RUTCOR
RESEARCH
REPORT

Lower and Upper Bounds on Probabilities of Boolean Functions of Events

András Prékopa[a] Béla Vizvári[b] Gábor Regős[c]

RRR 36-95, September 1995. Revised May 1996

RUTCOR • Rutgers Center for Operations Research • Rutgers University • P.O. Box 5062 • New Brunswick New Jersey • 08903-5062
Telephone: 908-445-3804
Telefax: 908-445-5472
Email: rrr@rutcor.rutgers.edu
http://rutcor.rutgers.edu/~rrr

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

[b]Visitor of RUTCOR, vizvari@rutcor.rutgers.edu, and Dept. of Operations Research, Eötvös Lóránd University of Budapest, H-1088 Budapest, Múzeum krt. 6-8., vizvari@cs.elte.hu

[c]Dept. of Operations Research, Eötvös Lóránd University of Budapest, H-1088 Budapest, Múzeum krt. 6-8., regos@cs.elte.hu

Rutcor Research Report
RRR 36-95, September 1995. Revised May 1996

Lower and Upper Bounds on Probabilities of Boolean Functions of Events

András Prékopa Béla Vizvári Gábor Regős

Abstract. Given n arbitrary events in a probability space, we assume that the individual probabilities as well as the joint probabilities of up to m events are known, where $m < n$. Using this information we give lower and upper bounds for some Boolean functions of events (e.g., at least one, or exactly one occurs). The available information can be cast in the form of a linear equality-inequality system. The bounds are obtained in such a way that we formulate linear objective functions and find feasible solutions to the dual problem.

Acknowledgements: This research was supported by the Air Force, the grant numbers are AFORS-89-05128, and F49620-93-0041.

1 Introduction

Let $A_1, ..., A_n$ be arbitrary events in some probability space, and introduce the notations

$$\begin{aligned} P(A_{i_1} \cap ... \cap A_{i_k}) &= p_{i_1...i_k}, \quad 1 \le i_1 < ... < i_k \le n, \\ S_k &= \sum_{1 \le i_1 < ... < i_k \le n} p_{i_1...i_k}, \quad k = 1, ..., n. \end{aligned}$$

Let $S_0 = 1$, by definition. If ν designates the number of those events (among $A_1, ..., A_n$) which occur, then we have the relation:

$$E\left[\binom{\nu}{k}\right] = S_k, \quad k = 0, ..., n. \tag{1}$$

The equations (1) can be written in the more detailed form

$$\sum_{i=0}^{n} \binom{i}{k} v_i = S_k, \quad k = 0, ..., n,$$

where $v_i = P(\nu = i)$, $i = 0, ..., n$.

The values (1) are called the binomial moments of ν. If we know all binomial moments of ν, then the probabilities $v_0, ..., v_n$, and also the value of any linear functional acting on the probability distribution $v_0, ..., v_n$, can be determined. If, however, we only know $S_1, ..., S_m$, where $m < n$, then linear programming problems provide us with lower and upper bounds on the true value of this functional. We formulate two closely related types of linear programming problems:

$$\begin{aligned} & \min(\max) \sum_{i=1}^{n} c_i x_i \\ & \text{subject to} \\ \sum_{i=1}^{n} \binom{i}{k} x_i &= S_k, \quad k = 1, ..., m \\ & x_i \ge 0, \quad i = 1, ..., n, \end{aligned} \tag{2}$$

and

$$\begin{aligned} & \min(\max) \sum_{i=0}^{n} c_i x_i \\ & \text{subject to} \\ \sum_{i=0}^{n} \binom{i}{k} x_i &= S_k, \quad k = 0, ..., m \\ & x_i \ge 0, \quad i = 0, ..., n. \end{aligned} \tag{3}$$

These provide us with lower and upper bounds on the linear functionals $\sum_{i=1}^{n} c_i v_i$, and $\sum_{i=0}^{n} c_i v_i$, respectively. Note that the first constraint in problem (3) does not appear in problem (2). The following objective functions are of particular interest:

$$c_0 = c_1 = ... = c_{r-1} = 0, \quad c_r = ... = c_n = 1 \tag{4}$$

$$c_r = 1, \quad c_i = 0 \text{ for } i \neq r. \tag{5}$$

If we use the objective function coefficient (4) in the linear programs (2), (3), then the optimum values provide us with lower and upper bounds for the probability that at least r out of n events occur. The objective function with coefficients (5) provides us with bounds for the probability that exactly r events occur. Any dual feasible basis of any of the problems (2) and (3) provides us with a bound. The best bound corresponds to the optimal basis which is both primal and dual feasible, and is called sharp.

