RUTCOR
RESEARCH
REPORT

Bounds on Probabilities and Expectations Using Multivariate Moments of Discrete Distributions

András Prékopa[a]

RRR 34-93, December 1993

RUTCOR • Rutgers Center for Operations Research • Rutgers University • P.O. Box 5062 • New Brunswick New Jersey • 08903-5062
Telephone: 908-932-3804
Telefax: 908-932-5472
Email: rrr@rutcor.rutgers.edu

[a]RUTCOR, Rutgers Center for Operations Research, Rutgers University, P.O. Box 5062, New Brunswick, New Jersey 08903, prekopa@rutcor.rutgers.edu

Rutcor Research Report
RRR 34-93, December 1993

Bounds on Probabilities and Expectations Using Multivariate Moments of Discrete Distributions

András Prékopa

Abstract. The paper deals with the multivariate moment problems in case of discrete probability distribution. Assuming the knowledge of a finite number of multivariate moments, lower and upper bounds are provided for probabilities and expectations of functions of the random variables involved. These functions obey higher order convexity formulated in terms of multivariate divided differences. As special cases, the multivariate Bonferroni inequalities are derived. The bounds presented are given by formulas as well as linear programming algorithms. Numerical examples are presented.
Key words: multivariate moment problem, stochastic inequalities, linear programming, Lagrange interpolation.

Acknowledgements: Research supported by the Air Force, grant numbers: AFORS-89-0512B, F49620-93-1-0041

1 Introduction

Moment problems are partly existence problems in which we give conditions on finite or infinite sequences of numbers that are necessary and/or sufficient for the existence of a probability distribution of which these numbers are the moments.

Another and, from the point of view of practical applications, even more important research direction is to find bounds on functionals of the unknown probability distribution under some moment information. Moments, at least some of them, are frequently easy to compute (even in experimental sciences, see, e.g., Wheeler and Gordon (1970)) and the bounds that can be obtained on this ground are frequently very good, in the sense that the lower and upper bounds for some value are close to each other. This means that these bounds can be used for approximation purpose as well.

In this paper we present results pertaining to the second research direction. Our functionals are expectations of higher order convex functions of random variables and probabilities of some events.

While the literature is rich in papers handling univariate moment problems, the multivariate case has not been studied enough until recently. The papers by Dula (1986), Kall (1991), Kemperman and Skibinski (1992) and Prékopa (1992b) can be mentioned as examples. The abstract moment problem formulations are in a way more general but they mainly rejoice at duality theory or some vague algorithmic considerations.

A few years ago the sharp Bonferroni inequalities of Dawson and Sankoff (1967), Kwerel (1975) and others, have been discovered as discrete moment problems by Samuels and Studden (1989) and Prékopa (1990). In this case the random variables of which some of the moments are known are occurrences concerning event sequences and the moments are binomial rather than power moments.

Given the information that a random variable is discrete, where the support is also known, the application of the general moment problem (where the support is unrestricted) provides us with weaker bounds than the application of the discrete moment problem. In fact, in the latter case the set of feasible solutions is smaller than in the former case. Discrete random variables with known support are quite frequent in the applications so that research in discrete moment problems is important both from the point of view of theory and applications.

Research in connection with the multivariate discrete moment problem has been initiated in the paper by Prékopa (1992b). This paper presents further and more important results in this respect.

Let $\xi_1, \cdots, \xi_s$ be discrete random variables and assume that the support of ξ_j is a known finite set $Z_j = \{z_{j0}, \cdots, z_{jn_j}\}$, where $z_{j0} < \cdots < z_{jn_j}$, $j = 1, \cdots, s$. Then the support of the random vector $\boldsymbol{\xi} = (\xi_1, \cdots, \xi_s)^T$ is part of the set $\mathbf{Z} = Z_1 \times \cdots \times Z_s$. We do not assume, however, the knowledge that which part of $\mathbf{Z}$ is the exact support of $\boldsymbol{\xi}$.

Let us introduce the notations

$$p_{i_1 \cdots i_s} = P(\xi_1 = z_{1i_1}, \cdots, \xi_s = z_{si_s}) \qquad 0 \le i_j \le n_j,\ j = 1, \cdots, s \tag{1.1}$$

$$\mu_{\alpha_1\cdots\alpha_s} = \sum_{i_1=0}^{n_1}\cdots\sum_{i_s=0}^{n_s} z_{1i_1}^{\alpha_1}\cdots z_{si_s}^{\alpha_s} p_{i_1\cdots i_s}, \tag{1.2}$$

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 $(\xi_1,\cdots,\xi_s)$. The sum $\alpha_1+\cdots+\alpha_s$ is called the total order of the moment.

We assume that the probabilities in (1.1) are unknown but known are some of the multivariate moments (1.2). We are looking for lower and upper bounds on the values

$$E\left[f(\xi_1,\cdots,\xi_s)\right] \tag{1.3}$$

$$P\left(\xi_1 \geq r_1,\cdots,\xi_s \geq r_s\right) \tag{1.4}$$

$$P\left(\xi_1 = r_1,\cdots,\xi_s = r_s\right), \tag{1.5}$$

where f is some function defined on the discrete set $\mathbf{Z}$ and $r_j \in Z_j,\ j=1,\cdots,s$. The problems of bounding the probabilities (1.4) and (1.5) are special cases of the problem of bounding the expectation (1.3). In fact, if

$$f(z_1,\cdots,z_s) = \begin{cases} 1, & \text{if } z_j \geq r_j,\ j=1,\cdots,s \\ 0, & \text{otherwise,} \end{cases} \tag{1.6}$$

then (1.3) is equal to (1.4), and if

$$f(z_1,\cdots,z_s) = \begin{cases} 1, & \text{if } z_j = r_j,\ j=1,\cdots,s \\ 0, & \text{otherwise,} \end{cases} \tag{1.7}$$

then (1.3) is equal to (1.5). In spite of this coincidence, the condition that we will impose on f, when bounding the expectation (1.3), does not always allow for the functions (1.6) and (1.7). Hence, separate attention has to be paid to the problems of bounding the probabilities.

As regards the moments (1.2), two different cases will be considered:

(a) there exist nonnegative integers $m_1,\cdots,m_s$ such that $\mu_{\alpha_1\cdots\alpha_s}$ are known for $0\leq\alpha_j\leq m_j$, $j=1,\cdots,s$;

(b) there exists a positive integer m such that $\mu_{\alpha_1\cdots\alpha_s}$ are known for $\alpha_1+\cdots+\alpha_s\leq m$, $\alpha_j\geq 0,\ j=1,\cdots,s$.

Case (b) is of course more practical than Case (a). If, e.g., we know all expectations, variances and covariances of the random variables $\xi_1,\cdots,\xi_s$, then Case (b) applies. If only the expectations and the covariances are known, then Case (a) applies. However, when the covariances are known then, in most cases, the variances are known too.

We formulate the bounding problems as linear programming problems. For the sake of simplicity we will use the notation $f_{i_1\cdots i_s} = f(z_{1i_1},\cdots,z_{si_s})$. In both problems formulated below the decision variables are the $p_{i_1\cdots i_s}$, all other entries are supposed to be known. In Case (a) the bounding problems are

$$\min(\max) \quad \sum_{i_1=0}^{n_1}\cdots\sum_{i_s=0}^{n_s} f_{i_1\cdots i_s}p_{i_1\cdots i_s}$$

subject to

$$\begin{aligned} &\sum_{i_1=0}^{n_1}\cdots\sum_{i_s=0}^{n_s} z_{1i_1}^{\alpha_1}\cdots z_{si_s}^{\alpha_s}p_{i_1\cdots i_s}=\mu_{\alpha_1\cdots\alpha_s}\\ &\quad\text{for } 0\le\alpha_j\le m_j,\ j=1,\cdots,s\\ &p_{i_1\cdots i_s}\ge 0,\quad \text{all } i_1,\cdots,i_s. \end{aligned} \tag{1.8}$$

In Case (b) the bounding problems are

$$\min(\max) \quad \sum_{i_1=0}^{n_1}\cdots\sum_{i_s=0}^{n_s} f_{i_1\cdots i_s}p_{i_1\cdots i_s}$$

subject to

$$\begin{aligned} &\sum_{i_1=0}^{n_1}\cdots\sum_{i_s=0}^{n_s} z_{1i_1}^{\alpha_1}\cdots z_{si_s}^{\alpha_s}p_{i_1\cdots i_s}=\mu_{\alpha_1\cdots\alpha_s}\\ &\quad\text{for } \alpha_j\ge 0,\ j=1,\cdots,s;\alpha_1+\cdots\alpha_s\le m\\ &p_{i_1\cdots i_s}\ge 0,\quad \text{all } i_1,\cdots,i_s. \end{aligned} \tag{1.9}$$

We reformulate these problems, using more concise notations. Let

$$A_j=\begin{pmatrix} 1 & 1 & \cdots & 1\\ z_{j0} & z_{j1} & \cdots & z_{jn_j}\\ \vdots & & \ddots & \\ z_{j0}^{m_j} & z_{j1}^{m_j} & \cdots & z_{jn_j}^{m_j}\end{pmatrix}, \qquad j=1,\cdots,s$$

$$A=A_1\otimes\cdots\otimes A_s,$$

where the symbol $\otimes$ refers to the tensor product. For example the tensor product of A_1 and A_2 equals

$$A_1\otimes A_2=\begin{pmatrix} A_1 & A_1 & \cdots & A_1\\ z_{20}A_1 & z_{21}A_1 & \cdots & z_{2n_2}A_1\\ \vdots & & \ddots & \\ z_{20}^{m_2}A_1 & z_{21}^{m_2}A_1 & \cdots & z_{2n_2}^{m_2}A_1\end{pmatrix}.$$

Note that the tensor product is noncommutative but it has the associative property. We further introduce the notations:

$$\begin{aligned}
\mathbf{b} &= E\left[(1,\xi_1,\cdots,\xi_1^{m_1})\otimes\cdots\otimes(1,\xi_s,\cdots,\xi_s^{m_s})\right]^T \\
&= (\mu_{00\cdots0},\mu_{10\cdots0},\cdots,\mu_{m_10\cdots0},\mu_{010\cdots0},\mu_{11\cdots0},\cdots)^T \\
\mathbf{p} &= (p_{i_1\cdots i_s},\ 0\le i_1\le m_1,\cdots,0\le i_s\le m_s)^T \\
\mathbf{f} &= (f_{i_1\cdots i_s},\ 0\le i_1\le m_1,\cdots,0\le i_s\le m_s)^T,
\end{aligned}$$

where the ordering of the components in $\mathbf{p}$ and $\mathbf{f}$ coincides with that of the corresponding columns in the matrix $A=(a_{i_1\cdots i_s})$.

The optimum values of the linear programming problems

$$\begin{aligned}
&\min(\max)\quad \mathbf{f}^T\mathbf{p} \\
&\text{subject to} \\
&\qquad A\mathbf{p} = \mathbf{b} \\
&\qquad \mathbf{p} \ge \mathbf{0}
\end{aligned} \tag{1.10}$$

provide us with the best lower and upper bounds for $E[f(\xi_1,\cdots,\xi_s)]$ in Case (a). We call these bounding problems.

In Case (b) we define $\widetilde{\mathbf{b}}$ as the vector obtained from $\mathbf{b}$ so that we delete those components $\mu_{\alpha_1\cdots\alpha_s}$ for which $\alpha_1+\cdots+\alpha_s>m$. Deleting the corresponding rows from A, let $\widetilde{A}$ designate the resulting matrix. Then in Case (b) the bounding problems are:

$$\begin{aligned}
&\min(\max)\quad \mathbf{f}^T\mathbf{p} \\
&\text{subject to} \\
&\qquad \widetilde{A}\mathbf{p} = \widetilde{\mathbf{b}} \\
&\qquad \mathbf{p} \ge \mathbf{0}.
\end{aligned} \tag{1.11}$$

The matrix A has size $[(m_1+1)\cdots(m_s+1)]\times[(n_1+1)\cdots(n_s+1)]$ and is of full rank. The matrix $\widetilde{A}$ has size $N\times[(n_1+1)\cdots(n_s+1)]$, where $N=\binom{s+m}{m}$ and is also of full rank.

It is well-known in linear programming theory that any dual feasible basis (i.e., that satisfies the optimality condition but is not necessarily primal feasible) has the property that the value of the objective function corresponding to the basic solution is smaller (greater) than or equal to the optimum value in case of a minimization (maximization) problem.

Let V_{min} (V_{max}) designate the minimum (maximum) value of any of the problems (1.10) and (1.11). Let further B_1 (B_2) designate a dual feasible basis in any of the minimization (maximization) problems (1.10) and (1.11). Then, in view of the above statement, we have the inequalities

$$\mathbf{f}_{B_1}^T\mathbf{p}_{B_1}\le V_{min}\le E\left[f(\xi_1,\cdots,\xi_s)\right]\le V_{max}\le \mathbf{f}_{B_2}^T\mathbf{p}_{B_2}, \tag{1.12}$$

where $\mathbf{f}_B$ and $\mathbf{p}_B$ designate the vectors of basic components of $\mathbf{f}$ and $\mathbf{p}$, respectively.

In this paper we use some of the basic facts from linear programming and the dual algorithm of Lemke (1954) for the solution of the linear programming problem. A simple and elegant presentation for both can be found in Prékopa (1992a). For the reader's convenience the dual algorithm is briefly summarized in Section 7.

Note that, as it is customary in linear programming, the term "basis" and the symbol "B" mean a matrix and, at the same time, the collection of its column vectors.

In this paper we look for dual feasible bases allowing for inequalities (1.12) and providing us with bounding formulas. If such a bound is not sharp then, starting from such a basis as an initial dual feasible basis, the dual method of linear programming provides us with a sharp algorithmic bound. It is shown by Prékopa (1990a, b) that in case of $s = 1$ the dual method can be executed in a very simple manner. We will show that some simplification is possible in the multidimensional case too.

We can look at the moment problems from a more general point of view, by replacing Chebyshev systems for the matrices $A_1, \cdots, A_s$. Such a generality in handling the problem does not present any new theoretical challenges as compared to the power moment problem, however. On the other hand, the nice formulas that we obtain through Lagrange interpolation polynomials would not be immediately at hand. Therefore we keep the discussion on a more specialized level.

There is one case, however, to which we pay special attention, in addition to the multivariate power moment problem. This is the multivariate binomial moment problem.

