots, \mu_m$. The sharp bounds correspond to bases B_1, B_2 which are optimal in the linear programs (11)-(12).

Remark: If the function f has only nonnegative divided differences of order $m+1$, then only the one way assertion holds: if I_B has the structure in Theorem 2.1, then B is a dual feasible basis in problem (11)-(12).

One can immediately derive bounding formulas for the cases where the number of utilized moments is small. If we only konw μ_1, then any dual feasible basis in the minimization problem has subscripts $j, j+1$ and the only dual feasible basis in the maximization problem has subscripts $0, n$. The latter one is optimal in the maximization problem since any LP that has feasible solution and finite optimum, has at least one primal-dual feasible basis. The former one is optimal in the minimization problem if the basis is also primal feasible, the condition for which is: $z_j \le \mu \le z_{j+1}$. With this j we have

$$\begin{aligned} &\tfrac{z_{j+1}-\mu_1}{z_{j+1}-z_j} f(z_j) + \tfrac{\mu_1 - z_j}{z_{j+1}-z_j} f(z_{j+1}) \\ &\quad \le E[f(X)] \\ &\le \tfrac{z_n - \mu_1}{z_n - z_0} f(z_0) + \tfrac{\mu_1 - z_0}{z_n - z_0} f(z_n). \end{aligned} \tag{21}$$

Note that if $f(z)$ is defined and is convex in the entire interval $[z_0, z_n]$ and the random variable may take any value in this interval, then for $E[f(X)]$ we have Jensen's inequality, as a lower bound and the Edmundson-Madansky inequality as an upper bound. The latter one coincides with the upper bound in (21) while the former one is equal to $f(\mu_1)$ which is different from the lower bound in (21). In fact, it is a weaker bound than the one in (21).

One may formulate the general statment that for a discrete random variable X the discrete moment bounds are always better than the general moment bounds, i.e., the bounds that we can obtain as optimum values of the problems

$$\begin{aligned} &\text{Min (Max)} \int_{z_0}^{z_n} f(z)dF(z) \\ &\text{subject to} \\ &\int_{z_0}^{z_n} z^k dF(z) = \mu_k, \quad k = 0, \ldots, m, \end{aligned} \tag{22}$$

where $F(z)$ is a probability distribution function with support $[z_0, z_n]$.

Bounding formulas for the cases when μ_1, μ_2 or μ_1, μ_2, μ_3 are known, are presented in section 5.9 of Prékopa (1995), for the cases of nonnegative integer valued random variables. An upper bound can be obtained (along the lines the upper bound in section 6.2.5 of the same book is obtained) if $\mu_1, \mu_2, \mu_3, \mu_4$ are known, otherwise we have to use algorithms to solve the problems.

3 Dual algorithm for the solution of problem (11)

We adapt the dual algorithm of Lemke (1954) to solve the problem. Assume that f has positive divided differences of order $m + 1$.

The algorithm starts by picking a dual feasible basis. After that, at each iteration, one vector leaves the basis and another one enters, while the dual feasibility is always preserved.

In the simplex algorithm first we determine the incoming vector and then the outgoing vector. In the dual algorithm it is the other way around. Once the outgoing vector has been determined, the determination of the incoming vector is very simple. In fact, if we remove a vector from the basis, then, by Theorem 2.1, there is one and only one way to restore dual feasibility, by an incoming vector, and it can be found by a simple search procedure. In view of this, we have to concentrate on the problem of identifying an outgoing vector.

Let $I_B = \{i_0, i_1, \ldots, i_m\}$, where $i_0 < i_1 < \ldots < i_m$. If

$$(B^{-1}\mathbf{b})_k < 0 \tag{23}$$

then $\boldsymbol{a}_{i_k}$ may be an outgoing vector. Since we have that

$$\mathbf{b} = E\left[\begin{pmatrix} 1 \\ X \\ \vdots \\ X^m \end{pmatrix}\right]$$

$$B^{-1}\mathbf{b} = E\left[B^{-1}\begin{pmatrix} 1 \\ X \\ \vdots \\ X^m \end{pmatrix}\right]$$

$$L_I(z) = f_B^T B^{-1}\begin{pmatrix} 1 \\ z \\ \vdots \\ z^m \end{pmatrix},$$

it follows that

$$(B^{-1}\mathbf{b})_k = E[L_{I,k}(X)], \tag{24}$$

where $L_{I,k}(z)$ is the kth Lagrange fundamental polynomial

$$L_{I,k}(z) = \prod_{j\in I\setminus\{i_k\}} \frac{z - z_{i_k}}{z_{i_j} - z_{i_k}}. \tag{25}$$

The sign of the denominator in (25) equals $(-1)^{m-k}$. Thus, to determine the sign of (25) we have to look at the polynomial

$$\prod_{j\in I\setminus\{i_k\}} (z - z_{i_j}),$$

replace X for z and take expectation. The obtained value equals

$$\begin{aligned} &\mu_m - \left(\textstyle\sum_{j\in I\setminus\{i_k\}} z_{i_j}\right)\mu_{m-1} \\ &+\cdots + (-1)^m \textstyle\prod_{j\in I\setminus\{i_k\}} z_{i_j}. \end{aligned} \tag{26}$$

If the value in (26), multiplied by $(-1)^{m-k}$, is negative, then the kth vector of the basis can be chosen as the outgoing vector.

The algorithm to solve problem (11) can be summarized as follows.

Step 1. Pick any dual feasible basis in agreement with Theorem 2.1. Let $I = \{i_0, i_1, \ldots, i_m\}$ designate the set of subscripts of the basis vectors.

Determination of the outgoing vector.

Step 2. Take any element $i_k \in I$ and compute the coefficients of $\mu_m, \mu_{m-1}, \ldots, \mu_1$ in (26).

Step 3. Compute the value in (26) and multply it by $(-1)^{m-i_k}$. If it is negative, then $\boldsymbol{a}_{i_k}$ may be an outgoing vector. Otherwise take another element of I. Repeat until an outgoing vector is identified. Otherwise go to step 5.

Determination of the incoming vector.