Lower and upper bounds for the probability that at least one out of n events occurs, based on the knowledge of $S_1, ..., S_m$, were found by Bonferroni (1937). These bounds are not sharp. For the case of $m = 2$, sharp lower bound for the probability that at least one out of n events occurs was proposed by Dawson and Sankoff (1967). For the case of $m \leq 3$, Kwerel (1975a,b) has obtained sharp lower and upper bounds. He applied linear programming theory in his proofs. For the case of $m = 2$ other results are due to Galambos (1977), and Sathe, Pradhan and Shah (1980). For a general m, the linear programs with objective functions (4), and (5) have been formulated and analyzed by Prékopa (1988, 1989). He also presented simple dual type algorithms to solve the problems. Boros and Prékopa (1989) utilized the results and presented closed form bounds.

Problems (2), and (3) use the probabilities $p_{i_1...i_k}$ in aggregated forms, i.e., $S_1, ..., S_m$ are used rather then the probabilities in these sums. This way we trade information for simplicity and size reduction of the problems. We call (2) and (3) aggregated problems.

The linear programs which make us possible to use the probabilities $p_{i_1...i_k}$, $1 \leq i_1 < ... < i_k \leq n$ individually, will be called disaggregated, and can be formulated as follows. Let D_1 be the $n \times 2^n - 1$ matrix, the columns of which are formed by all 0,1-component vectors which are different from the zero vector.

Let us call the collection of those columns of D_1, which have exactly k components equal to 1, the k^{th} block, $1 \leq k \leq n$. Assume that the columns in D_1 are arranged in such a way that first come all vectors in the first block, then all those in the second block, etc. Within each block the vectors are assumed to be arranged in a lexicographic order, where the 1's precede the 0's. Let $d_1, ..., d_n$ designate the rows of D_1, and define the matrix D_k, $2 \le k \le m$, as the collection of all rows of the form: $d_{i_1}...d_{i_k}$, where the product of the rows $d_{i_1}, ..., d_{i_k}$ is taken component-wise. Assume that the rows in D_k are arranged in such a way that the row subscripts $(i_1, ..., i_k)$ admit a lexicographic ordering, where smaller numbers precede larger ones. Let

$$A = \begin{pmatrix} D_1 \\ \cdot \\ \cdot \\ \cdot \\ D_m \end{pmatrix}.$$

In addition, we define the matrix $\hat{A}$ by

$$\hat{A} = \begin{pmatrix} 1 & \mathbf{1}^T \\ 0 & D_1 \\ \cdot & \cdot \\ \cdot & \cdot \\ \cdot & \cdot \\ 0 & D_m \end{pmatrix},$$

where $\mathbf{1}$ is the $2^n - 1$-component vector, all components of which are 1, and the zeros in the first column mean zero vectors of the same sizes as the numbers of rows in the corresponding D_i matrices.

Let $p^T = (p_{i_1...i_k},\ 1 \le i_1 < ... < i_k \le n,\ k = 1, ..., m)$, where the order of the components follow the order of the rows in A, and $\hat{p}^T = (1, p^T)$. The disaggregated problems are:

$$\begin{aligned} \min(\max) f^T x \\ \text{subject to} \\ Ax &= p \\ x \ge 0, \end{aligned} \tag{6}$$

and

$$\begin{aligned} \min(\max) \hat{f}^T \hat{x} \\ \text{subject to} \\ \hat{A}\hat{x} &= \hat{p} \\ \hat{x} &\ge 0, \end{aligned} \tag{7}$$

where $\hat{f}^T = (f_0, f^T),\ \hat{x}^T = (x_0, x^T)$.

The duals of the above problems are:

$$\begin{aligned} \max(\min) p^T y & \\ \text{subject to} & \\ A^T y \ \leq (\geq)\, f, & \end{aligned} \tag{8}$$

and

$$\begin{aligned} \max(\min) \hat{p}^T \hat{y} & \\ \text{subject to} & \\ \hat{A}^T \hat{y} \ \leq (\geq)\, \hat{f}, & \end{aligned} \tag{9}$$

where $\hat{y}^T = (y_0, y^T)$. Since the dual vector y multiplies the vector p in problem (8), it is appropriate to designate the components of y by $y_{i_1 \dots i_k}$, $1 \leq i_1 < .. < i_k \leq n, \ k = 1, \dots, m$.