We take $Z_j = \{0, \cdots, n_j\}$, $j = 1, \cdots, s$, introduce the cross binomial moments of $\xi_1, \cdots, \xi_s$ as

$$S_{\alpha_1 \cdots \alpha_s} = E\left[\binom{\xi_1}{\alpha_1} \cdots \binom{\xi_s}{\alpha_s}\right]$$

and formulate the problems

$$\min(\max) \quad \sum_{i_1=0}^{n_1} \cdots \sum_{i_s=0}^{n_s} f_{i_1 \cdots i_s} p_{i_1 \cdots i_s}$$

subject to

$$\begin{aligned}
&\sum_{i_1=0}^{n_1} \cdots \sum_{i_s=0}^{n_s} \binom{i_1}{\alpha_i} \cdots \binom{i_s}{\alpha_s} p_{i_1 \cdots i_s} = S_{\alpha_1 \cdots \alpha_s} \\
&\text{for} \quad 0 \le \alpha_j \le m_j, \quad j = 1, \cdots, s \\
&p_{i_1 \cdots i_s} \ge 0, \quad \text{all } i_1, \cdots, i_s,
\end{aligned} \tag{1.13}$$

and

$$\min(\max) \quad \sum_{i_1=0}^{n_1} \cdots \sum_{i_s=0}^{n_s} f_{i_1 \cdots i_s} p_{i_1 \cdots i_s}$$

subject to

$$\begin{aligned} & \sum_{i_1=0}^{n_1} \cdots \sum_{i_s=0}^{n_s} \binom{i_1}{\alpha_i} \cdots \binom{i_s}{\alpha_s} p_{i_1 \cdots i_s} = S_{\alpha_1 \cdots \alpha_s} \\ & \text{for} \quad \alpha_j \geq 0, \quad j = 1, \cdots, s, \quad \alpha_1 + \cdots + \alpha_s \leq m \\ & p_{i_1 \cdots i_s} \geq 0, \quad \text{all } i_1, \cdots, i_s. \end{aligned} \tag{1.14}$$

Problems (1.8) and (1.13) can easily be transformed into each other by the use of the Stirling numbers. Designating the first and second kind Stirling numbers by $s(l,k)$ and $S(l,k)$, respectively, they obey the equations

$$\begin{aligned} \binom{z}{l} &= \sum_{k=0}^{l} \frac{s(l,k)}{l!} z^k \\ z^l &= \sum_{k=0}^{l} S(l,k) k! \binom{z}{k}. \end{aligned}$$

Let us define the matrices

$$T_{1i} = \begin{pmatrix} \frac{s(0,0)}{0!} & & & \\ \frac{s(1,0)}{1!} & \frac{s(1,1)}{1!} & & \\ \vdots & & \ddots & \\ \frac{s(m_i,0)}{m_i!} & \cdots & \cdots & \frac{s(m_i,m_i)}{m_i!} \end{pmatrix}, \quad i = 1, \cdots, s$$

$$T_{2i} = \begin{pmatrix} S(0,0)0! & & & \\ S(1,0)1! & S(1,1)1! & & \\ \vdots & & \ddots & \\ S(m_i,0)m_i! & \cdots & \cdots & S(m_i,m_i)m_i! \end{pmatrix}, \quad i = 1, \cdots, s$$

and the tensor products

$$T_1 = T_{11} \otimes \cdots \otimes T_{1s}$$

$$T_2 = T_{21} \otimes \cdots \otimes T_{2s}.$$

Then, multiplying from the left the matrix of the equality constraints in problem (1.8) (problem (1.13)) by the matrix T_1 (T_2), we obtain problem (1.13) (problem (1.8)). Problems (1.9) and (1.14) can also be transformed into each other by another but still simple rule.

It follows that a basis in problem (1.8) is primal (dual) feasible if and only if it is primal (dual) feasible in problem (1.13).

This simple correspondence does not carry over to problems (1.9) and (1.14).

2 Divided Differences and Lagrange Interpolation

First, let $s = 1$ and, for the sake of simplicity, designate the elements of Z_1 simply by $z_0, \cdots, z_n$.

The divided difference of order 0, corresponding to z_i, is $f(z_i)$, by definition. The first order divided difference corresponding to z_{i_1}, z_{i_2} is designated and defined by

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

where $z_{i_1} \neq z_{i_2}$. The k-th order divided difference is defined recursively by

$$\left[z_{i_1}, z_{i_2}, \cdots, z_{i_k}, z_{i_{k+1}}; f\right] = \frac{[z_{i_2}, \cdots, z_{i_{k+1}}; f] - [z_{i_1}, \cdots, z_{i_k}; f]}{z_{i_{k+1}} - z_{i_1}},$$

where $z_{i_1}, \cdots, z_{i_{k+1}}$ are pairwise different. It is known (see, e.g., Jordan (1965)) that if all divided differences of order k, corresponding to consecutive points, are positive, then all divided differences of order k are positive and we have the equality

$$[z_{i_1}, \cdots, z_{i_{k+1}}; f] = \frac{\begin{vmatrix} 1 & 1 & \cdots & 1 \\ z_{i_1} & z_{i_2} & \cdots & z_{i_{k+1}} \\ \vdots & & \ddots & \\ z_{i_1}^{k-1} & z_{i_2}^{k-1} & \cdots & z_{i_{k+1}}^{k-1} \\ f(z_{i_1}) & f(z_{i_2}) & \cdots & f(z_{i_{k+1}}) \end{vmatrix}}{\begin{vmatrix} 1 & 1 & \cdots & 1 \\ z_{i_1} & z_{i_2} & \cdots & z_{i_{k+1}} \\ \vdots & & \ddots & \\ z_{i_1}^{k-1} & z_{i_2}^{k-1} & \cdots & z_{i_{k+1}}^{k-1} \\ z_{i_1}^{k} & z_{i_2}^{k} & \cdots & z_{i_{k+1}}^{k} \end{vmatrix}}. \tag{2.1}$$

It follows from this that $[z_{i_1}, \cdots, z_{i_{k+1}}; f]$ is independent of the ordering of the points $z_{i_1}, \cdots, z_{i_{k+1}}$.

If f is defined for every z in the interval $[z_0,\ z_n]$ and $f^{(k)}(z) > 0$ for every interior points of the interval then all k-th order divided differences of f are positive on any subset of $k+1$ elements of $\{z_0, \cdots, z_n\}$ (see Jordan (1965)).

The positivity of the first order divided differences means that f is increasing and the positivity of the second order divided differences means that f is a convex discrete function, i.e., the polygon connecting the points $(z_i,\ f(z_i))$, $i = 0, \cdots, n$ is convex. Thus, the second order divided differences of f are positive on $\{z_0, \cdots, z_n\}$ if and only if

$$\frac{f(z_{i_3}) - f(z_{i_1})}{z_{i_3} - z_{i_1}} > \frac{f(z_{i_2}) - f(z_{i_1})}{z_{i_2} - z_{i_1}} \qquad \text{for } z_0 \le z_{i_1} < z_{i_2} < z_{i_3} \le z_n.$$

Given a subset $\{z_i; i \in I\}$ of the set $\{z_0, \cdots, z_n\}$, the Lagrange polynomial corresponding to it is defined by

$$L_I(z) = \sum_{i \in I} L_{Ii}(z) f(z_i), \tag{2.2}$$

where

$$L_{Ii}(z) = \prod_{j \in I - \{i\}} \frac{z - z_j}{z_i - z_j} \tag{2.3}$$

is the i-th fundamental polynomial. Another form of it, using divided differences of f, is Newton's form, that corresponds to an ordering of the elements of the set $\{z_i; i \in I\}$.

If $\{z_i; i \in I\}$ has $m+1$ elements and $I^{(k)}$ is the set of subscripts of the first $k+1$ ones, $0 \le k \le m$ with $I^{(0)} = \{i_0\}$, then Newton's form of the polynomial (2.2) is

$$\begin{aligned} L_I(z) &= f(z_{i_0}) + \sum_{k=1}^{m} \prod_{j \in I^{(k-1)}} (z - z_j) \left[z_i, i \in I^{(k)}; f\right] \\ &= \sum_{k=0}^{m} \prod_{j \in I^{(k-1)}} (z - z_j) \left[z_i, i \in I^{(k)}; f\right], \end{aligned}$$

where $I^{(-1)} = \emptyset$ by definition.

Let $\mathbf{b}(z) = (1, z, \cdots, z^m)^T$. Clearly we have $\mathbf{b}(z_i) = a_i$, $i = 0, \cdots, n$, where a_i is the $i+1$-st column of A in problem (1.10) for the case of $s = 1$. We also have the equality $E[\mathbf{b}(\xi)] = \mathbf{b}$. If B is any basis of this problem and $\mathbf{f}_B$ is the vector of basic components of $\mathbf{f}$, then the polynomial

$$f(z) - \frac{1}{|B|} \begin{vmatrix} f(z) & \mathbf{f}_B^T \\ \mathbf{b}(z) & B \end{vmatrix} = \mathbf{f}_B^T B^{-1} \mathbf{b}(z) \tag{2.4}$$

is of degree m and at any z_i, $i \in I$ its value coincides with $f(z_i)$. Hence, the uniqueness of the Lagrange polynomial implies the equality

$$L_I(z) = \mathbf{f}_B^T B^{-1} \mathbf{b}(z). \tag{2.5}$$

This is, at the same time, the value of the objective function corresponding to that variant of problem (1.10), where we replace $\mathbf{b}(z)$ for $\mathbf{b}$.

In view of (2.5), the dual feasibility of the basis B can be written in the form

$$f(z) \geq L_I(z), \qquad \text{all } z \in \{z_0, \cdots, z_n\}, \tag{2.6}$$

in case of the minimization problem and

$$f(z) \leq L_I(z), \qquad \text{all } z \in \{z_0, \cdots, z_n\}, \tag{2.7}$$

in case of the maximization problem.

It is shown by Prékopa (1990b) that no dual degeneracy occurs in this case, i.e., if (2.6) or (2.7) hold then they hold with strict inequalities for any $z \notin \{z_i; i \in I\}$.

It is well-known in interpolation theory and it can easily be derived from (2.4) and the determinental form of the divided difference that

$$f(z) - L_I(z) = \prod_{j \in I}(z - z_j)\,[z, z_i, i \in I; f]\,. \tag{2.8}$$

This was used by Prékopa (1990b) to establish structural theorems for dual feasible bases. The importance of characterizing the dual feasible bases is shown by the inequalities (2.6), (2.7). In fact, having one of these, we can derive the inequalities for the expectations

$$E\,[f(\xi)] \geq E\,[L_I(\xi)] \tag{2.9}$$

or

$$E\,[f(\xi)] \leq E\,[L_I(\xi)]\,. \tag{2.10}$$

For the case of an arbitrary s, the divided difference corresponding to a subset

$$\mathbf{Z}_{I_1 \cdots I_s} = \{z_{1i}, i \in I_1\} \times \cdots \times \{z_{si}, i \in I_s\} = Z_{1I_1} \times \cdots \times Z_{sI_s}$$

of the set $\mathbf{Z}$ can be defined in an iterative manner so that first we take the 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 $[z_{1i}, i \in I_1; \cdots; z_{si}, i \in I_s; f]$ designate this divided difference and call it of order $(k_1, \cdots, k_s)$. The sum $k_1 + \cdots + k_s$ will be called the total order of the divided difference.

The set on which the above divided difference is defined is 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 Lagrange interpolation polynomial corresponding to the points in $\{z_{1i}, i \in I_1\} \times \cdots \times \{z_{si}, i \in I_s\}$ is defined by the equation

$$L_{I_1 \cdots I_s}(z_1, \cdots, z_s) = \sum_{i_1 \in I_1} \cdots \sum_{i_s \in I_s} f(z_{1i_1}, \cdots, z_{si_s}) L_{I_1 i_1}(z_1) \cdots L_{I_s i_s}(z_s), \tag{2.11}$$

where

$$L_{I_j i_j}(z_j) = \prod_{h\in I_j-\{i_j\}} \frac{z_j - z_{jh}}{z_{ji_j} - z_{jh}}, \qquad j = 1,\cdots,s. \tag{2.12}$$

The polynomial (2.11) coincides with the function f at every point of the set $Z_{I_1\cdots I_s}$ and is of degree $m_1\cdots m_s$.

Newton's form of the Lagrange polynomial (2.11) can be given as follows. Let us order each set Z_{jI_j} and let $I_j^{(k_j)}$ designate the first k_j+1 elements of I_j, $0\le k_j\le m_j$, $j=1,\cdots,s$. Then the required form is

$$L_{I_1\cdots I_s}(z_1,\cdots,z_s) = \sum_{k_1=0}^{m_1}\cdots\sum_{k_s=0}^{m_s}\prod_{j=1}^{s}\prod_{h\in I_j^{(k_j-1)}}(z_j-z_{jh})\left[z_{1h}, h\in I_1^{(k_1)};\cdots;z_{sh}, h\in I_s^{(k_s)}; f\right]. \tag{2.13}$$

The form (2.13) of the polynomial (2.11) does not allow for a straightforward generalization of the formula (2.8) to the multivariate case. Still, there is one generalization of it, using pseudo-polynomials, as follows.

Picking a subset $z_{j_1},\cdots,z_{j_k}$, $j_1<\cdots<j_k$ of the variables $z_1,\cdots,z_s$, we define the *pseudo-polynomial* $L_{I_{j_1}\cdots I_{j_k}}(z_1,\cdots,z_s)$ as the Lagrange polynomial (2.11) of the function obtained from f so that only the z_{j_i}, $i=1,\cdots,k$ are treated as variables while all z_j, $j\notin\{j_1,\cdots,j_k\}$ are fixed for a moment. The resulting function is polynomial of degree $m_{j_1}\cdots m_{j_s}$ in the variables $z_{j_1},\cdots,z_{j_k}$ for any fixed values of the rest of the variables. Let us introduce the notation

$$L^{(k)}(z_1,\cdots,z_s) = \sum_{1\le j_1<\cdots<j_k\le s} L_{I_{j_1}\cdots I_{j_k}}(z_1,\cdots,z_s).$$

Then the following theorem holds.

Theorem 2.1 *We have the formula*

$$\begin{aligned} f(z_1,\cdots,z_s) &= \sum_{k=1}^{s}(-1)^{k-1}L^{(k)}(z_1,\cdots,z_s) \\ &\quad + \prod_{j=1}^{s}\prod_{h\in I_j}(z_j-z_{jh})\left[z_1, z_{1h}, h\in I_1;\cdots;z_s,z_{sh}, h\in I_s; f\right]. \end{aligned} \tag{2.14}$$

Proof. Let us construct an $(m_1+2)\times\cdots\times(m_s+2)$ array, the elements of which are assigned to those elements of $\mathbb{R}^s$ which are in the product set $\{z_1, z_{1h}, h\in I_1\}\times\cdots\times\{z_s, z_{sh}, h\in I_s\}$.