Step 4. If the outgoing vector is identified, then find that vector which restores dual feasibility of the basis. Choose it as the incoming vector. Go to Step 3.

Step 5. Stop, we have $B^{-1}\mathbf{b} \geq \mathbf{0}$, hence the basis is primal feasible too, i.e., it is optimal. The value $f_B^T B^{-1}\mathbf{b}$ is a lower or upper bound for $E[f(X)]$, depending on which type of dual feasible bases have been used in the algorithm.

Since all bases are dual non-degenerate, it follows that cycling cannot occur and the algorithm terminates in a finite number of steps.

Discrete (as well as general) moment problems are frequently solved for large m, n values, e.g., $m \geq 20$ and n is several hundred or even larger. Since large size Vandermonde matrices are numerically unstable, problem(11) cannot always be solved by general purpose LP packages. This is the reason why we need the above described algorithm to efficiently solve the problem. In this algorithm we work with $(m+1) \times (m+1)$ matrices, at each iteration and the calculation of the values (26) can be carried out in a stable manner, in a reasonable time, if $m \approx 20$, say. At the final step, however, when we compute the optimal value: $f_B^T B^{-1}\mathbf{b}$ (and not only check the signs of some values), special care has to be taken and the use of special algorithms to solve Vandermonde systems of equations (see, e.g., Chun and Kailath (1991) and the references there) are advisable.

If the divided differences of order $m+1$ of the function f are only nonnegative, then the above described algorithm needs some modification: as long as the dual feasible bases are dual non-degenerate we do the same as before but whenever dual degeneracy occurs, some anti-cycling rule should be applied (see,e.g., Prékopa (1995), Chapter 1).

4 Discrete Binomial Moment Problems and the Use of Chebyshev Systems

The discrete binomial moment problem can be stated as

$$\begin{aligned}
&\text{Min (Max)} \sum_{i=0}^{n} f(z_i)x_i \\
&\text{subject to} \\
&\sum_{i=0}^{n} \binom{z_i}{k} x_i = S_k, \quad k = 0, \ldots, m \\
&x_i \geq 0, \quad i = 0, \ldots, n,
\end{aligned} \tag{27}$$

where

$$S_k = E\left[\binom{X}{k}\right] \tag{28}$$

is the kth binomial moment of the random variable with support set $\{z_0, z_1, \ldots, z_n\}$. We have the relations: $S_1 = \mu_1$, $S_2 = \frac{1}{2}(\mu_2 - \mu_1)$, $S_3 = \frac{1}{6}(\mu_3 - 3\mu_2 + 2\mu_1)$ etc. In view of these and similar relations between the binomial coefficients and powers, the equality constraints of problem (27) can be transformed into those of problem (11) (ansd vice versa) by simple linear transformation. This implies that if the $m+1$st order divided differences of the function f are positive, then Theorem 2.1 holds true, further, B is dual (primal) feasible in problem (27) iff it enjoys the same property in problem (11). Thus, we can find an optimal basis to problem (27) by finding one to problem (11). However, there are important problems of the type (27), where the above condition, for the function f, does not hold.

The binomial moment problems are connected with n events $A_1, \ldots, A_n$ and the random variable

$$X = X_1 + \ldots + X_n,$$

where

$$X_i = \begin{cases} 1, & \text{if } A_i \text{ occurs} \\ 0, & \text{otherwise, } i = 1, \ldots, n. \end{cases}$$

A frequently asked question is: what is the probability that at least one out of the n events occur. One answer to this question is given by the inclusion-exclusion formula:

$$\begin{aligned}
&P(A_1 \cup \cdots \cup A_n) \\
&= \sum_{k=1}^{n} (-1)^{k-1} \sum_{1 \leq i_1 < \cdots < i_k \leq n} P(A_{i_1} \cap \cdots \cap A_{i_k}).
\end{aligned} \tag{29}$$

However, terms corresponding to large k values are freqently impossible to compute and in such cases we have to be satisfied with bounds. We make use of the following:

Theorem 4.1. The following equations hold true:

$$\begin{aligned} S_k &= E\left[\binom{X}{k}\right] \\ &= \sum_{1\le i_1<\cdots<i_k\le n} P(A_{i_1}\cap\cdots\cap A_{i_k}). \end{aligned} \tag{30}$$

To obtain bounds for the probability of the union $A_1\cup\cdots\cup A_n$ we may proceed in the following way. We formulate the linear programs

$$\begin{aligned} &\text{Min (Max)} \sum_{i=1}^{n} x_i \\ &\text{subject to} \\ &\sum_{i=1}^{n} \binom{i}{k} x_i = S_k,\ k=1,\ldots,m \\ &x_i \ge 0, \quad i=1,\ldots,n. \end{aligned} \tag{31}$$

If we compare problems (27) and (31), where in problem (27) we have $z_i = i, f(i) = 1, i = 1,\ldots,n,\ f(0) = 0$, we see that the difference is that in problem (31) the constraint of S_0 and the variable x_0 are missing. On the other hand, we can establish relationships between the optimum values of the two problems.

Let $V_{min}(V_{max})$ designate the optimum value of the minimization (maximization) problem in (31). It is easy to see that the optimum value of the minimization problem(27) is the same $V_{\min}$, while the optimum value of the maximization problem (27) is $\min(V_{\max},1)$. In view of this relationship we work with problem (31) rather than with problem (27). The sharp bounds for the union are given by the inequalities:

$$V_{\min} \le P(A_1\cup\cdots\cup A_n) \le \min(V_{\max},1). \tag{32}$$

Discrete moment problems, more general than the power or binomial moment problems can be formulated by the use of the concept of a discrete Chebyshev system.

Definition. We say that the functions $g_0(z), g_1(z),\ldots,g_k(z),\ z\in Z$ form a (discrete variable) Chebyshev system if for every $z_0<z_1<\cdots<z_k$ we have

$$\begin{vmatrix} g_0(z_0) & g_0(z_1) & \ldots & g_0(z_k) \\ g_1(z_0) & g_1(z_1) & \ldots & g_1(z_k) \\ \vdots & \vdots & & \vdots \\ g_k(z_0) & g_k(z_1) & \ldots & g_k(z_k) \end{vmatrix} > 0. \tag{33}$$

If only the $\ge$ sign holds in (33), then we call it a weak Chebyshev system. For more details about Chebyshev systems the reader is referred to Karlin- Studden (1996) or Krein- Nudelman (1977).