If we construct bounds for $P(A_1 \cup \dots \cup A_n)$, then we should take $f^T = (1, \dots, 1)$, and $\hat{f}^T = (0, f^T)$ in the above problems. In this case the more detailed form of problems (8) is the following

$$\begin{gathered} \max(\min) \sum_{k=1}^{m} \sum_{1 \leq i_1 < \dots < i_k \leq n} p_{i_1 \dots i_k} y_{i_1 \dots i_k} \\ \text{subject to} \\ \sum_{k=1}^{m} \sum_{1 \leq i_1 < \dots < i_k \leq n} y_{i_1 \dots i_k} \ \leq (\geq)\, 1. \end{gathered} \tag{10}$$

The more detailed form of problem (9) is the following

$$\begin{gathered} \max(\min) \left\{ y_0 \ + \sum_{k=1}^{m} \sum_{1 \leq i_1 < \dots < i_k \leq n} p_{i_1 \dots i_k} y_{i_1 \dots i_k} \right\} \\ \text{subject to} \\ y_0 \ + \sum_{k=1}^{m} \sum_{1 \leq i_1 < \dots < i_k \leq n} y_{i_1 \dots i_k} \ \leq (\geq)\, 1. \end{gathered} \tag{11}$$

The above probability approximation scheme was first proposed by George Boole (1854). A detailed account on it was presented by Hailperin (1956). Kounias and Marin (1976) made use of problem (11) to generate bounds for the case of $m = 2$.

Concerning problems (6), (7), (8) and (9) the following objective functions are of particular interest:

$$f_i \ = \begin{cases} 1 & \text{if i corresponds to a column in } D_1 \text{ which has at least r 1's} \\ 0 & \text{otherwise,} \end{cases} \tag{12}$$

and

$$f_i = \begin{cases} 1 & \text{if i corresponds to a column in } D_1 \text{ which has exactly r 1's} \\ 0 & \text{otherwise.} \end{cases} \quad (13)$$

If we take the objective function given by (12), then the optimum value of the minimization (maximization) problem (7) gives lower(upper) bound for the probability that at least r out of the n events occur. If we take the objective function given by (13), then the optimum value of the minimization(maximization) problem (7) gives lower(upper) bound for the probability that exactly r out of the n events occur. Any dual feasible basis of any of the above problems provides us with a bound. The sharp (best) bounds correspond to optimal bases.

In case of the objective function (12), $r = 1$, the optimum values of the minimization problems (6) and (7) are the same. The optimum values of the maximization problems (6) and (7) are the same provided that the optimum value corresponding to (6) is smaller than or equal to 1. Otherwise, they are different but in that case we should take 1 as the upper bound.

2 Connection Between the Aggregated and Disaggregated Problems

Any feasible solution of problem (6) gives rise, in a natural was to a feasible solution of problem (2). Similarly, any feasible solution of problem (7) gives rise to a feasible solution of problem (3).

Conversely, any feasible solution of the aggregated problem (2) or (3) gives rise to a feasible solutions of the corresponding disaggregated problem. In fact, we obtain problem (6) or (7) from problem (2) or (3) in such a way that we split rows and columns. Splitting a column in the aggregated problem means its representation as a sum of columns taken from the corresponding disaggregated problem.

Another question is that which bases in the aggregated problem produce bases in the disaggregated problem. Consider problem (2) for the case $m = 2$. Then, in the corresponding disaggregated problem we have $n + \binom{n}{2}$ rows. The i^{th} and j^{th} columns in problem (2) split into $\binom{n}{i}$ and $\binom{n}{j}$ columns, respectively. A necessary condition that these columns form a basis in problem (6) is that $\binom{n}{i} + \binom{n}{j} = n + \binom{n}{2}$, where $i < j$. This condition holds if $i = 1$ and $j = 2$, or $i = n-2$ and $j = n-1$. On the other hand these are in fact bases in problem (6), as it is easy to see.

The structures of the dual feasible bases of problems (2) and (3) have been discovered by Prékopa (1988, 1990a, 1990b) for the cases of the objective functions (4) and (5) and some others, too. We recall one theorem of this kind.