Let us introduce $z_{1m_1+1},\cdots,z_{sm_s+1}$ as alternative notations for $z_1,\cdots,z_s$, respectively, and define the subscript sets

$$I_j^{(m_j+1)} = I_j^{(m_j)}\bigcup\{m_j+1\}, \quad j=1,\cdots,s.$$

Then, assign to the point $(z_{1k_1},\cdots,z_{sk_s})$, $0\le k_i\le m_i+1$, $i=1,\cdots,s$ the value

$$\prod_{j=1}^{s}\prod_{h\in I_j^{(k_j-1)}}(z_j-z_{jh})\left[z_{1h},h\in I_1^{(k_1)};\cdots;z_{sh},h\in I_s^{(k_s)};f\right].$$

Note that this is the same as the general term in the equation (2.13) but now the range of k_j is extended by the element m_j+1, $j=1,\cdots,s$.

Equations (2.8), (2.13) and the inclusion-exclusion formula imply equation (2.14). □

Let us introduce the notations $\mathbf{b}(z_1,\cdots,z_s)$, $\tilde{\mathbf{b}}(z_1,\cdots,z_s)$, where

$$\mathbf{b}(z_1,\cdots,z_s)=(1,z_1,\cdots,z_1^{m_1})\otimes\cdots\otimes(1,z_s,\cdots,z_s^{m_s})$$

and $\tilde{\mathbf{b}}(z)$ is obtained from $\mathbf{b}(z)$ by deleting those components $z_1^{\alpha_1}\cdots z_s^{\alpha_s}$ for which $\alpha_1+\cdots+\alpha_s>m$. Then we have the equalities

$$\mathbf{b}=E[\mathbf{b}(\xi_1,\cdots,\xi_s)],$$

$$\tilde{\mathbf{b}}=E[\tilde{\mathbf{b}}(\xi_1,\cdots,\xi_s)].$$

If B is a basis of the columns of A or $\tilde{A}$ and the notation $\mathbf{z}$ is used for $(z_1,\cdots,z_s)$, then equation (2.4) remains valid.

Let $U=\{u_1,\cdots,u_M\}$ be a set of points in $\mathbb{R}^s$ and $H=\{(\alpha_1,\cdots,\alpha_s)\}$ a finite set of s-tuples of nonnegative integers $(\alpha_1,\cdots,\alpha_s)$.

We say that the set U admits Lagrange interpolation of type H if for any real function $f(\mathbf{z})$, $\mathbf{z}\in U$, there exists a polynomial $p(\mathbf{z})$ of the form

$$p(\mathbf{z})=\sum_{(\alpha_1,\cdots,\alpha_s)\in H}c(\alpha_1,\cdots,\alpha_s)z_1^{\alpha_1}\cdots z_s^{\alpha_s},\tag{2.15}$$

where all $c(\alpha_1,\cdots,\alpha_s)$ are real, such that

$$p(u_i)=f(u_i),\qquad i=1,\cdots,M.\tag{2.16}$$

Equations (2.16) form a system of linear equations for the coefficients $c(\alpha_1,\cdots,\alpha_s)$. If $|H|=M$, then in (2.16) the number of equations is the same as the number of unknowns. Simple linear algebraic facts imply that if U admits Lagrange interpolation then it admits a unique Lagrange interpolation of type H.

Let B be a basis of the columns of the matrix A and H the collection of all power s-tuples of the components of the vector $\mathbf{b}(z_1,\cdots,z_s)$. In this case $|H|=(m_1+1)\cdots(m_s+1)$. Let

$$I=\{(i_1,\cdots,i_s)\,|\,a_{i_1\cdots i_s}\in B\}.\tag{2.17}$$

Then the unique H type Lagrange polynomial corresponding to the set

$$U=\{(z_{1i_1},\cdots,z_{si_s})\,|\,(i_1,\cdots,i_s)\in I\}\tag{2.18}$$

is equal to

$$L_I(z_1, \cdots, z_s) = \mathbf{f}_B^T B^{-1} \mathbf{b}(z_1, \cdots, z_s). \tag{2.19}$$

Since $\mathbf{b}(z_{1i_1}, \cdots, z_{si_s}) = a_{i_1 \cdots i_s}$, it follows that the basis B is dual feasible in the minimization (maximization) problem (1.10) if and only if

$$\begin{aligned} &f(z_1, \cdots, z_s) \geq L_I(z_1, \cdots, z_s), \qquad \text{all } (z_1, \cdots, z_s) \in \mathbf{Z} \\ &(f(z_1, \cdots, z_s) \leq L_I(z_1, \cdots, z_s), \qquad \text{all } (z_1, \cdots, z_s) \in \mathbf{Z}) \,. \end{aligned} \tag{2.20}$$

Note that in (2.20) equality holds for all $(z_1, \cdots, z_s) \in U$.

Let $\widetilde{B}$ be a basis of the columns of $\widetilde{A}$ and H the collection of all power s-tuples of the components of $\widetilde{\mathbf{b}}(z_1, \cdots, z_s)$. If we define I and U as

$$I = \{(i_1, \cdots, i_s) \,|\, \widetilde{a}_{i_1 \cdots i_s} \in \widetilde{B}\} \tag{2.21}$$

$$U = \{(z_{1i_1}, \cdots, z_{si_s}) \,|\, (i_1, \cdots, i_s) \in I\} \tag{2.22}$$

then

$$L_I(z_1, \cdots, z_s) = \mathbf{f}_{\widetilde{B}}^T \widetilde{B}^{-1} \widetilde{\mathbf{b}}(z_1, \cdots, z_s) \tag{2.23}$$

is the unique H type Lagrange polynomial corresponding to the set U.

The dual feasibility of the basis $\widetilde{B}$ in the minimization (maximization) problem means that

$$\begin{aligned} &f(z_1, \cdots, z_s) \geq L_I(z_1, \cdots, z_s), \qquad \text{all } (z_1, \cdots, z_s) \in \mathbf{Z} \\ &(f(z_1, \cdots, z_s) \leq L_I(z_1, \cdots, z_s), \qquad \text{all } (z_1, \cdots, z_s) \in \mathbf{Z}) \end{aligned} \tag{2.24}$$

where equality holds in case of $(z_1, \cdots, z_s) \in U$.

The inequalities (2.20) and (2.24) are the conditions of optimality of the minimization (maximization) problems (1.10) and (1.11), respectively.

Replacing $(\xi_1, \cdots, \xi_s)$ for $(z_1, \cdots, z_s)$ and taking expectations, relations (2.20) and (2.24) provide us with bounds for $E[f(\xi_1, \cdots, \xi_s)]$ in Cases (a) and (b), respectively. If the basis is also primal feasible, then it is optimal and thus, the obtained inequality is sharp.

An important special case, where the set U admits unique Lagrange interpolation of type H, where

$$H = \{(\alpha_1, \cdots, \alpha_s) \,|\, \alpha_i \text{ integer}, \ \alpha_i \geq 0, \ i = 1, \cdots, s, \ \alpha_1 + \cdots + \alpha_s \leq m\} \tag{2.25}$$

is presented in Chung and Yao (1977). We assume that $U = \{u_1, \cdots, u_N\}$ and $N = \binom{s+m}{m}$. The condition that U has to satisfy is termed Condition GC (condition on geometric configuration) and is the following.

Condition GC: For each i, $1 \leq i \leq N$, there exist m distinct hyperplanes $H_{i1}, \cdots, H_{im}$ such that u_i does not lie on any of these hyperplanes and all other points of U lie on at least one of these hyperplanes.

The result by Chung and Yao is the following.

Theorem 2.2 *Suppose that U satisfies Condition GC. Then it admits unique Lagrange interpolation of type H and the interpolating polynomial is*

$$p(\mathbf{z}) = \sum_{i=1}^{N} \frac{\pi_i(\mathbf{z})}{\pi_i(u_i)} f(u_i),$$

where $\pi_i(\mathbf{z}) = u_{i1}(\mathbf{z}) \cdots u_{im}(\mathbf{z})$ and $u_{i1}(\mathbf{z}) = 0, \cdots, u_{im}(\mathbf{z}) = 0$ are the equations of the hyperplanes $H_{i1}, \cdots, H_{im}$, respectively, for $1 \leq i \leq N$.

The converse is also true. If U admits unique Lagrange interpolation of type H and the interpolation polynomial has the form

$$p(\mathbf{z}) = \sum_{i=1}^{N} p_i(\mathbf{z}) f(u_i),$$

where for each i, $1 \leq i \leq N$, the polynomial $p_i(\mathbf{z})$ is the product of m first degree polynomials, then U satisfies Condition GC.

The simplest type of a set that satisfies Condition GC is called a *principal lattice of a simplex* and was first used by Nicolaides (1972) for interpolation purpose.

Let m be a positive integer and $v_1, \cdots, v_{s+1}$ the vertices of a nondegenerate simplex in $\mathbb{R}^s$. Further, let

$$\Lambda = \left\{ \lambda = (\lambda_1, \cdots, \lambda_{s+1}) \mid \lambda_i = \frac{j_i}{m}, 0 \leq j_i \leq m, j_i \text{ integer}, i = 1, \cdots, s+1, j_1 + \cdots + j_{s+1} = m \right\}.$$

Then the set of vectors $\left\{ \sum_{i=1}^{s+1} \lambda_i v_i, \lambda \in \Lambda \right\}$ is called a principal lattice of the simplex. It has $N = \binom{s+m}{m}$ elements and it is easy to see that it satisfies Condition GC. In fact, designating by $\mu_1(\mathbf{z}), \cdots, \mu_{s+1}(\mathbf{z})$ the barycentric coordinates of $\mathbf{x} \in \mathbb{R}^s$ with respect to $\{v_1, \cdots, v_{s+1}\}$, the hyperplanes

$$H_{ij} = \left\{ \mathbf{z} \mid \mu_i(\mathbf{z}) = \frac{j}{m} \right\}, \quad j = 0, \cdots, m-1; \quad i = 1, \cdots, s+1$$

have the required property.

Principal lattices and more general lattices satisfying Condition GC come up in the paper as special cases. In most cases, however, more general lattices are used that still guarantee the uniqueness of the Lagrange interpolation polynomial on those points.

3 Inequalities Based on Rectangular Dual Feasible Bases.

In this section we assume that $f(z_1, \cdots, z_s) = f_1(z_1) \cdots f_s(z_s)$ for $z_i \in Z_i$, $i = 1, \cdots, s$.

For each j, $1 \le j \le s$, we consider the one-dimensional moment problem

$$\begin{aligned} &\min(\max) \quad \sum_{i=0}^{n_j} f_j(z_{ji}) p_i^{(j)} \\ &\text{subject to} \\ &\qquad \sum_{i=0}^{n_j} z_{ij}^{\alpha} p_i^{(j)} = \mu_{\alpha}^{(j)}, \quad \alpha = 0, \cdots, m_j \\ &\qquad\qquad p_i^{(j)} \ge 0, \quad i = 0, \cdots, n_j, \end{aligned} \tag{3.1}$$

where $\mu_{\alpha}^{(j)} = E(\xi_j^{\alpha})$, $\alpha = 0, \cdots, m_j$, $j = 1, \cdots, s$ are known, together with the set $Z_j = \{z_{ji},\ i = 0, \cdots, m_j\}$ and the unknown decision variables are the $p_i^{(j)} = P(\xi_j = z_{ji})$, $i = 0, \cdots, n_j$, $j = 1, \cdots, s$.

We will use the dual feasible basis structure theorems established by Prékopa (1990) for the univariate case.

Theorem 3.1 *Suppose that $f_j(z) \ge 0$ for all $z \in Z_j$. If for each j, $1 \le j \le s$, we are given a B_j that is a dual feasible basis relative to the maximization problem (3.1), then $B = B_1 \otimes \cdots \otimes B_s$ is a dual feasible basis relative to the maximization problem (1.10).*

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

$$E\left[f(\xi_1, \cdots, \xi_s)\right] \le E\left[L_{I_1}(\xi_1) \cdots L_{I_s}(\xi_s)\right]. \tag{3.2}$$

Proof. The dual feasibility of the bases $B_1, \cdots, B_s$ means that

$$\begin{aligned} L_{I_1}(z_1) &\ge f_1(z_1), \quad z_1 \in Z_1 \\ &\vdots \\ L_{I_s}(z_s) &\ge f_s(z_s), \quad z_s \in Z_s. \end{aligned} \tag{3.3}$$

On the other hand, the unique H-type Lagrange polynomial, with $H = \{(\alpha_1, \cdots, \alpha_s) \mid 0 \le \alpha_j \le m_j, \alpha_j \text{ integer}, j = 1, \cdots, s\}$, is given by (2.11). Since $f(z_1, \cdots, z_s) = f_1(z_1) \cdots f_s(z_s)$, it follows that the polynomial (2.11) takes the form

$$L_{I_1 \cdots I_s}(z_1, \cdots, z_s) = L_{I_1}(z_1) \cdots L_{I_s}(z_s). \tag{3.4}$$

Since the dual feasibility of B relative to the maximization problem (1.10) is the same as the second inequality in (2.20), the theorem follows by (3.3) and (3.4). □

Theorem 3.2 *Suppose that $L_{I_j}(z) \ge 0$ for all $z \in Z_j$. If for each j, $1 \le j \le s$, we are given a B_j that is a dual feasible basis relative to the minimization problem (3.1), then $B = B_1 \otimes \cdots \otimes B_s$ is a dual feasible basis relative to the minimization problem (1.10).*

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

$$E\left[f(\xi_1, \cdots, \xi_s)\right] \ge E\left[L_{I_1}(\xi_1) \cdots L_{I_s}(\xi_s)\right]. \tag{3.5}$$

Proof. The proof is the same as that of Theorem 3.1, with a slight modification. □

3.1 Upper Bounds for Probabilities $P(\xi_1 \geq z_{r_1}, \cdots, \xi_s \geq z_{r_s})$

Theorem 3.1, combined with the one-dimensional upper bounds presented by Prékopa (1990b) provides us with a variety of upper bounds for probabilities mentioned in the title of this subsection.