Definition. Let $g_0(z), g_1(z), \ldots, g_{k-1}(z)$, $z \in Z$ be a Chebyshev system. We say that the function $f(z), z \in Z$ is convex of order k with respect to this Chebyshev system if for every $z_{i_0} < z_{i_1} < \cdots < z_{i_k}$ we have

$$\begin{vmatrix} g_0(z_{i_0}) & g_0(z_{i_1}) & \ldots & g_0(z_{i_k}) \\ g_1(z_{i_0}) & g_1(z_{i_1}) & \ldots & g_1(z_{i_k}) \\ \vdots & \vdots & & \vdots \\ g_{k-1}(z_{i_0}) & g_{k-1}(z_{i_1}) & \ldots & g_{k-1}(z_{i_k}) \\ f(z_{i_0}) & f(z_{i_1}) & \ldots & f(z_{i_k}) \end{vmatrix} \geq 0. \tag{34}$$

If always the strict inequality holds, then f is said to be strictly convex of order k with respect to the system $g_0(z), g_1(z), \ldots, g_{k-1}(z)$.

Now we formulate the discrete moment problem:

$$\begin{aligned} & \text{Min (Max)} \sum_{i=0}^{n} f(z_i)x_i \\ & \text{subject to} \\ & \sum_{i=0}^{n} g_k(z_i)x_i = \mu_k, \quad k = 0, 1, \ldots, m \\ & x_i \geq 0, \quad i = 0, 1, \ldots, n, \end{aligned} \tag{35}$$

where $g_0(z), g_1(z), \ldots, g_m(z)$, $z \in Z$ form a Chebyshev system and the function $(-1)^{m+1}f(z)$, $z \in Z$ is strictly convex of order $m+1$ with respect to this system. This means that if in (34), written up for $k = m+1$, we put $f(z_{i_0}), f(z_{i_1}), \ldots, f(z_{i_m})$ in the first row, then the resulting determinant is always strictly positive.

A dual feasible basis structure theorem, similar to Theorem 2.1, can be derived for problem (35). We will enunciate it at the end of the derivation.

In what follows, we use problem (12) as another form of problem (34). Let B be a basis in problem (34). In view of the formula

$$\begin{pmatrix} 1 & f_B^T \\ 0 & B \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -f_B^T B^{-1} \\ 0 & B^{-1} \end{pmatrix},$$

we may write

$$\begin{pmatrix} 1 & f_B^T \\ 0 & B \end{pmatrix}^{-1} \begin{pmatrix} f_i \\ \boldsymbol{a}_i \end{pmatrix} = \begin{pmatrix} f_i - f_B^T B^{-1}\boldsymbol{a}_i \\ B^{-1}\boldsymbol{a}_i \end{pmatrix},$$

or

$$\begin{pmatrix} 1 & f_B^T \\ 0 & B \end{pmatrix} \begin{pmatrix} f_i - f_B^T B^{-1}\boldsymbol{a}_i \\ B^{-1}\boldsymbol{a}_i \end{pmatrix} = \begin{pmatrix} f_i \\ \boldsymbol{a}_i \end{pmatrix}.$$

If we use Cramer's rule, we obtain the relation for $f_i - f_B^T B^{-1}\boldsymbol{a}_i$, the first component of the solution of the above system of linear equations:

$$f_i - f_B^T B^{-1}\boldsymbol{a}_i = \frac{1}{|B|}\begin{vmatrix} f_i & f_B^T \\ \boldsymbol{a}_i & B \end{vmatrix}, \tag{36}$$

where $|B|$ designates the determinant of the matrix B.

Let I_B designate the subscript set of the basic vectors. If $i \in I_B$, then the value in (36) is 0. If $i \notin I_B$, then, by assumption, $|B| > 0$ and since the columns of the second determinant are columns of a positive determinant, if suitably rearranged, it follows that the right hand side of (36) is different from 0, and thus, all bases are dual nondegenerate.

If we look at a minimization problem, then B is dual feasible iff (36) is positive for every $i \notin I_B$. This immediately implies that if $m+1$ is even, then the basic vectors form consecutive pairs and if m is odd, then any basis consists of consecutive pairs and $\boldsymbol{a}_n$. Similar statement holds for the maximization problem. Both cases are summarized in the following

Theorem 4.2. The basis B is dual feasible in problem (34) iff the subcript set I_B of the basic vectors has the following structure:

$$\begin{array}{lcc} & m+1\,\text{even} & m+1\,\text{odd} \\ \text{min} & i, i+1, \ldots, j, j+1 & 0, i, i+1, \ldots, j, j+1 \\ \text{max} & 0, i, i+1, \ldots, j, j+1, n & i, i+1, \ldots, j, j+1, n \end{array}$$

Before presenting the application of this theorem, we state another one on binomial coefficients.

Theorem 4.3 (Gessel and Viennot (1985), Prékopa (1988)). Consider the matrix

$$D = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & \cdots & 1 \\ & 1 & 2 & 3 & 4 & \cdots & n \\ & & 1 & \binom{3}{2} & \binom{4}{2} & \cdots & \binom{n}{2} \\ & & & 1 & \binom{4}{3} & \cdots & \binom{n}{3} \\ & & & & & \ddots & \\ & & & & & & \binom{n}{n} \end{pmatrix}. \tag{37}$$

All minors (determinants of square submatrices) of D that have all positive entries in the main diagonal, are positive.

Corollary. Consider the following submatrix of D:

$$E = \begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 & 1 & \cdots & 1 \\ 1 & 2 & 3 & 4 & \cdots & m & m+1 & \cdots & n \\ & 1 & \binom{3}{2} & \binom{4}{2} & \cdots & \binom{m}{2} & \binom{m+}{2} & \cdots & \binom{n}{2} \\ & . & . & . & \cdots & . & . & \cdots & . \\ & & & & & 1 & \binom{m+1}{m} & \cdots & \binom{n}{m} \end{pmatrix}.$$

Theorem 4.3 implies that all minors of order $m+1$ from E and all minors of order m from the last m rows of E are positive.