Theorem 2.1 *Let $a_1, ..., a_n$ designate the columns of the matrix of problem (2), $I \subset \{1, ..., n\}$, $\mid I \mid = m$, and assume that the objective function coefficients are: $c_1 = ... = c_n = 1$. Then, $\{a_i,\ i \in I\}$ a dual feasible basis if and only if I has the structure:*

	m even	*m odd*
min *problem*	$i, i+1, ..., j, j+1$	$i, i+1, ..., j, j+1, n$
max *problem*	$1, i, i+1, ..., j, j+1, n$	$1, i, i+1, ..., j, j+1.$

$\square$

In view of this theorem, the first $n + \binom{n}{2}$ columns of the matrix of problem (6) form a dual feasible basis. Similarly, the $n + \binom{n}{2}$ columns in the second to the last, and third to the last blocks of problem (6) form a dual feasible basis. The corresponding dual vectors can be computed from the equations produced by the aggregated problem:

$$(y_1, y_2)(a_1, a_2) \;=\; (1, 1),$$

and

$$(y_1, y_2)(a_{n-2}, a_{n-1}) \;=\; (1, 1),$$

respectively. The detailed forms of these equations are:

$$\begin{aligned} y_1 \qquad\quad &= 1 \\ 2y_1 + y_2 &= 1 \end{aligned}$$

and

$$\begin{aligned} (n-2)y_1 + \binom{n-2}{2} y_2 &= 1 \\ (n-1)y_1 + \binom{n-1}{2} y_2 &= 1, \end{aligned}$$

respectively. The first system of equations gives $y_1 = 1,\ y_2 = -1$, and the second one gives: $y_1 = 2/(n-1),\ y_2 = -2/(n-1)(n-2)$. If we assign $y_1 = 1$ to all vectors in the first block and $y_2 = -1$ to all vectors in the second block of problem (6), then we obtain the the dual vector corresponding to the first dual feasible disaggregated basis. Similarly, if we assign $y_1 = 2/(n-1)$ to vectors in the block $n-2$ and $y_2 = -2/(n-2)(n-1)$ to all vectors in block $n-1$ of problem (6), then we obtain the dual vector to the other dual feasible disaggregated basis. The first dual vector gives the Bonferroni lower bound:

$$P(A_1 \cup \ldots \cup A_n) \ \geq \ \sum_{i=1}^{n} p_i \ - \sum_{1 \leq i < j \leq n} p_{ij} \ = \ S_1 - S_2.$$

The second dual vector gives the lower bound

$$P(A_1 \cup \ldots \cup A_n) \ \geq \ \frac{2}{n-1} S_1 \ - \ \frac{2}{(n-2)(n-1)} S_2.$$

The optimal lower bound corresponds to that dual feasible basis (a_i, a_{i+1}) of problem (2), which is also primal feasible. This gives $i = 1 + \lfloor 2S_2/S_1 \rfloor$, and the bound is:

$$P(A_1 \cup \ldots \cup A_n) \ \geq \ \frac{2}{i+1} S_1 \ - \ \frac{2}{i(i+1)} S_2.$$

This formula was first derived by Dawson and Sankoff (1967).

If we want to find the sharp lower bound for $P(A_1 \cup \ldots \cup A_n)$, by the use of problem (6) for $m = 2$, then we may start from any of the above mentioned two dual feasible bases and use the dual method of linear programming, to solve the problem. Since we want lower bound, we have a minimization problem. This suggests that the second dual feasible basis is a better one to serve as an initial dual feasible basis. The reason is that in blocks $n-2$, and $n-1$ the coefficients of the variables are larger, and since it is a minimization problem we may expect that we are closer to the optimal basis than in case of the first dual feasible basis.

Numerical Example. Let $n = 6$, and assume that

$p_1 = 0.30$	$p_2 = 0.35$	$p_3 = 0.55$	$p_4 = 0.40$	$p_5 = 0.30$	$p_6 = 0.35$
$p_{12} = 0.15$	$p_{13} = 0.25$	$p_{14} = 0.05$	$p_{15} = 0.15$	$p_{16} = 0.15$	
$p_{23} = 0.25$	$p_{24} = 0.15$	$p_{25} = 0.15$	$p_{26} = 0.05$		
$p_{34} = 0.25$	$p_{35} = 0.25$	$p_{36} = 0.25$			
$p_{45} = 0.15$	$p_{46} = 0.15$				
$p_{56} = 0.15$					

We used the dual method to solve the minimization problem (6). As initial dual feasible basis we chose the collection of vectors in blocks $n-2 = 4$ and $n-1 = 5$.