Recall from the above mentioned paper that if for each $j = 1, \cdots, s$, the function in the maximization problem (3.1) is the following

$$f_j(z) = \begin{cases} 0, & \text{if } z < z_{r_j} \\ 1, & \text{if } z \geq z_{r_j} \end{cases}$$

then every dual feasible basis subscript set I_j has one of the following structures:
$m_j + 1$ even

$$\begin{array}{ll} I_j \subset \{r_j, \cdots, n_j\}, & \text{if } n_j - r_j \geq m_j \\ \{i, i+1, \cdots, h, h+1, r_j, k, k+1, \cdots, t, t+1, n_j\}, & \text{if } 2 \leq r_j \leq n_j - 1 \\ \{0, i, i+1, \cdots, h, h+1, r_j, k, k+1, \cdots, t, t+1\}, & \text{if } 1 \leq r_j \leq n_j \end{array}$$

$m_j + 1$ odd

$$\begin{array}{ll} I_j \subset \{r_j, \cdots, n_j\}, & \text{if } n_j - r_j \geq m_j \\ \{i, i+1, \cdots, h, h+1, r_j, k, k+1, \cdots, t, t+1\}, & \text{if } 1 \leq r_j \leq n_j - 1 \\ \{0, i, i+1, \cdots, h, h+1, r_j, k, k+1, \cdots, t, t+1, n_j\}, & \text{if } 1 \leq r_j \leq n_j - 1 \end{array}$$

where the number in the parentheses are arranged in increasing order. The second and third type bases in both groups are dual nondegenerate, while the first type bases are dual degenerate.

Now, picking arbitrarily the dual feasible bases $I_1, \cdots, I_s$, in accordance with the above structures, the inequality (3.2) provides us with the upper bound.

A rectangular basis of this type is illustrated in Figure 1 for the case of $s = 2$, $Z_1 = Z_2 = \{0, \cdots, 9\}$.

Example 1. Let $s = 2$, $Z_j = \{0, \cdots, n_j\}$, $j = 1, 2$, $m_1 = 1$, $m_2 = 3$ and choose the dual feasible bases as follows:

$$\begin{aligned} I_1 &= \{0, r_1, k, k+1\}, \qquad r_1 \geq 1 \\ I_2 &= \{0, r_2, t, t+1\}, \qquad r_2 \geq 1. \end{aligned}$$

Then we have

$$L_{I_1}(z) = \frac{(z-0)(z-k)(z-k-1)}{(r_1-0)(r_1-k)(r_1-k-1)} + \frac{(z-0)(z-r_1)(z-k-1)}{(k-0)(k-r_1)(k-k-1)} + \frac{(z-0)(z-r_1)(z-k)}{(k+1-0)(k+1-r_1)(k+1-k)}$$

$$L_{I_2}(z) = \frac{(z-0)(z-t)(z-t-1)}{(r_2-0)(r_2-t)(r_2-t-1)} + \frac{(z-0)(z-r_2)(z-t-1)}{(t-0)(t-r_2)(t-t-1)} + \frac{(z-0)(z-r_2)(z-t)}{(t+1-0)(t+1-r_2)(t+1-t)}.$$

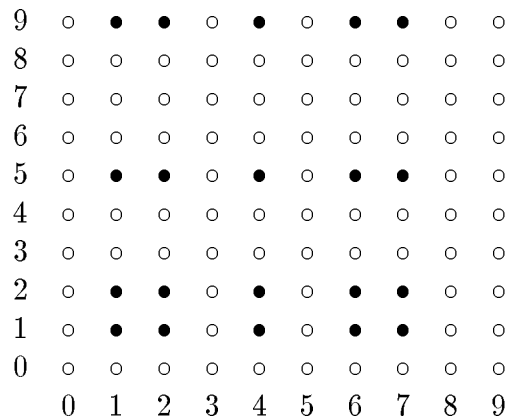

Figure 1: Illustration of a rectangular dual feasible basis through the planar points to which the basic columns of A correspond in the maximization problem (1.10). We choose $m_1+1=5, r_1=4, m_2+1=4, r_2=5$.

To be more specific, let $r_1=4,\ r_2=5,\ k=6,\ t=6$. Then the above polynomials take the forms

$$L_{I_1}(z)=\frac{z}{168}\left(z^2-17z+94\right)$$

$$L_{I_2}(z)=\frac{z}{210}\left(z^2-18z+107\right).$$

The inequality (3.2) specializes to

$$\begin{aligned} P(\xi_1\geq 4,\ \xi_2\geq 5) \quad \leq \quad & \frac{1}{35280}\left(\mu_{33}-18\mu_{32}+107\mu_{31}\right. \\ & -17\mu_{23}+306\mu_{22}-1819\mu_{21} \\ & \left.+94\mu_{13}-1619\mu_{12}+10058\mu_{11}\right). \end{aligned}$$

Example 2. Let $m_j=2,\ r_j=1,\ j=1,\cdots,s$. Then the only dual feasible bases are those that correspond to the subscript sets $I_j=\{0,1,n_j\},\ j=1,\cdots,s$. The Lagrange polynomials take the form

$$L_{I_j}(z)=-\frac{(z-z_{j0})(z-z_{jn_j})}{z_{jn_j}-z_{j1}}+\frac{(z-z_{j0})(z-z_{j1})}{(z_{jn_j}-z_{j0})(z_{jn_j}-z_{j1})},\quad j=1,\cdots,s.$$

In case of $s=1$ the bound (3.2) is sharp because the unique dual feasible basis must be primal feasible too (we have assumed that the right hand side values in problem (3.1) are moments of some random variable which implies that the problem has feasible solution; it has finite optimum too because the set of feasible solutions is compact). The bound (3.2) is sharp in the multivariate case too, in the sense that the basis in problem (1.10), corresponding to the subscript set $I=I_1\times\cdots\times I_s$ is primal and dual feasible, hence optimal.

Of particular interest is the case where $Z_j = \{0, \cdots, n_j\}$, $j = 1, \cdots, s$ and the random variable ξ_j is equal to the number of events that occur among some events $E_{j1}, \cdots, E_{jn_j}$, $j = 1, \cdots, s$. We may write $L_{I_j}(z)$ in the form

$$L_{I_j}(z) = z - \frac{z(z-1)}{2}\frac{2}{n_j}, \quad j = 1, \cdots, s$$

from which we derive

$$L_{I_1}(z_1) \cdots L_{I_s}(z_s) = \left[z_1 - \frac{z_1(z_1-1)}{2}\frac{2}{n_1}\right] \cdots \left[z_s - \frac{z_s(z_s-1)}{2}\frac{2}{n_s}\right].$$

Using the cross binomial moments $S_{\alpha_1 \cdots \alpha_s}$, the inequality (3.2) can be obtained. It is also a sharp one. For example, if $s = 2$ then we obtain

$$\begin{aligned} P((A_{11} \cup \cdots \cup A_{1n_1}) \cap \cdots \cap (A_{s1} \cup \cdots \cup A_{sn_s})) &= P(\xi_1 \geq 1, \xi_2 \geq 1) \\ &\leq S_{1,1} - \frac{2}{n_2}S_{1,2} - \frac{2}{n_1}S_{2,1} + \frac{4}{n_1 n_2}S_{2,2}. \end{aligned}$$

This result was first obtained by Galambos and Xu (1991).

The general formula can be written in the form

$$\begin{aligned} P\left(\bigcap_{j=1}^{s}\bigcup_{i=1}^{n_j} A_{ji}\right) &= P(\xi_1 \geq 1, \cdots, \xi_s \geq 1) \\ &\leq \sum_{1 \leq \alpha_j \leq 2, j=1,\cdots,s} (-1)^{\alpha_1 + \cdots + \alpha_s - s} S_{\alpha_1 \cdots \alpha_s} \frac{1}{n_1^{\alpha_1 - 1}} \cdots \frac{1}{n_s^{\alpha_s - 1}}. \end{aligned} \tag{3.6}$$

Upper bounds for probabilities $P(\xi = z_{r_1}, \cdots, \xi_s = z_{r_s})$ can be established in a similar fashion, based on the one-dimensional basis structure theorems of Prékopa (1990).

3.2 Lower Bounds for Probabilities $P(\xi_1 \geq z_{r_1}, \cdots, \xi_s \geq z_{r_s})$

First we recall from Prékopa (1990b) that if $f_j(z)$ are the functions defined in Section 3.1, $j = 1, \cdots, s$ then every dual feasible basis subscript set in the minimization problem (3.1) has one of the following structures:

$m_j + 1$ even

$$\begin{array}{ll} I_j \subset \{0, \cdots, r_j - 1\}, & \text{if } r_j \geq m_j + 1 \\ \{0, i, i+1, \cdots, h, h+1, r_j - 1, k, k+1, \cdots, t, t+1\}, & \text{if } 2 \leq r_j \leq n_j - 1 \\ \{i, i+1, \cdots, h, h+1, r_j - 1, k, k+1, t, t+1, \cdots, n_j\}, & \text{if } 1 \leq r_j \leq n_j \end{array}$$

$m_j + 1$ odd

$$\begin{array}{ll} I_j \subset \{0, \cdots, r_j - 1\}, & \text{if } r_j \geq m_j + 1 \\ \{0, i, i+1, \cdots, h, h+1, r_j - 1, k, k+1, \cdots, t, t+1, n_j\}, & \text{if } 2 \leq r_j \leq n_j \\ \{i, i+1, \cdots, h, h+1, r_j - 1, k, k+1, t, t+1\}, & \text{if } 1 \leq r_j \leq n_j - 1. \end{array}$$

Looking at Theorem 3.2, we see that if for each $j = 1, \cdots, s$, I_j is a dual feasible basis subscript set for problem (3.1), then we need the additional condition: $L_{I_j}(z) \geq 0$, $j = 1, \cdots, s$ to be able to infer to the dual feasibility of the basis corresponding to $I = I_1 \times \cdots \times I_s$.

This creates difficulty especially because some of the basis subscript sets do not contain the number n_j and thus the same Lagrange polynomial can be nonnegative on Z_j for a smaller n_j and my take negative values too for larger n_j. In view of this, we turn our attention to special cases.

Let $r_j = 1$, $z_{j0} = 0$, $m_j = 2$, $I_j = \{0, z_{ji_j}, z_{ji_j+1}\}$, $j = 1, \cdots, s$. Then the basis corresponding to the subscript set I_j is dual feasible in problem (3.1). The Lagrange polynomial $L_{I_j}(z)$ takes the form

$$\begin{aligned} L_{I_j}(z) &= \frac{z(z - z_{ji_j+1})}{z_{ji_j}(z_{ji_j} - z_{ji_j+1})} + \frac{z(z - z_{ji_j})}{z_{ji_j+1}(z_{ji_j+1} - z_{ji_j})} \\ &= \frac{z(z_{ji_j} + z_{ji_j+1} - z)}{z_{ji_j} z_{ji_j+1}} \end{aligned}$$

This polynomial is nonnegative for $0 \leq z \leq z_{n_j}$ iff $z_{ji_j} + z_{ji_j+1} \geq z_{n_j}$.

In the special case where $\{z_{j0}, \cdots, z_{jn_j}\} = \{0, \cdots, n_j\}$, then nonnegativity condition for $L_{I_i}(z)$ is that $2i_j + 1 \geq n_j$. Assuming these to be the case, for each j, $1 \leq j \leq s$, we may write

$$P(\xi_1 \geq 1, \cdots, \xi_s \geq 1) \geq E\left[\prod_{j=1}^{s} \frac{\xi_j(2i_j + 1 - \xi_j)}{i_j(i_j + 1)}\right]. \tag{3.7}$$

If ξ_j, $j = 1, \cdots, s$ designate the occurrences concerning the event sets $E_{j1}, \cdots, E_{jn_j}$, $j = 1, \cdots, s$, respectively, as described in Section 3.1, then it is desirable to give (3.7) another form, expressed in terms of the cross binomial moments. For the case of $s = 2$ the inequality (3.7) gives the following result

$$\begin{aligned} P(\xi_1 \geq 1, \xi_2 \geq 1) &\geq E\left[\prod_{j=1}^{2}\left[\frac{2\xi_j}{i_j + 1} - \frac{2}{i_j(i_j + 1)}\binom{\xi_j}{2}\right]\right] \\ &= \frac{4}{(i_1 + 1)(i_2 + 1)} S_{1,1} - \frac{4}{(i_1 + 1)i_2(i_2 + 1)} S_{1,2} \\ &\quad - \frac{4}{(i_2 + 1)i_1(i_1 + 1)} S_{2,1} + \frac{4}{i_1(i_1 + 1)i_2(i_2 + 1)} S_{2,2}. \end{aligned} \tag{3.8}$$

For an arbitrary s the formula is:

$$\begin{aligned} P\left(\bigcap_{j=1}^{s} \bigcup_{i=1}^{n_j} A_{ji}\right) &= P(\xi_1 \geq 1, \cdots, \xi_s \geq 1) \\ &\geq \sum_{\substack{\alpha_1 + \cdots + \alpha_s \leq 2s \\ 1 \leq \alpha_j \leq 2, j = 1, \cdots, s}} (-1)^{\alpha_1 + \cdots + \alpha_s - s} S_{\alpha_1 \cdots \alpha_s} \frac{4}{i_1^{\alpha_1 - 1}(i_1 + 1) \cdots i_s^{\alpha_s - 1}(i_s + 1)}, \end{aligned} \tag{3.9}$$

where it is assumed that $i_j \geq (n_j - 1)/2$.

Lower bounds for probabilities $P(\xi_1 = z_{r_1}, \cdots, \xi_s = z_{r_s})$ can be established in a similar fashion, based on the relevant one-dimensional dual feasible basis structure theorems of Prékopa (1990b).

It should be mentioned, in connection with problem (1.10), that the optimal basis is not necessarily a rectangular one as it has been shown by Prékopa (1992a). We can reach the optimal basis by starting from any dual feasible basis and carry out the dual method for solving problem (1.10).

In order to find a good rectangular dual feasible basis we can choose I_j, $j = 1, \cdots, s$ so that I_j is optimal for problem (3.1), provided that it is a maximization problem. In case of the minimization problem we choose the best among those dual feasible bases for which the Lagrange polynomial is nonnegative. A dual feasible basis is best in the sense that any dual step that improves on the objective function does not preserve the nonnegativity of the Lagrange polynomial.

Note that having the best univariate bases $I_1, \cdots, I_s$, $I = I_1 \times \cdots \times I_s$ is not necessarily the best rectangular basis.