The above Corollary can be applied to discover the collection of dual feasible bases of problems (31). In fact, the Corollary implies that if $Z = \{1, 2, \ldots, n\}$, then the functions $g_k(i) = \binom{i}{k}, i \in Z,\ k = 1, \ldots, m$ form a Chebyshev system and the constant function $(-1)^m$ is strictly convex of order m with respect to this system. Thus , Theorem 4.2 implies

Theorem 4.4. The basis B is dual feasible in problem (31) iff the subscript set I_B has the structure in Theorem 4.2, where $m+1$ is replaced by m and 0 is replaced by 1.

As special cases we can easily derive the classical bounding formulas for the probability of the union, if $m = 2$. In the minimization problem any dual feasible basis consists of two consicutive vectors. In order tofind out which one of them is primal feasible, i.e., optimal, we have to look for that i, for which the solution of the equations:

$$ix_i + (i+1)x_{i+1} = S_1$$
$$\binom{i}{2}x_i + \binom{i+1}{2}x_{i+1} = S_2$$

satisfies $x_i \geq 0,\ x_{i+1} \geq 0$. It turns out that this i equals

$$i = 1 + \left\lfloor \frac{2S_2}{S_1} \right\rfloor. \tag{38}$$

The lower bound, i.e., the sum of x_1 and x_2 is

$$\frac{2}{i+1}S_1 - \frac{2}{i(i+1)}S_2. \tag{39}$$

This result was first established by Dawson and Sankoff (1967).

As regards the maximization problem, there is just one dual feasible basis and it corresponds to the subscript set $\{1, n\}$. If we solve the equation

$$\begin{aligned} x_1 + nx_n &= S_1 \\ \tbinom{n}{2}x_n &= S_2 \end{aligned}$$

and add x_1 and x_2, then we obtain the upper bound for the union:

$$S_1 - \frac{2}{n} S_2. \tag{40}$$

It is interesting to remark that the classical Bonferroni inequalities

$$P(A_1 \cup \cdots \cup A_n) \geq S_1 - S_2 + \cdots + S_{m-1} - S_m \tag{41}$$

if m is even and

$$P(A_1 \cup \cdots \cup A_n) \leq S_1 - S_2 + \cdots + S_{m-2} - S_{m-1} + S_m\,, \tag{42}$$

if m is odd, can also be derived from problems (27). In fact, the basis corresponding to $I_B = \{1, 2, \ldots, m\}$ is dual feasible in the minimization problem, if m is even and dual feasible in the maximization problem, if m is odd and the right hand side values in (41) and (42) are the corresponding objective function values.

For further closed form inequalities, derived from problem (23), see Kwerel (1975a, b), Boros and Prékopa (1989), Prékopa (1995).

If we look for the probability that at least r or exactly r out of n events occur, then to obtain bounds for them we have to use problem (11) or (27), for the cases of $Z = \{0, 1, \ldots, n\}$ and the functions

$$f(i) = \begin{cases} 1, & \text{if } i \geq r \\ 0, & \text{if } i < r \end{cases} \tag{43}$$

and

$$f(i) = \begin{cases} 1, & \text{if } i = r \\ 0, & \text{if } i \neq r. \end{cases} \tag{44}$$

However, these functions do not have all positive or all negative divided differences on Z, hence the above theory cannot directly be applied to these cases. Still, the collection of dual feasible bases is known in these cases too and dual methods for the solution of the relevant moment problems have been developed (see Prékopa (1990a,b,1995)).

5 Multivariate Discrete Moment Problems

Let $X_1, \ldots, X_s$ be random variables with supports $Z_1, \ldots, Z_s$, respectively, and introduce the notations

$$\begin{array}{c} x_{i_1 \cdots i_s} = P(X_1 = z_{1i_1}, \cdots, X_s = z_{si_s}) \\ 0 \leq i_j \leq n_j, \ j = 1, \ldots, s \end{array} \tag{45}$$

$$\mu_{\alpha_1 \cdots \alpha_s} = \sum_{i_1=0}^{n_1} \cdots \sum_{i_s=0}^{n_s} z_{1i_1}^{\alpha_1} \cdots z_{si_s}^{\alpha_s} x_{i_1 \cdots i_s}, \tag{46}$$

where $\alpha_1, \cdots, \alpha_s$ are nonnegative integers. The number $\mu_{\alpha_1 \cdots \alpha_s}$ is called the $(\alpha_1, \cdots, \alpha_s)$-order moment of the random vector $(X_1, \cdots, X_s)$ and $\alpha_1 + \cdots + \alpha_s$ the total order of the moment.

We are looking for lower and upper bounds on

$$E\left[f(X_1, \ldots, X_s)\right], \tag{47}$$

where f is some function defined on the discrete set Z. In most cases f is supposed to be some multivariate higher order convex function or one of the following types:

$$f(z_1, \ldots, z_s) = \begin{cases} 1, & \text{if } z_i \geq z_{r_i},\ i = 1, \ldots, s \\ 0, & \text{if otherwise} \end{cases} \tag{48}$$

and

$$f(z_1, \ldots, z_s) = \begin{cases} 1, & \text{if } z_i = z_{r_i} \\ 0, & \text{if otherwise.} \end{cases} \tag{49}$$

The choice of the function (48) enables us to create bounds for $P(X_1 \geq z_{r_1}, \ldots, X_s \geq z_{r_s})$, while the function (49) is used for bounding $P(X_1 = z_{r_1}, \ldots, X_s = z_{r_s})$.