In Figure 2 we illustrate a basis of this type for the case of $s = 2$, $Z_1 = Z_2 = \{0, \cdots, 9\}$.

9	●	○	●	●	○	●	○	●	●	○
8	○	○	○	○	○	○	○	○	○	○
7	○	○	○	○	○	○	○	○	○	○
6	●	○	●	●	○	●	○	●	●	○
5	○	○	○	○	○	○	○	○	○	○
4	○	○	○	○	○	○	○	○	○	○
3	○	○	○	○	○	○	○	○	○	○
2	○	○	○	○	○	○	○	○	○	○
1	○	○	○	○	○	○	○	○	○	○
0	●	○	●	●	○	●	○	●	●	○
	0	1	2	3	4	5	6	7	8	9

Figure 2: Illustration of a rectangular dual feasible basis through the planar points to which the basic columns of A correspond in the minimization problem (1.10). We chose $m_1 + 1 = 6, r_1 = 6, m_2 + 1 = 3, r_2 = 7$.

4 Bounds Based on Bivariate Moments of Total Order m

In this section we assume that $s = 2$ and the known bivariate moments are $\mu_{\alpha\beta}$, where $\alpha \geq 0$, $\beta \geq 0$ are integers and $\alpha + \beta \leq m$. We present lower and upper bounds on $E[f(\xi_1, \cdots, \xi_s)]$, using problem (1.9), (1.11).

Theorem 4.1 *Let* $I = \{(i_1, i_2) | i_1 \geq 0, i_2 \geq 0$ *integer,* $i_1 + i_2 \leq m\}$ *and assume that all divided differences of total order* $m+1$*, of the function* f*, are nonnegative. Then the following assertions hold.*

(a) *The set of columns* $\{a_{i_1 i_2} | (i_1, i_2) \in I\}$ *is a basis* B *for the columns of* $\tilde{A}$ *in problem (1.11).*

(b) *The Lagrange polynomial* $L_I(z_1, z_2)$ *corresponding to the points* $\{(z_{1i_1}, z_{2i_2}) | (i_1, i_2) \in I\}$ *is unique and is the following*

$$L_I(z_1, z_2) = \sum_{\substack{i_1 + i_2 \leq m \\ 0 \leq i_j \leq n_j, j=1,2}} [z_{10}, \cdots, z_{1i_1}; z_{20}, \cdots, z_{2i_2}; f] \prod_{j=1}^{2} \prod_{h=0}^{i_j - 1} (z_j - z_{jh}), \tag{4.1}$$

where $\prod_{h=0}^{i_j-1} (z_j - z_{jh}) = 1$ *for* $i_j = 0$*, by definition.*

(c) *We have the inequalities*

$$f(z_1, z_2) \geq L_I(z_1, z_2) \qquad \text{for } (z_1, z_2) \in \mathbf{Z}, \tag{4.2}$$

i.e., B *is dual feasible in the minimization problem (1.11),*

$$E[f(\xi_1, \xi_2)] \geq E[L_I(\xi_1, \xi_2)]. \tag{4.3}$$

If B *is also a primal feasible basis in problem (1.11), then the inequality (4.3) is sharp.*

(d) *If all divided differences of total order* $m+1$ *are nonpositive, then all assertions hold with the difference that* B *is dual feasible in the maximization problem (1.11) and the inequalities (4.2) and (4.3) are reversed.*

Proof. The proofs of assertions (a) and (b) can be combined.

First we observe that the polynomial (4.1) coincides with f at the points $\{(z_{1i_1}, z_{2i_2}) | (i_1, i_2) \in I\}$. In fact, if (z_{1k_1}, z_{2k_2}) is an elements of this set and we replace z_{1k_1} and z_{2k_2} for z_1 and z_2, respectively, then only those terms in (4.1) may be different from 0 for which $i_1 \leq k_1$, $i_2 \leq k_2$. The obtained expression is

$$\sum_{i_2=0}^{k_2} \sum_{i_1=0}^{k_1} [z_{10}, \cdots, z_{1i_1}; z_{20}, \cdots, z_{2i_2}; f] \prod_{j=1}^{2} \prod_{h=0}^{i_j-1} (z_{jk_j} - z_{jh}).$$

This can be recognized as the interpolating polynomial, in Newton's form, corresponding to the set $\{z_{10}, \cdots, z_{1k_1}\} \times \{z_{20}, \cdots, z_{2k_2}\}$, taken at the point (z_{1k_1}, z_{2k_2}). Thus, $L_I(z_{1k_1}, z_{2k_2}) = f(z_{1k_1}, z_{2k_2})$.

Since the above assertion holds for any function $f(z_1, z_2)$, $(z_1, z_2) \in Z$, it follows that the interpolating polynomial is unique which implies that B is a basis for the columns of A.

This last assertion can be proved by another way too. Applying a method similar to what is usually applied to find the determinant of a Vandermonde matrix, we obtain that

$$|B| = \prod_{j=1}^{2} \prod_{h=0}^{m-1} \prod_{i=0}^{m-(h+1)} (z_{jm-h} - z_{ji})^{h+1} > 0.$$

Proof of (c). Let us define the function

$$\begin{aligned} R_I(z_1, z_2) \quad = \quad & [z_{10}, \cdots, z_{1m}, z_1; z_{20}; f](z_1 - z_{10}) \cdots (z_1 - z_{1m}) \\ & +[z_{10}, \cdots, z_{1m-1}, z_1; z_{20}, z_{21}; f](z_1 - z_{10}) \cdots (z_1 - z_{1m-1})(z_2 - z_{20}) \\ & \vdots \\ & +[z_{10}, z_1; z_{20}, \cdots, z_{2m}; f](z_1 - z_{10})(z_2 - z_{20}) \cdots (z_2 - z_{2m-1}) \\ & +[z_1; z_{20}, \cdots, z_{2m}, z_2; f](z_2 - z_{20}) \cdots (z_2 - z_{2m}). \end{aligned} \tag{4.4}$$

It may happen that in the above formula $z_1 \in \{z_{10}, \cdots, z_{1i}\}$ for some i values and/or for $z_2 \in \{z_{20}, \cdots, z_{2m}\}$. Then the corresponding divided differences can be defined arbitrarily because these terms are 0 anyway. The number of terms defining $R_I(z_1, z_2)$ is $m + 1$. We will show that

$$L_I(z_1, z_2) + R_I(z_1, z_2) = f(z_1, z_2) \qquad \text{for all } (z_1, z_2) \in \mathbf{Z}. \tag{4.5}$$

In fact, combining terms from (4.1) with terms from (4.4), we may write

$$\begin{aligned} L_I(z_1, z_2) + R_I(z_1, z_2) \quad = \quad & \left\{ \sum_{i=0}^{m} [z_{10}, \cdots, z_{1i}; z_{20}; f] \prod_{h=0}^{i-1} (z_1 - z_{1h}) \right. \\ & \left. +[z_{10}, \cdots, z_{1m}, z_1; z_{20}; f] \prod_{h=0}^{m} (z_1 - z_{1h}) \right\} \\ & + \left\{ \sum_{i=0}^{m-1} [z_{10}, \cdots, z_{1i}; z_{20}, z_{21}; f] \prod_{h=0}^{i-1} (z_1 - z_{1h})(z_2 - z_{20}) + \right. \\ & \left. [z_{10}, \cdots, z_{1m-1}, z_1; z_{20}, z_{21}; f] \prod_{h=0}^{m-1} (z_1 - z_{1h})(z_2 - z_{20}) \right\} \\ & \vdots \\ & + \left\{ \sum_{i=0}^{1} [z_{10}, z_{1i}; z_{20} \cdots, z_{2m-1}; f] \prod_{h=0}^{i-1} (z_1 - z_{1h}) \prod_{k=0}^{m-2} (z_2 - z_{2k}) \right. \\ & \left. +[z_{10}, z_1; z_{20}, \cdots, z_{2m}; f](z_1 - z_{10}) \prod_{k=0}^{m-1} (z_2 - z_{2k}) \right\} \\ & + \left\{ [z_{10}; z_{20} \cdots, z_{2m}; f] \prod_{k=0}^{m-1} (z_2 - z_{2k}) \right. \end{aligned} \tag{4.6}$$

$$\begin{aligned} & \left. +[z_1; z_{20}, \cdots, z_{2m}, z_2; f] \prod_{k=0}^{m} (z_2 - z_{2k}) \right\} \\ = \; & f(z_1, z_{20}) + [z_1; z_{20}, z_{21}; f](z_2 - z_{20}) + \cdots \\ & +[z_1; z_{20} \cdots, z_{2m}, z_2; f](z_2 - z_{20}) \cdots (z_2 - z_{2m}) \\ = \; & f(z_1, z_2). \end{aligned}$$

The divided differences that appear in the definition of $R_I(z_1, z_2)$ are of total order $m+1$, hence they are all nonnegative. The factors that multiply them are also nonnegative for any $(z_1, z_2) \in \mathbf{Z}$. Hence $R_I(z_1, z_2) \geq 0$ for $(z_1, z_2) \in Z$. This proves (4.2) and (4.3) follows from (4.2).

If B is also a primal feasible basis in problem (1.11), then it is optimal and, consequently, (4.3) is sharp. $\square$

Remark. If $z_{ji} = i$, $i = 0, \cdots, n_j$, $j = 1, \cdots, s$, then the inequality (4.3) has another form:

$$E[f(\xi_1, \xi_2)] \geq \sum_{\substack{i_1+i_2 \leq m \\ 0 \leq i_j \leq n_j, j=1,2}} [0, \cdots, i_1; 0, \cdots, i_2; f] i_1! i_2! S_{i_1 i_2}. \tag{4.7}$$

Figure 3 illustrates a dual feasible basis of Theorem 4.1 for the case of $Z_1 = Z_2 = \{0, \cdots, 9\}$.

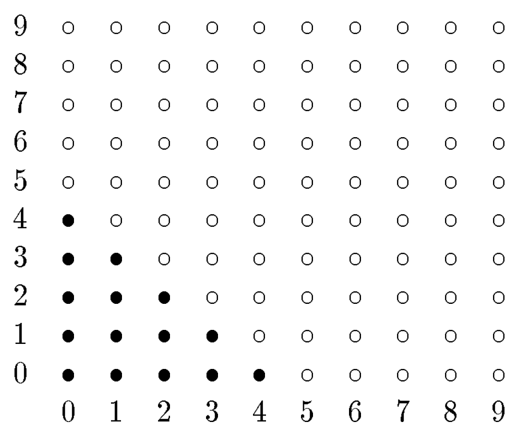

Figure 3: Illustration of a dual feasible basis through the planar points to which the basic columns of A correspond in the minimization problem (1.11). We chose $m + 1 = 4$.

Theorem 4.2 *Let $I = \{(n_1 - i_1, n_2 - i_2) | i_1, i_2$ integers, $0 \leq i_1 \leq n_1, 0 \leq i_2 \leq n_2, i_1 + i_2 \leq m\}$ and assume that all divided differences of total order $m+1$ of the function f are nonnegative. Then the following assertions hold.*

(a) *The set of columns $\{a_{i_1 i_2} | (i_1, i_2) \in I\}$ is a basis for the columns of $\widetilde{A}$ in problem (1.11).*

(b) *The Lagrange polynomial $L_I(z_1, z_2)$ corresponding to the points $\{(z_{1i_1}, z_{2i_2})|(i_1, i_2) \in I\}$ is unique and is the following*

$$L_I(z_1, z_2) = \sum_{\substack{i_1+i_2 \le m \\ 0 \le i_j \le n_j, j=1,2}} [z_{1n_1}, \cdots, z_{1n_1-i_1}; z_{2m2}, \cdots, z_{2n_2-i_2}; f] \prod_{j=1}^{2} \prod_{h=n_j-(i_j-1)}^{n_j} (z_j - z_{jh}). \quad (4.8)$$

(c) *If $m+1$ is odd (even), then*

$$f(z_1, z_2) \le (\ge) L_I(z_1, z_2) \qquad \text{for } (z_1, z_2) \in \mathbf{Z}, \quad (4.9)$$

i.e., B is a dual feasible basis in the maximization (minimization) problem (1.11) and

$$E[f(\xi_1, \xi_2)] \le (\ge) E[L_I(\xi_1, \xi_2)]. \quad (4.10)$$

If B is also primal feasible then the inequality (4.10) is sharp.

(d) *If all divided differences of total order $m+1$ are nonpositive, then all assertions hold with the difference that B is a dual feasible basis in the minimization (maximization) problem (1.11) and the inequalities (4.9) and (4.10) are reversed.*

Proof. Assertions (a) and (b) can be proved in the same way as we have proved the corresponding assertions of Theorem 4.1. Assertion (c) follows by applying Theorem 4.1 for the function

$$g(z_1, z_2) = f(z_1 - (z_{1n_1} - z_{10}), z_2 - (z_{2n_2} - z_{20})), \quad (z_1, z_2) \in Z.$$

Assertion (d) holds trivially. □

Figures 4.a and 4.b illustrate dual feasible bases of Theorem 4.2 for the case of $Z_1 = Z_2 = \{0, \cdots, 9\}$.