We formulate two variants for the multivariate discrete power moment problem. For the sake of simplicity we introduce the notation $f_{i_1 \cdots i_s} = f(z_{1i_1}, \cdots, z_{si_s})$. The first problem is

$$\begin{aligned} & \text{Min (Max)} \sum_{i_1=0}^{n_1} \cdots \sum_{i_s=0}^{n_s} f_{i_1 \cdots i_s} x_{i_1 \cdots i_s} \\ & \text{subject to} \\ & \sum_{i_1=0}^{n_1} \cdots \sum_{i_s=0}^{n_s} z_{1i_1}^{\alpha_1} \cdots z_{si_s}^{\alpha_s} x_{i_1 \cdots i_s} = \mu_{\alpha_1 \cdots \alpha_s} \\ & \text{for } 0 \leq \alpha_j \leq m_j,\ j = 1, \cdots, s \\ & x_{i_1 \cdots i_s} \geq 0, \quad \text{all } i_1, \cdots, i_s. \end{aligned} \tag{50}$$

The second problem is

$$\begin{aligned} & \text{Min (Max)} \sum_{i_1=0}^{n_1} \cdots \sum_{i_s=0}^{n_s} f_{i_1 \cdots i_s} x_{i_1 \cdots i_s} \\ & \text{subject to} \\ & \sum_{i_1=0}^{n_1} \cdots \sum_{i_s=0}^{n_s} z_{1i_1}^{\alpha_1} \cdots z_{si_s}^{\alpha_s} x_{i_1 \cdots i_s} = \mu_{\alpha_1 \cdots \alpha_s} \\ & \text{for } \alpha_j \geq 0,\ j = 1, \cdots, s;\ \alpha_1 + \cdots \alpha_s \leq m \\ & x_{i_1 \cdots i_s} \geq 0, \quad \text{all } i_1, \cdots, i_s. \end{aligned} \tag{51}$$

In these problems the moments $\mu_{\alpha_1 \cdots \alpha_s}$ are supposed to be known and the decision variables are the unknown probabilities $x_{i_1 \cdots i_s}$. The two problems differ only in the assumptions regarding $\alpha_1, \ldots, \alpha_s$. The assumption in problem (51) is more natural than the one in (50). In fact, if, e.g., $s = 2$, then it is natural to assume that we know all moments up to the second order: $E(X_1), E(X_2), E(X_1^2), E(X_2^2), E(X_1X_2)$, as required in problem (51), rather than the collection of moments that appears in problem (50) for $m_1 = 2, m_2 = 2$.

However, it is much easier to create dual feasible bases for problem (50), than for problem (51). The reason is that we can easily create bases for problem (50) by forming the tensor product of univariate bases and if the function f is the product of s suitably defined univariate functions, then we may derive dual feasible basis structure theorems, by the use of the univariate counterparts. Bases of problem (50) which are tensor products of univariate bases will be called rectangular.

Assume now that the function f is of the form

$$f(z_1, \ldots, z_s) = f_1(z_1) \cdots f_s(z_s). \tag{52}$$

Under this assumption the following two theorems (see Prékopa (1998)) provide us with simple methods to construct lower and upper bounds for the value(47). Before formulating the statments we write up the univariate discrete moment problems corresponding to the components of the random vector $(X_1, \ldots, X_s)$:

$$\begin{aligned}
&\text{Min (Max)} \quad \sum_{i=0}^{n_j} f_j(z_{ji}) x_i \\
&\text{subject to} \\
&\sum_{i=0}^{n_j} z_{ji}^{\alpha} x_i = \mu_{\alpha}^{(j)}, \ \alpha = 0, 1, \ldots, m_j \\
&x_i \ge 0, \ i = 0, 1, \ldots, n_j,
\end{aligned} \tag{53}$$

where $\mu_{\alpha}^{(j)} = E(X_j^{\alpha})$, $\alpha = 0, 1, \ldots, m_j$, $j = 1, \ldots, s$.

Theorem 5.1. Suppose that $f_j(z) \ge 0$ for all $z \in Z_j$, $j = 1, \ldots, s$. If for each $j \, (1 \le j \le s)$ we are given a B_j that is dual feasible relative to the maximization problem (53), then their tensor product $B = B_1 \otimes \cdots \otimes B_s$ is a dual feasible basis relative to the maximization problem (50).

Moreover, if I_j is the set of subscripts of B_j and $L_{I_j}(z)$ is the corresponding Lagrange polinomial, then we have the inequality

$$E[f(X_1, \ldots, X_s)] \le E[L_{I_1}(X_1) \cdots L_{I_s}(X_s)] \tag{54}$$

Theorem 5.2. If we replace the assumptions $f_j(z) \geq 0$, $z \in Z_j$ by the assumptions $L_{I_j}(z) \geq 0$, $z \in Z_j$, in Theorem 5.1, and write minimization instead of maximization, then the assertion of the theorem holds true with reversed inequality in (54).

For the cases of problem (51) we have only a few types of dual feasible bases. We briefly mention what are these but disregard their detailed presentation. The interested reader may consult with the paper Prékopa (1998).

If we write up problem (51) in a more compact form, then each column of the matrix of the equality constraints can be represented by a subscript vector $(i_1, \ldots, i_s)$. Let I designate the subscript set of a basis.

Assumption. All divided differences of $f(z_1, \ldots, z_s)$, $(z_1, \ldots, z_s) \in Z_1 \times \cdots \times Z_s$ of total order $m+1$ are nonnegative and all moments $\mu_{\alpha_1 \ldots \alpha_s}$ of total order m are known.

Under this assumption we have the following dual feasible basis sreucture:

$$I = \{(i_1, \ldots, i_s) | i_j \geq 0,\ j = 1, \ldots, s;\ i_1 + \cdots + i_s \leq m\}$$

is a dual feasible basis subscript set for minimization problem(51);

$$I = \{(i_1, \ldots, i_s) | 0 \leq i_j \leq n_j,\ j = 1, \ldots, s;\ n_1 i_1 + \cdots + n_s - i_s \leq m\}$$

is a dual feasible basis subscript set in the minimization (maximization) problem (51), if $m+1$ is even (odd).

If we write up the Lagrange polynomial corresponding to the points

$$(z_{1i_1}, \ldots, z_{si_s}), (i_1, \ldots, i_s) \in I,$$

replace X_j for $z_j, j = 1, \ldots, s$ and take expectation, then we obtain a lower (upper) bound for the value (47) if I corresponds to a minimization (maximization) problem.