Theorem 4.3 *Let $I = \{(0,0), (1,0), (0,1), (n_1, 0), (0, n_2), (n_1, n_2)\}$ and assume that all divided differences of total order 3 of the function f are nonnegative. Then the following assertions hold.*

(a) *The set of columns $\{a_{i_1 i_2}|(i_1, i_2) \in I\}$ is a basis B for the columns of $\tilde{A}$ in problem (1.11).*

(b) *The Lagrange polynomial $L_I(z_1, z_2)$ corresponding to the points $\{(z_{i_1}, z_{i_2}) \,|\, (i_1, i_2) \in I\}$ is unique and is the following*

$$\begin{aligned} L_I(z_1, z_2) \quad = \quad & f(z_{10}, z_{20}) + [z_{10}, z_{11}; z_{20}; f](z_1 - z_{10}) \\ & + [z_{10}; z_{20}, z_{21}; f](z_2 - z_{20}) \\ & + [z_{10}, z_{11}, z_{1n_1}; z_{20}; f](z_1 - z_{10})(z_1 - z_{11}) \\ & + [z_{10}; z_{20}, z_{21}, z_{2n_2}; f](z_2 - z_{20})(z_2 - z_{21}) \\ & + [z_{10}, z_{11}; z_{20}, z_{21}; f](z_1 - z_{10})(z_2 - z_{20}). \end{aligned} \quad (4.11)$$

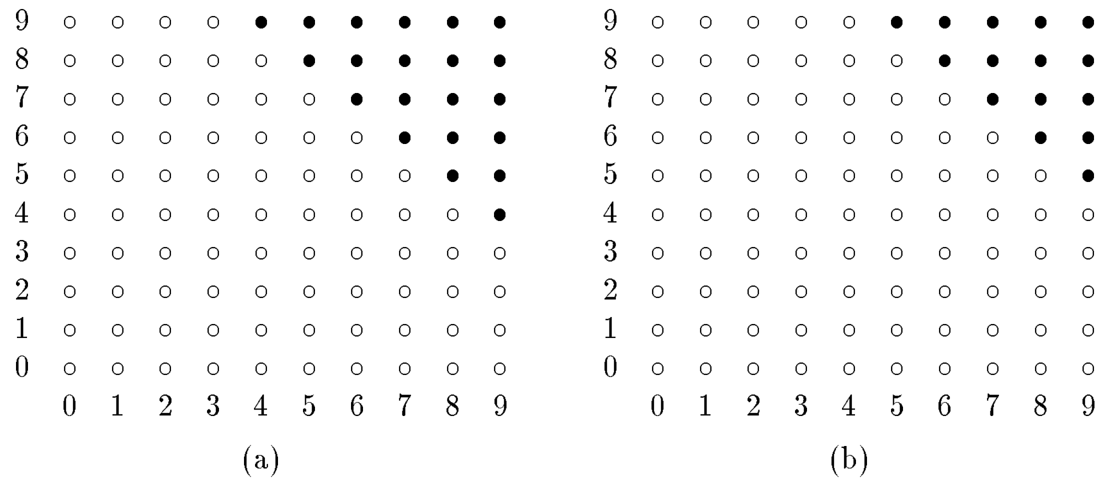

Figure 4: Illustration of dual feasible bases through the planar points to which the basic columns of $\widetilde{A}$ correspond in the minimization problem (1.11). The basis in Figure 4.a (4.b) yields an upper (lower) bound because $m+1=5$ is odd ($m+1=4$ is even).

(c) *We have the inequalities*

$$f(z_1,z_2) \le L_I(z_1,z_2) \qquad \text{for } (z_1,z_2) \in \mathbf{Z}, \tag{4.12}$$

i.e., B is dual feasible in the maximization problem (1.11),

$$E[f(\xi_1,\xi_2)] \le E[L_I(\xi_1,\xi_2)]. \tag{4.13}$$

If B is also a dual feasible basis in problem (1.11), then the inequality (4.12) is sharp.

(d) *If all divided differences of total order 3 are nonpositive, then all assertions hold with the difference that B is dual feasible in the minimization problem (1.11) and the inequalities (4.12) and (4.13) are reversed.*

Proof of (a). It is a simple exercise to check that $|B| \ne 0$. Thus, B is in fact a basis.
Proof of (b). The uniqueness of the Lagrange polynomial follows from (a). Note that Theorem 2.2 cannot be applied here because the point system $\{(z_{1i_1}, z_{2i_2}) \,|\, (i_1,i_2) \in I\}$ does not satisfy Condition GC.

That the polynomial $L_I(z_1,z_2)$, given by (4.11), is in fact the Lagrange polynomial corresponding to the points $\{(z_{1i_1} \mid z_{2i_2}), (i_1,i_2) \in I\}$ follows from the fact that $L_I(z_1,z_2)$ coincides with $f(z_1,z_2)$ on them (as it is easy to see).

Now we show that (4.12) holds. First we assume that $z_1 > z_{10}$, $z_2 > z_{20}$.

In view of the assumption that the $(2,1),(1,2)$ order divided differences are nonnegative, we have the inequalities

$$[z_{10}, z_1; z_{20}, z_2; f] \le [z_{10}, z_{1n_1}; z_{20}, z_2; f] \le [z_{10}, z_{1n_1}; z_{20}, z_{2n_2}; f].$$

It follows from this that

$$f(z_1,z_2) \le f(z_{10},z_2) + f(z_1,z_{20}) - f(z_{10},z_{20}) + [z_{10},z_{1n_1};z_{20},z_{2n_2};f](z_1-z_{10})(z_2-z_{20}). \quad (4.14)$$

On the other hand, the nonnegativity of the $(3,0)$ order divided differences and the fact that $\{0,1,n_1\}$ is a univariate dual feasible basis structure in the problem

$$\begin{aligned} &\min \quad \sum_{i=0}^{n_1} f(z_{1i},z_{20})p_i^{(1)} \\ &\text{subject to} \\ &\qquad \sum_{i=0}^{n_1} z_{1i}^{\alpha} p_i^{(1)} = \mu_{\alpha 0}, \quad \alpha = 0,1,2 \\ &\qquad p_i^{(1)} \ge 0, \qquad i = 0,\cdots,n_1, \end{aligned} \quad (4.15)$$

(see Prékopa (1990b)) imply that

$$f(z_1,z_{20}) \le f(z_{10},z_{20}) + [z_{10},z_{11};z_{20};f](z_1-z_{10}) + [z_{10},z_{11},z_{12};z_{20};f](z_1-z_{10})(z_1-z_{11}). \quad (4.16)$$

In a similar way we obtain

$$f(z_{10},z_2) \le f(z_{10},z_{20}) + [z_{10};z_{20},z_{21};f](z_2-z_{20}) + [z_{10};z_{20},z_2;z_{22};f](z_2-z_{20})(z_2-z_{21}). \quad (4.17)$$

The inequalities (4.14), (4.16) and (4.17) imply (4.12).

The expectation inequality (4.13) follows from (4.12). If B is also primal feasible, then it is optimal in the maximization problem (1.11), hence the inequality is sharp. Assertion (d) follows from the fact that in this case the function $-f$ has nonnegative divided differences of total order 3, hence the inequalities (4.14), (4.16) and (4.17) hold if we replace $-f$ for f. These imply the reversed inequalities of (4.12) and (4.13).

If $z_1 = z_{10}$ or $z_2 = z_{20}$ the (4.12) reduces to (4.17) or (4.16). This completes the proof. □

A dual feasible basis of Theorem 4.3 is illustrated in Figure 5.

Remark. If $Z_1 = \{0,\cdots,n_1\}$, $Z_2 = \{0,\cdots,n_2\}$, then the inequality in (4.13) takes the form

$$\begin{aligned} E[f(\xi_1,\xi_2)] \;\le\; & f(0,0) + [0,1;0;f]S_{10} + [0;0,1;f]S_{01} + 2[0,1,n_1;0;f]S_{20} \\ & +2[0;0,1,n_2;f]S_{02} + [0,n_1;0,n_2;f]S_{11}. \end{aligned} \quad (4.18)$$

Theorem 4.4 *Let $I = \{(0,0),(0,1),(1,0),(1,1),(0,n_2),(n_1,0)\}$ and assume that all divided differences of orders $(2,1),(1,2)$ of the function f are nonnegative while the divided differences of orders $(3,0),(0,3)$ are nonpositive. Then the following assertions hold.*

(a) *The set of columns $\{a_{i_1 i_2} | (i_1,i_2) \in I\}$ is a basis B for the columns of $\tilde{A}$ in problem (1.11).*

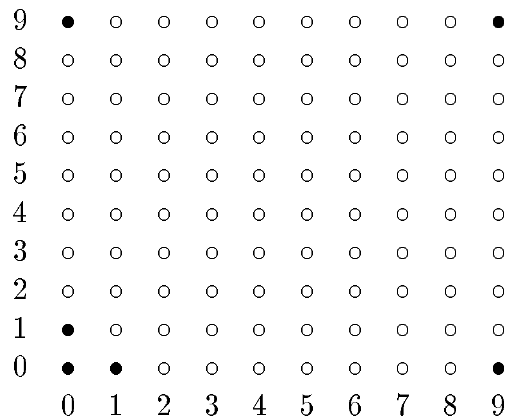

Figure 5: Illustration of a dual feasible basis through the planar points to which the basic columns of $\widetilde{A}$ correspond in the maximization problem (1.11).

(b) *The Lagrange polynomial $L_I(z_1, z_2)$, corresponding to the points $\{(z_{i_1}, z_{i_2})|(i_1, i_2) \in I\}$ is unique and is the following*

$$\begin{aligned} L_I(z_1, z_2) \quad = \quad & f(z_{10}, z_{20}) + [z_{10}, z_{11}; z_{20}; f](z_1 - z_{10}) \\ & +[z_{10}; z_{20}, z_{21}; f](z_2 - z_{20}) \\ & +[z_{10}, z_{11}, z_{1n_1}; z_{20}; f](z_1 - z_{10})(z_1 - z_{11}) \\ & +[z_{10}; z_{20}, z_{21}, z_{2n_2}; f](z_2 - z_{20})(z_2 - z_{21}) \\ & +[z_{10}, z_{11}; z_{20}, z_{21}; f](z_1 - z_{10})(z_2 - z_{20}). \end{aligned} \tag{4.19}$$

(c) *We have the inequalities*

$$f(z_1, z_2) \geq L_I(z_1, z_2) \qquad \textit{for } (z_1, z_2) \in \mathbf{Z}, \tag{4.20}$$

i.e., B is dual feasible in the minimization problem (1.11),

$$E[f(\xi_1, \xi_2)] \geq E[L_I(\xi_1, \xi_2)]. \tag{4.21}$$

If B is also a primal feasible basis in problem (1.11), then the expectation inequality (4.21) is sharp.

(d) *If all divided differences of orders $(2,1),(1,2)$ are nonpositive, and those of orders $(3,0),(0,3)$ are nonnegative, then all assertions hold with the difference that B is dual feasible in the maximization problem (1.11) and the inequalities (4.20) and (4.21) are reversed.*

Proof. The proof is very similar to that of Theorem 4.3 and is omitted. □

Figure 6 illustrates a dual feasible basis of Theorem 4.4.

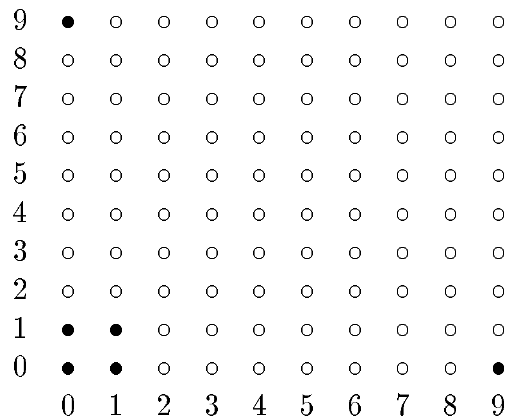

Figure 6: Illustration of a dual feasible basis through the planar points to which the basic columns of $\tilde{A}$ correspond in the minimization problem (1.11).

Remark. If $Z_1 = \{0, \cdots, n_1\}$, $Z_2 = \{0, \cdots, n_2\}$, then the inequality in (4.21) can be written in the form

$$\begin{aligned} E[f(\xi_1, \xi_2)] \quad \geq \quad & f(0,0) + [0,1;0;f]S_{10} + [0;0,1;f]S_{01} + 2[0,1,n_1;n_2 f]S_{20} \\ & +2[0;0,1,n_2;f]S_{02} + [0,1;0,1;f]S_{11}. \end{aligned} \tag{4.22}$$

The proof of Theorem 4.1 allows for the derivation of simular dual feasibility assertions for other lattices, using other assumptions. For example, if $m = 3$ and the divided differences of orders $(2,0), (2,1), (1,3)$ and $(0,4)$ are nonnegative, then the set of points $\{(z_{1i_1}, z_{2i_2}) | (i_1, i_2) \in I\}$, with $I = \{(0,0), (0,1), (0,2), (0,3), (1,0), (1,1)\}$, determines a unique Lagrange interpolation and a dual feasible basis in the minimization problem (1.11). The Lagrange polynomial is

$$\begin{aligned} L_I(z_1, z_2) \quad = \quad & [z_{10}; z_{20}; f] + [z_{10}; z_{10}, z_{21}; f](z_2 - z_{20}) \\ & +[z_{10}; z_{20}, z_{21}, z_{22}; f](z_2 - z_{20})(z_2 - z_{21}) \\ & +[z_{10}; z_{20}, z_{21}, z_{22}, z_{23}; f](z_2 - z_{20})(z_2 - z_{21})(z_2 - z_{22}) \\ & +[z_{10}, z_{11}; z_{20}; f](z_1 - z_{10}) \\ & +[z_{10}, z_{11}; z_{20}, z_{21}; f](z_1 - z_{10})(z_2 - z_{20}). \end{aligned}$$

From here inequalities of the type (4.2) and (4.3) can be derived.

5 Bounds Based on Multivariate Moments of Total Order m

We assume that the known moments are: $\mu_{\alpha_1 \cdots \alpha_s}$, where $\alpha_j \geq 0$, $j = 1, \cdots, s$, $\alpha_1 + \cdots + \alpha_s \leq m$.

The theorems presented in this section are generalizations of the Theorems 4.1 and 4.2.

Theorem 5.1 *Let* $I = \{(i_1, \cdots, i_s) | i_j \geq 0,$ *integers,* $j = 1, \cdots, s, i_1 + \cdots + i_s \leq m\}$ *and assume that all divided differences of total order* $m+1$*, of the function* f*, are nonnegative. Then the following assertion hold.*

(a) *The set of columns* $\{a_{i_1 \cdots i_s} | (i_1, \cdots, i_s) \in I\}$ *is a basis* B *for the columns of* $\tilde{A}$ *in problem (1.11).*

(b) *The Lagrange polynomial* $L_I(z_1, \cdots, z_s)$*, corresponding to the points* $\{(z_{1i_1}, \cdots, z_{si_s}) | (i_1, \cdots, i_s) \in I\}$ *is unique and is the following*

$$L_I(z_1, \cdots, z_s) = \sum_{\substack{i_1 + \cdots i_s \leq m \\ 0 \leq i_j \leq n_j, j=1,\cdots,s}} [z_{10}, \cdots, z_{1i_1}; \cdots; z_{s0}, \cdots, z_{si_s}; f] \prod_{j=1}^{s} \prod_{h=0}^{i_j - 1} (z_j - z_{jh}), \quad (5.1)$$

where, by definition, $\prod_{h=1}^{i_j - 1} (z_j - z_{jh}) = 1$*, for* $i_j = 0$*.*

(c) *We have the inequalities*

$$f(z_1, \cdots, z_s) \geq L_I(z_1, \cdots, z_s), \qquad \text{for } (z_1, \cdots, z_s) \in Z, \quad (5.2)$$

i.e., B *is a dual feasible basis in the minimization problem (1.11),*

$$E[f(\xi_1, \cdots, \xi_s)] \geq E[L_I(\xi_1, \cdots, \xi_s)]. \quad (5.3)$$

If B *is also a primal feasible basis in problem (1.11), then the inequality (5.3) is sharp.*

(d) *If all divided differences of total order* $m+1$ *are nonpositive, then all assertions hold with the difference that* B *is a dual feasible basis in the maximization problem (1.11) and the inequalities (5.2), (5.3) are reversed.*

Proof. The proof of assertions (a) and (b) is the same as those in the proof of Theorem 4.1, therefore it will not be detailed.