The bounds, corresponding to these bases, are not sharp, however. The sharp bounds can be obtained if we carry out the dual method starting form these bases. An algorithm, similar to that presented in Section 3, can be designed for these cases too. More dual feasible bases are presented in Nagy and Prékopa (2000), for the bivariate case.

We can combine the two problems (50), (51) in such a way that we rely on the simpler problem (50) to find good dual feasible bases but we use moments of total order m. We illustrate the procedure in case of $s = 2$, for the sake of simplicity. We assume that all divided differences of $f(z_1, z_2)$, of total order at most $m+1$, are nonnegative.

First we pick some functions $f_0(z_1), f_1(z_1), \ldots, f_m(z_1)$; $g_0(z_2), g_1(z_2), \ldots, g_m(z_2)$, where we assume that $f_i(z_1)$ and $g_i(z_2)$ have nonnegative divided differences of order $i+1$, as single-variable functions. We may choose these functions, e.g., in such a way that we pick $m+1$ values for $z_2 : z_{20}, z_{21}, \ldots, z_{2m}$ and form $f_i(z_1) = f(z_1, z_{2i}), i = 0, 1, \ldots, m.$ and do the same with the first variable to obtain $g_i(z_2) = f(z_{1i}, z_2)$, $i = 0, 1, \ldots, m$.

Assume that all functions $f_i(z_1)g_{m-i}(z_2)$ and their lower bounding polynomials are nonnegative.

The next step is to write up the linear programming problem

$$\begin{aligned}
&\text{Min} \sum_{i=0}^{m} E[f_i(X_1)g_{m-i}(X_2)]y_i \\
&\text{subject to} \\
&\sum_{i=0}^{m} f_i(z_1)g_{m-i}(z_2)y_i \geq f(z_1, z_2),\ (z_1, z_2) \in Z_1 \times Z_2 \\
&\sum_{i=0}^{m} y_i = 1 \\
&y_i \geq 0,\ i = 0, 1, \ldots, m,
\end{aligned} \tag{55}$$

if we want to find upper bound for $E[f(X_1, X_2)]$ or

$$\begin{aligned}
&\text{Max} \sum_{i=0}^{m} E[f_i(X_1)g_{m-i}(X_2)]y_i \\
&\text{subject to} \\
&\sum_{i=0}^{m} f_i(z_1)g_{m-i}(z_2)y_i \leq f(z_1, z_2),\ (z_1, z_2) \in Z_1 \times Z_2 \\
&\sum_{i=0}^{m} y_i = 1 \\
&y_i \geq 0,\ i = 0, 1, \ldots, m,
\end{aligned} \tag{56}$$

if we want to find lower bound.

Each term $f_i(z_1)g_{m-i}(z_2)$ in problems (55),(56) can be bounded from above (below) by the product of two polynomials with degrees i and $m-i$, respectively. If we replace the terms $f_i(z_1)g_{m-i}(z_2)$ by the upper bounding polynomials in the constraints and lower bounding polynomials in the objective function and then solve problem (55), then the optimum value is a lower bound for $E[f(X_1, X_2)]$. An upper bound for $E[f(X_1, X_2)]$ can be obtained from Problem (56) in a similar way.

The multivariate binomial moment problems are very similar to problems (50), (51), we only have to replace $z_{ki_k}^{\alpha_k}$ by $\binom{z_{ki_k}}{\alpha_k}$. In the simplest case we take $Z_i = \{0, 1, \ldots, n_i\}$. These problems are suitabale to create bounds for Boolean functions of finite sequences of events $A_{i1}, \ldots, A_{in_i}$, $i = 1, \ldots, s$.

We can improve on the univariate discrete moment bounds, that can be obtained for a single sequence of events $A_1, \ldots, A_n$, in such a way that we subdivide these sequence into s subsequences and create the multivariate bounds. If each subsequence consists of one single event., then we obtain the Boolean probability bounding scheme (see Prékopa, Vizvári, Regös (1997)).

6 Applications

Probability bounds based on discrete moment problems have been applied to a large number of practical problems. Among these which use LP formulation and higher order moments we mention the reliability bounds for transportation systems(Prékopa and Boros (1991)), communication systems (Prékopa, Boros and Lih (1991)), PERT problem (Prékopa and Long (1992)) and the use of probability bounds for probabilistic constrained stochastic programming problems (Prékopa (1999)). In this section we present some applications in economics and finance.

A utility function $u(z)$ usually indicates a function with the property that $u'(z) > 0$ and $u''(z) < 0$. Some authors (see, e.g., Ingersoll (1987)) argue, however, that in a standard situation the higher order derivatives of the function $u(z)$ also satisfy $u^{(k)} > 0$ for k odd and $u^{(k)} < 0$ for k even. Having this property and a discrete random variable X, we can create lower and upper bounds for $E[u(X)]$ for any number m for which the moments $\mu_k = E(X^k), k = 1, \ldots, m$ are known. For an odd m we use the discrete moment expectation bounds in their original form while for an even m we create the bounds for $E[-u(X)]$ and then take the negatives of the bounds. If $m \leq 4$, then we can present the formulas in closed forms, otherwise we obtain them by executing the algorithm presented in Section 2.

An interesting, new field of application is the valuation of financial instruments, e.g., options.

Bounding option prices (values) has a past history. Without aiming completeness, we mention the papers by Ritchken (1985), Lo (1987), Grundy (1991), Zhang (1994), and Dulá (1995). The bounds, presented in these papers, are based on continuous moment problems, where mostly the knowledge of the first two moments is assumed. The paper by Dulá uses multivariate distribution and assumes the knoeledge of the expectations and the covariance matrix. The paper by Grundy (1991) presents results where higher order moment is used but it is the nth moment alone the knowledge of which is assumed.

Bounding option prices is important because the assumption that the price of the asset follows geometric Brownian motion proves to be invalid in some cases and no other analytic form may be available to replace it. However, from past history data we may obtain moment information and, working with discrete random variables, we can obtain quite good bounds, even if the number of utilized moments is small.

We illustrate the moment bound construction on the European call option. Let us introduce the notations:

$$\begin{aligned} t &= \text{time now} \\ T &= \text{future time, expiration of option} \\ S(t) &= \text{price of underlying asset now} \\ S(T) &= \text{random future price of asset} \\ K &= \text{striking price} \\ r &= \text{rate of interest, we assume continuous compounding} \\ c &= \text{price of the option.} \end{aligned}$$

If we use "risk neutral valuation", then the price of the option is given by the equation:

$$c = e^{-r(T-t)} E([S(T) - K]_+ | S(t) = s). \tag{57}$$

The Black-Scholes formula gives the value of c for the case where $S(\tau)$, $\tau \geq 0$ has the form

$$S(\tau) = S(0) e^{\sigma Z(\tau) + \mu\tau}, \tag{58}$$

where $Z(\tau)$, $\tau \geq 0$ is the standard Brownion motion process, i.e., (a) $Z(0) = 0$, (b) the process has independent increments, (c) $Z(\tau)$ has the distribution $N(0, \tau)$ and $\sigma > 0$, μ are constants (see Black and Scholes (1973)). The process (58) is called multiplicative Brownion motion process.

Now we drop the assumption that $S(\tau)$, $\tau \geq 0$ is a multiplicative Brownion motion process. Let $t = 0$, for the sake of simplicity and assume that $S(T)$ has the form

$$S(T) = e^{\alpha Z + \beta}, \tag{59}$$

where Z is a random variable with support $\{0, \Delta, 2\Delta, \ldots, n\Delta\}$ with $\Delta > 0$. We also assume that there exists an h such that

$$e^{\alpha h \Delta + \beta} - K = 0. \tag{60}$$

Thus, the payoff is 0, if $Z \leq h\Delta$ and is $e^{\alpha k \Delta} - K$, if $Z = k \geq h$. We present two methods for bounding the option price c.

In the first method we assume that the first m conditional moments of Z, given that $e^{\alpha Z+\beta} > K$, are known. Let $\nu_1, \ldots, \nu_m$ designate them. Here too, $\nu_0 = 1$, by definition. Since we have the equation

$$\begin{aligned} c &= e^{-rT}E([S(T)-K]_+) \\ &= e^{-rT}E(e^{\alpha Z+\beta}-K|e^{\alpha Z+\beta} > K)P(e^{\alpha Z+\beta} > K), \end{aligned} \tag{61}$$

we derive lower and upper bounds for the factors in the second line of (61), for c. The bounds for $E(e^{\alpha Z+\beta}-K|e^{\alpha Z+\beta} > K)$ are the optimum values of the LP's:

$$\begin{aligned} &\text{Min (Max) } \sum_{i=h}^{n}(e^{\alpha i\Delta+\beta}-K)x_i \\ &\text{subject to} \\ &\sum_{i=h}^{n}(i\Delta)^k x_i = \nu_k,\ k=0,1,\ldots,m \\ &x_i \geq 0,\ i=h,\ldots,n. \end{aligned} \tag{62}$$

Note that the function

$$f(i) = e^{\alpha i\Delta+\beta}-K,\ i=h,\ldots,n$$

has positive divided differences of all orders that can be defined on the support set $\{h, \ldots, n\}$. Thus , there is no limitation, from this point of view, regarding the number of moments that we can use in the bounding LP.

The lower and upper bounds for

$$P(e^{\alpha i\Delta+\beta} > K) = P(\frac{1}{\Delta}Z \geq h+1)$$

can be based on the moments

$$\mu_k = \frac{1}{\Delta^k}E(Z^k),\ k=0,1,\ldots,m,$$

using the discrete moment problems (11) with the function (43) and $r = h+1$.

The bounds, obtained this way for the option price c, may not be tight enough because we get them as products of bounds obtained by the use of two separate optimization problems.

In the second method we assume that we know the first m moments of Z: $\mu_1, \ldots, \mu_m$. To obtain bounds for c we solve the LP's:

$$\begin{aligned} &\text{Min (Max) } \sum_{i=0}^{n}[e^{\alpha i\Delta+\beta}-K]_+x_i \\ &\text{subject to} \\ &\sum_{i=0}^{n}(i\Delta)^k x_i = \mu_k,\ k=0,1,\ldots,m \\ &x_i \geq 0,\ i=0,1,\ldots,n. \end{aligned} \tag{63}$$

Our discrete moment problem methodology, however, requires that the coefficients in the objective function of problem (63) should have positive or at least nonnegative divided differences of order $m+1$. We will prove that, under some condition, this function has nonnegative divided differences of order $m+1$.

We keep the assumption, mentioned in the first method, that for some positive integer $h\,(0<h<n)$ we have the equality $e^{\alpha h\Delta+\beta}=K$.

Theorem 5.1. The function

$$f(i\Delta)=[e^{\alpha i\Delta+\beta}-K]_+,\; i=0,1,\dots,n \tag{64}$$

has positive divided differences of order up to m, if

$$\alpha\geq\frac{\ln m}{\Delta}.$$

Proof. Since the support set $\{0,\Delta,2\Delta,\dots,n\Delta\}$ has equidistant points, it is enough to show the nonnegativity of the differences (rather than the divided differences) of order up to m. In addition, we may look at the function

$$g(i)=e^{-\beta}f(i\Delta),\; i=0,1,\dots,n,$$

instead of the function (64), and prove the assertion for this, where the suport set is $\{0,1,\dots,n\}$.