We also mention, without proof, that the determinant of B has a simple form

$$|B| = \prod_{j=1}^{s} \prod_{h=0}^{m-1} \prod_{i=0}^{m-(h+1)} (z_{jm-h} - z_{ji})^{h+1}.$$

The factors in this formula are the denominators that appear in connection with the divided differences in (5.1). Since $|B| > 0$, assertion (a) and the uniqueness of the Lagrange polynomial can also be derived from this.

In (5.1), as well as in the next formula, we use only the numbers $z_{i0}, \cdots, z_{im}$, from $Z_i = \{z_{10}, \cdots, z_{1n_i}\}$, $1 \leq i \leq s$. Hence, it will not be misleading, if for each $i = 1, \cdots, s$ we designate z_i by z_{im+1}. This is only for the simplicity of the next formula.

With this notation we define the function $R_I(z_1,\cdots,z_s)$, $(z_1,\cdots,z_s)\in Z$ as follows

$$R_I(z_1,\cdots,z_s)=\sum_{\substack{i_1+\cdots+i_s=m+1\\ 0\le i_j\le n_j, j=1,\cdots,s}}[z_{10},\cdots,z_{1i_1};\cdots;z_{s0},\cdots,z_{si_s};f]\prod_{j=1}^{s}\prod_{h=0}^{i_j-1}(z_j-z_{jh}). \quad (5.4)$$

It may happen that in the above formula $z_j=z_{jm_j+1}\in\{z_{10},\cdots,z_{jm_j}\}$, for some j. Then the corresponding divided difference can be defined arbitrarily because that term is 0 anyway. (This remark applies to (5.5), too). We show that $L_I(z_1,\cdots,z_s)+R_I(z_1,\cdots,z_s)=f(z_1,\cdots,z_s)$.

The proof can be carried out by induction. For $s=1$ it is the same as equation (2.8). For $s=2$ the equation (5.4) reduces to (4.5) that has already been proved.

Assuming that (5.4) holds for the case of $s-1$, for any function f, we derive the equality

$$\begin{aligned}L_I(z_1,\cdots,z_s)+R_I(z_1,\cdots,z_s) &= \sum_{i_s=0}^{m}[z_1,\cdots,z_{s-1};z_{s0},\cdots,z_{si_s};f]\prod_{h=0}^{i_s-1}(z_s-z_{sh})\\ &+[z_1,\cdots z_{s-1};z_{s0},\cdots,z_{sm},z_s;f]\prod_{h=0}^{m}(z_s-z_{sh}). \quad (5.5)\end{aligned}$$

By (2.8) we see that this is further equal to $[z_1;\cdots;z_{s-1};z_s;f]$ which is the same as $f(z_1,\cdots,z_s)$.

Since $R(z_1,\cdots,z_s)\ge 0$ for every $(z_1,\cdots,z_s)\in Z$, we have the inequality (5.2) and its consequence (5.3). The rest of the theorem follows in a straightforward manner. $\Box$

Remark. If $Z_j=\{0,\cdots,n_j\}$, $j=1,\cdots,s$, then (5.3) can be written in the form

$$E[f(\xi_1,\cdots,\xi_s)]\ge\sum_{\substack{i_1+\cdots+i_s\le m\\ 0\le i_j\le n_j, j=1,\cdots,s}} i_1!\cdots i_s![z_{10},\cdots,z_{1i_1};\cdots;z_{s0},\cdots,z_{si_s};f]S_{i_1\cdots i_s}. \quad (5.6)$$

Theorem 5.2 *Let $I=\{(i_1,\cdots,i_s)|i_j\ge 0,$ integers, $j=1,\cdots,s,n_1-i_1+\cdots+n_s-i_s\le m\}$ and assume that all divided differences of total order $m+1$, of the function f, are nonnegative. Then the following assertion hold.*

(a) *The set of columns $\{a_{i_1\cdots i_s}|(i_1,\cdots,i_s)\in I\}$ is a basis B for the columns of $\tilde{A}$ in problem (1.11).*

(b) *The Lagrange polynomial $L_I(z_1,\cdots,z_s)$, corresponding to the points $\{(z_{1i_1},\cdots,z_{si_s})|(i_1,\cdots,i_s)\in I\}$, is unique and is the following*

$$\begin{aligned}L_I(z_1,\cdots,z_s) \le & \sum_{\substack{i_1+\cdots i_s\ge n_1+\cdots n_s-s\\ 0\le i_j\le n_j, j=1,\cdots,s}}[z_{1n_1},\cdots,z_{1n_1-i_1};\cdots;z_{sn_s},\cdots,z_{sn_s-i_s};f]\\ &\times\prod_{j=1}^{s}\prod_{h=n_j-i_j-1}^{n_j}(z_j-z_{jh}). \quad (5.7)\end{aligned}$$

(c) *If $m+1$ is even, then*

$$f(z_1,\cdots,z_s) \le L_I(z_1,\cdots,z_s), \qquad \text{for } (z_1,\cdots,z_s) \in Z, \tag{5.8}$$

i.e., B is a dual feasible basis in the maximization problem (1.11),

$$E[f(\xi_1,\cdots,\xi_s)] \le E[L_I(\xi_1,\cdots,\xi_s)]. \tag{5.9}$$

If $m+1$ is odd, then the inequalities (5.7) and (5.8) are reversed, i.e., B is a dual feasible basis in the minimization problem (1.11). In either case the expectation inequality is sharp, if B is also a primal feasible basis in problem (1.11).

(d) *If all divided differences of total order $m+1$ are nonpositive, then all assertions hold with the difference that (5.7) and (5.8) hold for $m+1$ odd, and the reversed inequalities hold for $m+1$ even.*

Proof. The polynomial (5.6) coincides with f at the points $\{(z_{1i_1},\cdots,z_{si_s}),\ (i_1,\cdots,i_s) \in I\}$, for every f. This proves assertions (a) and (b).

Assertion (c) can be proved in the same way as that of Theorem 5.1. If all divided differences that appear in the suitably defined $R_I(z_1,\cdots,z_s)$ are nonnegative (nonpositive), then still the sign of $R_I(z_1,\cdots,z_s)$ depends on the number of factors that multiply the divided differences in each term. Since all factors are nonpositive for all $(z_1,\cdots,z_s) \in Z$ and there are m factors in each term, the assertion in (c) follows. Assertion (d) is a trivial modification of assertion (c). □

Remark. If $Z_j = \{0,\cdots,n_j\}$, $j=1,\cdots,s$, then the inequality (5.8) can be written in the form

$$E[f(\xi_1,\cdots,\xi_s)] \ge \sum_{\substack{i_1+\cdots+i_s \ge n_1+\cdots n_s - s \\ 0 \le i_j \le n_j, j=1,\cdots,s}} (n_1-i_1)!\cdots(n_s-i_s)![n_1,\cdots,n_1-i_1;\cdots;n_s,\cdots,n_s-i_s;f]S_{n_1-i_1\cdots n_s-i_s}. \tag{5.10}$$

In Theorems 5.1 and 5.2 the lattice of the points determining the interpolation polynomial is pointed at the origin and the point $(n_1,\cdots,n_s)$, respectively. In the sepcial case of $m=1$ it can be pointed arbitrarily, under some further assumption, to derive similar statements.

Theorem 5.3 *Suppose that there exists a continuous convex function $g(z_1,\cdots,z_s)$, $(z_1,\cdots,z_s) \in \mathbb{R}^s$ such that $f(z_1,\cdots,z_s) = g(z_1,\cdots,z_s)$ for $(z_1,\cdots,z_s) \in \mathbf{Z}$. Let $I = \{(i_1+\delta_1,\cdots,i_s+\delta_s)\}$ be a set of points such that $\delta_1,\cdots,\delta_s$ are 0, 1 or -1, the point with $\delta_1=\cdots=\delta_s=0$ is in I and for each $j=1,\cdots,s$, we have $\delta_j \ne 0$ in exactly one point. Then the following assertions hold.*

(a) *The set of vectors $\{a_{i_1\cdots i_s} | (i_1,\cdots,i_s) \in I\}$ is a basis for the columns of $\tilde{A}$ in problem (1.11).*

(b) *The Lagrange polynomial corresponding to the points* $\{(z_{1i_1}, \cdots, z_{si_s}) | (i_1, \cdots, i_s) \in I\}$ *is unique and is the following*

$$L_I(z_1, \cdots, z_s) = f(z_{1i_1}, \cdots, z_{si_s}) + \sum_{j=1}^{s} \frac{f(z_{1i_1}, \cdots, z_{ji_j+\delta_j}, \cdots, z_s) - f(z_{1i_1}, \cdots, z_{ji_j-1}, \cdots, z_{si_s})}{\delta_j}(z_j - z_{ji_j}).$$

(c) *We have the inequalities*

$$f(z_1, \cdots, z_s) \leq L_I(z_1, \cdots, z_s), \qquad \text{for} \quad (z_1, \cdots, z_s) \in \mathbf{Z}, \tag{5.11}$$

i.e., B is dual feasible in the maximization problem (1.11) and

$$E[f(\xi_1, \cdots, \xi_s)] \leq E[L_I(\xi_1, \cdots, \xi_s)]. \tag{5.12}$$

The last inequality is sharp if B is also a primal feasible basis in problem (1.11).

Proof. Inequality (5.11) is the immediate consequence of the convexity assumption concerning the functions g and f. Inequality (5.12) and the rest of the assertions follow in a simple way. □

Sometimes we can utilize Theorem 2.1 too, in order to obtain lower and upper bounds on $E[f(\xi_1, \cdots, \xi_s)]$. An example is presented in Prékopa (1992b). The disadvantage of the theorem is that the sum in the first term, on the right hand side of (2.14), is not a polynomial. Hence, replacing $\xi_1, \cdots, \xi_s$ for $z_1, \cdots, z_s$, and taking its expectation, the obtained value cannot be experssed in terms of the multivariate moments. The remainder term, however, can be handled easily. In fact, replacing for $I_1, \cdots, I_s$ some dual feasible subscript structures from Section 3.1 (or 3.2), the obtained remainder term will be nonnegative (nonpositive).

6 Multivariate Bonferroni Inequalities

In this section we assume that $Z_j = \{0, \cdots, n_j\}$, $j = 1, \cdots, s$. Defining

$$g(z) = \begin{cases} 0, & \text{if } z = 0 \\ 1, & \text{if } z \geq 1, \end{cases}$$

we easily see that

$$[0, \cdots, i; g] = (-1)^{i-1}\frac{1}{i!}, \qquad \text{for } i \geq 1.$$

Let $f(z_1, \cdots, z_s) = g(z_1) \cdots g(z_s)$, $(z_1, \cdots, z_s) \in \mathbf{Z}$. Then we have

$$[0, \cdots, i_1; \cdots; 0, \cdots, i_s; f] = \prod_{j=1}^{s} [0, \cdots, i_j; g] = \prod_{j=1}^{s} (-1)^{i_j-1}\frac{1}{i_j!} \quad \text{for } i_j \geq 1, j = 1, \cdots s. \tag{6.1}$$

If for at least one j we have $i_j = 0$, then the above divided difference is 0.

Let $A_{j1}, \cdots, A_{jn_j}$, $j = 1, \cdots, s$ be s finite sequences of arbitrary events and let ξ_j designate the number of those, in the jth sequence, that occur. Then $\xi_j \geq 1$ is the same as $\cup_{i=1}^{n_j} A_{ji}$. Now, Theorem 5.1 and relation (6.2) imply

Theorem 6.1 *If $m+1-s$ is even, then we have*

$$P\left(\bigcap_{j=1}^{s}\bigcup_{i=1}^{n_j}A_{ji}\right)\geq\sum_{\substack{i_1+\cdots+i_s\leq m\\ 1\leq i_j\leq n_j, j=1,\cdots,s}}(-1)^{i_1+\cdots+i_s-s}S_{i_1\cdots i_s}\tag{6.2}$$

and if $m+1-s$ is odd, then we have

$$P\left(\bigcap_{j=1}^{s}\bigcup_{i=1}^{n_j}A_{ji}\right)\leq\sum_{\substack{i_1+\cdots+i_s\leq m\\ 1\leq i_j\leq n_j, j=1,\cdots,s}}(-1)^{i_1+\cdots+i_s-s}S_{i_1\cdots i_s}.\tag{6.3}$$

For the case of $s=1$ we obtain

$$P\left(\bigcup_{i=1}^{n}A_i\right)\geq\sum_{i=1}^{m}(-1)^{i-1}S_i,$$

if m is even, and

$$P\left(\bigcup_{i=1}^{n}A_i\right)\leq\sum_{i=1}^{m}(-1)^{i-1}S_i,$$

if m is odd. These are the original Bonferroni inequalities (see Bonferroni (1937)).

We can also deduce inequalities for $P(\xi_1=r_1,\cdots,\xi_s=r_s)$ and $P(\xi_1\geq r_1,\cdots,\xi_s\geq r_s)$. In the first case we define

$$f_j(z)=\begin{cases}0, & \text{if } z\neq r_j\\ 1, & \text{if } z=r_j\end{cases}$$

and $f(z_1,\cdots,z_s)=f_1(z_1)\cdots f_s(z_s)$ for $(z_1,\cdots,z_s)\in Z$. By the determinantal form of the univariate divided differences we easily deduce that

$$[0,\cdots,i_j;f_j]=(-1)^{i_j-r_j}\frac{1}{i_j!}\binom{i_j}{r_j}.$$

This implies that

$$[0,\cdots,i_1;\cdots;0,\cdots,i_s;f]=\prod_{j=1}^{s}[0,\cdots,i_j;f_j]=\prod_{j=1}^{s}(-1)^{i_j-r_j}\frac{1}{i_j!}\binom{i_j}{r_j}.\tag{6.4}$$

This is nonnegative if $i_1+\cdots i_s-(r_1+\cdots r_s)$ is even, otherwise it is nonpositive. Hence Theorem 6.1 implies

Theorem 6.2 *If $m+1-(r_1+\cdots+r_s)$ is even, then we have*

$$P(\xi_1=r_1,\cdots,\xi_s=r_s) \geq \sum_{\substack{i_1+\cdots+i_s\leq m \\ r_j\leq i_j\leq n_j, j=1,\cdots,s}} \prod_{j=1}^{s}(-1)^{i_j-r_j}\binom{i_j}{r_j}S_{i_1\cdots i_s} \tag{6.5}$$

and if $m+1-(r_1+\cdots+r_s)$ is odd, then we have

$$P(\xi_1=r_1,\cdots,\xi_s=r_s) \leq \sum_{\substack{i_1+\cdots+i_s\leq m \\ r_j\leq i_j\leq n_j, j=1,\cdots,s}} \prod_{j=1}^{s}(-1)^{i_j-r_j}\binom{i_j}{r_j}S_{i_1\cdots i_s}. \tag{6.6}$$

Finally, in order to obtain inequalities for $P(\xi_1\geq r_1,\cdots,\xi_s\geq r_s)$ we define

$$f_j(z)=\begin{cases}0, & \text{if } z<r_j\\ 1, & \text{if } z\geq r_j\end{cases}$$

and $f(z_1,\cdots,z_s)=f_1(z_1)\cdots f_s(z_s)$ for $(z_1,\cdots,z_s)\in\mathbf{Z}$. Again, using the determinantal form of the univariate divided differences, we get

$$[0,\cdots,i_j;f]=\sum_{h=r_j}^{i_j}(-1)^{i_j-h}\frac{1}{i_j!}\binom{i_j}{h}.$$

On the other hand, we have the combinatorial identity

$$\sum_{h=r}^{i}(-1)^{i-h}\binom{i}{h}=(-1)^{i-r}\binom{i-1}{r-1}.$$

Thus, we have the formula for the multivariate divided differences

$$[0,\cdots,i_1;\cdots;0,\cdots,i_s;f]=\prod_{j=1}^{s}[0,\cdots,i_j;f_j]=\prod_{j=1}^{s}(-1)^{i_j-r_j}\frac{1}{i_j!}\binom{i_j-1}{r_j-1}. \tag{6.7}$$

Theorem 6.1 and equation (6.7) imply

Theorem 6.3 *If $m+1-(r_1+\cdots+r_s)$ is even, then we have the inequality*

$$P(\xi_1\geq r_1,\cdots,\xi_s\geq r_s) \geq \sum_{\substack{i_1+\cdots+i_s\leq m \\ r_j\leq i_j\leq n_j, j=1,\cdots,s}} \prod_{j=1}^{s}(-1)^{i_j-r_j}\binom{i_j-1}{r_j-1}S_{i_1\cdots i_s} \tag{6.8}$$

and if $m+1-(r_1+\cdots+r_s)$ is odd, then we have the inequality

$$P(\xi_1\geq r_1,\cdots,\xi_s\geq r_s) \leq \sum_{\substack{i_1+\cdots+i_s\leq m \\ r_j\leq i_j\leq n_j, j=1,\cdots,s}} \prod_{j=1}^{s}(-1)^{i_j-r_j}\binom{i_j-1}{r_j-1}S_{i_1\cdots i_s} \tag{6.9}$$

The inequalities (6.2), (6.3), (6.4), (6.6), (6.7) and (6.8) have been derived first by Meyer (1969).

7 Algorithmic Bounds and Numerical Examples

The significance of the knowledge of a dual feasible basis is twofold. First, we can immediately present bound for the optimum value of the liner programming problem we are dealing with. Second, starting from this basis, we have an algorithmic tool with the aid of which we can improve on the bound or obtain the best possible bound. This tool is the dual method of linear programming, due to Lemke (1954). For a short and elegant description of it see Prékopa (1992a).

Given a linear programming problem

$$\begin{aligned} &\min(\max) \quad \mathbf{c}^T\mathbf{x} \\ &\text{subject to} \\ &\quad A\mathbf{x} = \mathbf{b} \\ &\quad \mathbf{x} \geq \mathbf{0}, \end{aligned} \tag{7.1}$$

where A is an $m \times n$ matrix ($m \leq n$), assumed to be of full rank, any basis B is a nonsingular $m \times m$ part of A.

We say that B is feasible or primal feasible if the solution of the equation $B\mathbf{x}_B = \mathbf{b}$ produces $\mathbf{x}_B \geq \mathbf{0}$.

Let I or I_B designate the set of subscripts of those columns of A which are in the basis. Further, let $\mathbf{c}_B$ designate the vector of components c_i, $i \in I$, arranged in the same order as they are in $\mathbf{c}$.

The basis B is said to be dual feasible if the solution of the equation $\mathbf{y}^T B = \mathbf{c}_B^T$ satisfies the constraints of the dual of problem (7.1):

$$\begin{aligned} &\max(\min) \quad \mathbf{b}^T\mathbf{y} \\ &\text{subject to} \\ &\quad A^T\mathbf{y} \leq (\geq)\mathbf{c}. \end{aligned} \tag{7.2}$$

If B is both primal and dual feasible, then it is optimal.

Let $A = (a_1, \cdots, a_n)$, $\mathbf{c}^T = (c_1, \cdots, c_n)$. With these notations the dual feasibility of B can be formulated as follows:

$$\mathbf{c}_B^T B^{-1} a_h \leq (\geq) c_h, \qquad h = 1, \cdots, n. \tag{7.3}$$

For $h \in I$ equality holds in (7.3).

The dual method of linear programming starts from a dual feasible basis B. Then the following steps are performed. We assume the problem is a minimization problem.

Step 1. Check if $B^{-1}\mathbf{b} \geq \mathbf{0}$, i.e., the basis B is primal feasible. If yes, then stop, optimal basis has been found. Otherwise go to Step 2.

Step 2. Pick any negative component of $B^{-1}\mathbf{b}$. If it is the ith one, then delete the ith vector from B. Go to Step 3.

Step 3. Determine the incoming vector maintaining dual feasibility of the basis and making the objective function value nondecreasing. Go to Step 1.

Step 3 is usually costly. In case of the univariate discrete moment problems (see Prékopa (1990b)) the structure of the dual feasible bases have been found and Step 3 can be carried out by performing simple combinatorial search.

In case of the multivariate discrete moment problems we have only a few dual feasible basis structures and we cannot spare Step 3 when solving the problem to obtain the best possible bound.

Still, the availability of an initial dual feasible basis is of great help. We can save the time needed to execute the first phase in a two-phase solution method that is roughly 50% of the time needed to solve the LP. In addition, since moment problems are numerically very sensitive, the knowledge of an initial dual feasible basis increases numerical stability.

The dual method, as applied to these problems, has many other features. For example, we may have more detailed information about the possible values of the random vector $(\xi_1, \cdots, \xi_s)$, i.e., we may know that some of the values in the set $Z = Z_1 \times \cdots \times Z_s$ are not possible, in other words, have probability 0. Information of this type has not been exploited so far in former sections of the paper. The dual method, however, allows to take such information into account, in a trivial way. In fact, we simply have to delete those columns from the problem that are multiplied by the probabilities known to be 0. The basis remains dual feasible with respect to the new problem. This way we even improve on the bound.

Below we present just one small numerical example for illustration.

Let $n_1 = n_2 = 9$, $m_1 = m_2 = 3$. The following power moments have been obtained from the uniform distribution: $p_{i_1 i_2} = 1/100$ for each $0 \le i_1, i_2 \le 9$:

$$\begin{aligned}
&\mu_{00} = 1, && \mu_{10} = 4.5, && \mu_{20} = 28.5, && \mu_{30} = 202.5\\
&\mu_{01} = 4.5, && \mu_{11} = 20.25, && \mu_{21} = 128.25, &&\\
&\mu_{02} = 28.5, && \mu_{12} = 128.25, && &&\\
&\mu_{03} = 202.5. && && &&
\end{aligned}$$

We want to obtain the sharp lower bound for $P(\xi_1 \ge 1, \xi_2 \ge 1)$. We start from the dual feasible basis with subscript set $I = \{(0,0),(1,0),(2,0),(0,1),(1,1),(0,2)\}$.

As optimal solution, for the minimization problem (1.11), we obtain

$$\begin{aligned}
&p_{40} = 0.075, && p_{50} = 0.125, && p_{90} = 0,\\
&p_{04} = 0.175, && p_{94} = 0.125, && p_{45} = 0.225,\\
&p_{95} = 0.075, && p_{09} = 0.025, && p_{59} = 0.175,\\
&p_{99} = 0, && \text{and all other } && p_{i_1 i_2} = 0.
\end{aligned}$$

The value of the objective function is the sum of those $p_{i_1 i_2}$ probabilities for which we have $i_1 \ge 1$, $i_2 \ge 1$. This sum equals 0.6. Thus, the result is

$$P(\xi_1 \ge 1,\, \xi_2 \ge 1) \ge 0.6.$$

Note that the true probability is the sum of those $p_{i_1 i_2} = 1/100$, for which $i_1 \geq 1, i_2 \geq 1$. This number is 0.81.

Suppose now that we have the information concerning ξ_1 and ξ_2 that $\xi_1 + \xi_2 \leq 12$. This means that the set of possible values of the random vector (ξ_1, ξ_2) is only a subset of the set $\{(i,j)|0 \leq i \leq 9,\ 0 \leq j \leq 9\}$. Thus, we may delete those columns, variables and objective function coefficients from problem (1.11) which correspond to (i,j) with $i+j > 12$. Solving the restricted problem, the optimal solution is

$$\begin{aligned}
&p_{30} = 0.11393, \quad p_{03} = 0.09749, \quad p_{04} = 0.08732, \\
&p_{45} = 0.23637, \quad p_{55} = 0.11857, \quad p_{56} = 0.00985, \\
&p_{97} = 0.12531, \quad p_{09} = 0.05191, \quad p_{49} = 0.10666, \\
&p_{59} = 0.05259, \quad \text{and all other } p_{i_1 i_2} = 0.
\end{aligned}$$

The value of the objective function is $p_{45}+p_{55}+p_{56}+p_{97}+p_{49}+p_{59} = 0.64935$. This improves on the former lower bound that is 0.6.

In case of $m = 3$ the Bonferroni inequality (6.3) produces the irrealistic result:

$$P\left(\bigcap_{i=1}^{2} \bigcup_{i=0}^{9} A_{ji}\right) \geq S_{11} - S_{12} - S_{21} = 20.25 - 54 - 54 = -87.75.$$

This number is, at the same time, the value of the objective function in case of the initial dual feasible basis.

For further numerical examples see Prékopa (1992b).

References

[1] Bonferroni, C. E. (1937). *Teoria Statistica Delle Classi e Calcolo Delle Probabilità.* Volume in onore di Riccardo Dalla Volta, Università di Firenze, 1–62.

[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. *Math. Oper. Res.* **14**, 317–342.

[3] Chung, K.C. and T.H. Yao (1977). On Lattices Admitting Unique Lagrange Interpolations. *SIAM Journal on Numerical Analysis* **14**, 735–743.

[4] Dawson, D. and A. Sankoff (1967). A Inequality for Probabilities. *Proc. Am. Math. Soc.* **18**, 504–507.

[5] Dulá, J. H. (1986). *Bounds on the Expectations of Convex Functions.* Ph. D. Dissertation. Dept. of Operations and Industrial Engineering, The University of Michigan, Ann Arbor, Michigan.

[6] Galambos, J. (1987). *The Asymptotic Theory of Extreme Order Statistics* (2nd ed.). Krieger, Malabar, Florida.

[7] Galambos, J. and Y. Xu (1991). Some Optimal Bivariate Bonferroni-Type Bounds. Presented at the 868th Meeting of the AMS, Temple University, Philadelphia, PA, October 12–13, 1991.

[8] Horn, R.A. and C.R. Johnson (1991). *Topics in Matrix Analysis.* Cambridge University Press, New York.

[9] Jordan, C. (1947). *Calculus of Finite Differences.* Chelsea Publishing Company, New York.

[10] Kall, P. (1991). An Upper Bound for SLP Using First and Total Second Moments. *Annals of Operations Research* **30**, 267–276.

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

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

[13] Kemperman, J.H.B. and M. Skibinsky (1992). Convariance Spaces for Measures on Polyhedral Sets. In: *Stochastic Inequalities* (M. Shaked and Y.L. Tong, eds.), Institute of Mathematical Statistics, Lecture Notes — Monograph Series, Vol 22, 182–195.

[14] Klein-Haneveld, W. (1992). Multilinear Approximation on Rectangles and the Related Moment Problem. *Mathematics of Operations Research*, to appear.

[15] Kwerel, S. M. (1975). Bounds on the Probability of a Union and Intersection of m Events. *Adv. Appl. Prob.* **7**, 431–448.

[16] Lemke, C. E. (1954). The Dual Method for Solving the Linear Programming Problem. *Naval Research Logistics Quarterly* **1**, 36–47.

[17] Lemke, C.E. (1954). The Dual Method for Solving the Linear Programming Problem. *Naval Research Logistics Quarterly* **1**, 36–40.

[18] Meyer, R.M. (1969). A Note on a Multivariate Form of Bonferroni's Inequalities. *Annals of Mathematical Statistics* **40**, 692–693.

[19] Nicolaides, R.A. (1972). On a Class of Finite Elements Generated by Lagrange Interpolation. *SIAM Journal on Numerical Analysis* **9**, 435–445.

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

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

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

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

[24] Prékopa, A. (1992a). A Very Short Introduction to Linear Programming. *RUTCOR Lecture Notes*, RLN # 2-92, July 1992.

[25] Prékopa, A. (1992b). 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, Vol 22, 309–331.

[26] Råde, L. (1989). Expected Time Failure of Reliability Systems. *Mathematical Scientist* **14**, 24–37.

[27] Riordan, J. (1968). *Combinatorial Identities.* Wiley, New York.

[28] Sathe, Y. S., M. Pradhan and S. P. Shah (1980). Inequalities for the Probability of the Occurence of at least m out of n Events. *J. Appl. Prob.* **17**, 1127–1132.

[29] Wheeler, J. C. and R. G. Gordon (1970). Bounds for Averages Using Moment Constraints. In: *The Padé Approximant in Theoretical Physics* (G. A. Baker, Jr. and J. L. Gammel, eds.), Academic Press, New York and London. Chapter 3, 99–127.