It is well-known that the kth order differences of any function g (defined on integers), corresponding to the points $x,x+1,...,x+k$, is given by the formula

$$\Delta^k g(x)=\sum_{i=0}^{k}(-1)^k\binom{k}{i}g(x+k-i). \tag{65}$$

Using (65) we can derive the following results for our function g. If $x\geq h$, then $g(i)=e^{\alpha i\Delta}-e^{\alpha h\Delta}$, $i=x,x=1,\dots,x+k$, hence we have $\Delta^k g(x)>0$. If, on the other hand, $x+k\leq h$, then trivially $\Delta^k g(x)=0$; if $x<h$ and $x+k>h$, then

$$\begin{aligned}\Delta^k g(x) &= \sum_{i=0}^{x+k-h}(-1)^i\tbinom{k}{i}g(x+k-i)\\ &= \sum_{i=0}^{x+k-h}(-1)^i\tbinom{k}{i}\left(e^{\alpha(x+k-i)\Delta}-e^{\alpha h\Delta}\right) \qquad (66)\\ &= e^{\alpha h\Delta}\sum_{i=0}^{x+k-h}(-1)^i\tbinom{k}{i}\left(e^{\alpha(x+k-h-i)\Delta}-1\right).\end{aligned}$$

The last term in the above sum is zero but we keep it, if $x+k-h$ is odd, otherwise we drop it. It is well-known that

$$-\sum_{i=0}^{j}(-1)^i\binom{k}{i} > 0$$

for any odd $j < k$. This implies that if $x+k-h$ is odd then

$$\Delta^k g(x) > e^{\alpha h\Delta}\sum_{i=0}^{x+k-h}(-1)^i\binom{k}{i}e^{\alpha(x+k-h-i)\Delta},$$

and if $x+k-h$ is even, then

$$\Delta^k g(x) > e^{\alpha h\Delta}\sum_{i=0}^{x+k-h-1}(-1)^i\binom{k}{i}e^{\alpha(x+k-h-i)\Delta}.$$

In both cases we combine each term, corresponding to an even i, with the next term. Any combined terms are nonnegative if

$$\frac{k-i}{i+1} \le e^{\alpha\Delta}. \tag{67}$$

Since $(k-i)/(i+1)$ is decreasing in i, its greatest value corresponds to $i=0$, it follows that the difference (66) is positive, if

$$\alpha \ge \frac{\ln k}{\Delta}. \tag{68}$$

Since we have assumed (68) to hold for $k=m$, it follows that it holds for every $k \le m$.

References

[1] Black, M. and M. Scholes (1973). The Pricing of Corporate Liabilities. Journal of Political Economy 81, 637-659.

[2] 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.

[3] Chun, J. and T. Kailath (1991). Displacement Structure for Hankel, Vandermonde, and Related (Derived) Matrices. Linear Algebra and its Applications 151, 199-227.

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

[5] Dulá, J.H. (1995). A Multidimensional Semi-Parametric Upper Bound for Pricng a General Class of Complex Call Options. University of Mississippi, MS 38677.

[6] Fekete, M. and G, Pólya (1912). Über ein Problem von Laguerre. Rendiconti del Circolo Matematico di Palermo 23, 89-120.

[7] Galambos, J. and J. Simonelli (1996). Bonferroni-Type Inequalities with Applications, Springer, New York.

[8] Gessel, I. and G. Viennot (1985). Binomial Determinants, Paths, and Hook Length Formulae. Advances in Mathematics 58, 300-321.

[9] Grundy, B.D. (1991). Option Prices and the Underlying Asset's Return. The Journal of Finance 46, 1045-1069.

[10] Karlin, S. and W.J. Studden (1966). Tchebycheff Systems: With Applications in Analysis and Statistics. Interscience, New York.

[11] Karlin,S., C.A, Micchelli and Y. Rinott (1986). Multivariate Splines: A Probabilistic Perspective. Journal of Multivariate Analysis 20, 69-90.

[12] Konno, H., H., Shirakawa and H. Yamazaki (1993). A Mean-Absolute Deviation-Skewness Portfolio Optimization Model. Annals of Operations Research 45, 205-220.

[13] Krein, M. and A. Nudelman (1977). The Markov Moment Problem and Extremal Problems. Translations of Mathematical Monographs 50, American Mathematical Society, Providence, R.I.

[14] Kwerel, S.M. (1975a). Most Stringent Bounds on Aggregated Probabilities of Partially Specified Dependent Probability Systems. Journal of the American Statistical Association 70, 472-479.

[15] Kwerel, S.M. (1975b) Bounds on the Probability of a Union and Intersection of m Events. Advances of Applied Probability 7, 431-448.

[16] Lemke, C.E. (1854) The Dual Method for Solving the Linear Programming Problem. Naval Research Logistic Quarterly 1, 36-47.

[17] Lo, A. W. (1987). Semi-Parametric Upper Bounds for Option Prices and Expected Payoffs. Journal of Financial Economics 19, 373-387.

[18] Nagy, G. and A. Prékopa (2000). On Higher Order Discrete Convex Functions. RUTCOR Research Report, to appear.

[19] Popoviciu, T. (1945). Les Functions Convex. Actualitiés Scientifique et Industrielles 992, Hermann, Paris.

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

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

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

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

[24] Prékopa, A. (1995). Stochastic Programming. Kluwer Academic Publishers, Dodrecht, Boston.

[25] Prékopa, A. (1998). Bounds on Probabilities and Expectations Using Multivariate Moments of Discrete Distributions. Studia Scientiarum Mathematiearum Hungarica 34, 349-378.

[26] Prékopa, A. (1999). The use of Discrete Moment Bounds in Probabilistic Constrained Stochastic Programming Models. Annals of Operations Research 85, 21-38.

[27] Prékopa, A. and E. Boros (1991). On the Existence of a Feasible Flow in a Stochastic Transportation Network. Operations Research 39, 119-129.

[28] Prékopa, A., E. Boros and Keh-Wei Lih (1991). The Use of Binomial Moments for Bounding Network Reliability. DIMACS Series in Discrete Mathematics and Theoretical Computer Science 5, 197-212.

[29] Prékopa, A. and J. Long (1992). New Bounds and Approximations for the Probability Distribution of the Length of the Critical Path. RUTCOR Research Report 16-92.

[30] Prékopa, A., B. Vizvári and G. Regös (1997). Lower and Upper Bounds on Probabilities of the Boolean Functions of Events. RUTCOR Research Report 21-97.

[31] Ritchken, P. (1985). On Option Pricing Bounds. The Journal of Finance 40, 1219-1233.

[32] Zhang, P. G. (1994). Bounds for Option Prices and Expected Payoffs. Review of Quantitative Finance and Accounting, 179-197.