These vectors have indices $42, ..., 62$. After twenty iterations an optimal basis was found, the indices of which are:
12, 13, 16, 20, 22, 25, 28, 29, 36, 38, 39, 43, 44, 45, 48, 50, 51, 52, 54, 55, 56.
The basic components of the primal optimal solution are:
$x_{12} = 0.01$, $x_{13} = 0.02$, $x_{16} = 0.10$, $x_{20} = 0.03$, $x_{22} = 0.08$, $x_{25} = 0.01$, $x_{28} = 0.03$, $x_{29} = 0.03$, $x_{36} = 0.05$, $x_{38} = 0.02$, $x_{39} = 0.05$, $x_{43} = 0.02$, $x_{44} = 0.03$, $x_{45} = 0.01$, $x_{48} = 0.01$, $x_{50} = 0.08$, $x_{51} = 0.00$, $x_{52} = 0.06$, $x_{54} = 0.05$, $x_{55} = 0.01$, $x_{56} = 0.01$.
The components of the dual optimal dual solution are:

$$\begin{array}{llllll}
y_1 = 0.4, & y_2 = 0.6, & y_3 = 0.6, & y_4 = 0.8, & y_5 = -0.2, & y_6 = 0.0, \\
y_{12} = -0.2, & y_{13} = -0.2, & y_{14} = -0.4, & y_{15} = 0.2, & y_{16} = -0.2, & \\
y_{23} = -0.2, & y_{24} = -0.4, & y_{25} = 0.0, & y_{26} = -0.2, & & \\
y_{34} = -0.4, & y_{35} = 0.0, & y_{36} = -0.2, & & & \\
y_{45} = 0.2, & y_{46} = -0.4, & & & & \\
y_{56} = 0.0. & & & & &
\end{array}$$

The optimum value equals 0.71.

We generated the right-hand side vector p in problem (6) in such a way that we defined $x^0 = (x_j^0,\ j = 1, ..., 63)^T$, where x_j^0 is different from zero only if $j = 4k,\ k = 1, ..., 15$, and for these j values we made the assignments $x_j^0 = 0.05$; then set $p = Ax^0$. In this case $\sum_{j=1}^{63} x_j^0 = 0.75$ and $x_0^0 = 0.25$. The optimum value of the maximization problem (6) is 1.

3 A Method of Partial Aggregation-Disaggregation to Generate Bounds

Let $E_1, ..., E_s$ be pairwise disjoint nonempty subsets of the set $\{1, ..., n\}$ exhausting the set $\{1, ..., n\}$, and introduce the notation $n_j = \mid E_j \mid,\ j = 1, ..., s$.

Out of the events $A_1, ..., A_n$ we create s event sequences, where the i^{th} one is $\{A_i,\ i \in E_j,\ 1 \leq j \leq s\}$. Any of the events $A_1, ..., A_n$ is contained in one and only one event sequence. For these event sequences we will use the alternative notations:

$$\begin{array}{c} A_{11}, ..., A_{1n_1} \\ \cdots \\ A_{s1}, ..., A_{sn_s}. \end{array} \tag{14}$$

Let ξ_j designate the number of those events in the j^{th} sequence, which occur, and

$$
\begin{aligned}
S_{\alpha_1 \ldots \alpha_s} &= E\left[\binom{\xi_1}{\alpha_1} \ldots \binom{\xi_s}{\alpha_s}\right] \\
& 0 \le \alpha_j \le n_j, \quad j = 1, \ldots, s.
\end{aligned} \tag{15}
$$

We formulate the multivariate binomial moment problem:

$$
\begin{aligned}
\min(\max) \sum_{i_1=0}^{n_1} \ldots \sum_{i_s=0}^{n_s} f_{i_1 \ldots i_s} x_{i_1 \ldots i_s} \\
\text{subject to} \\
\sum_{i_1=0}^{n_1} \ldots \sum_{i_s=0}^{n_s} \binom{i_1}{\alpha_1} \ldots \binom{i_s}{\alpha_s} x_{i_1 \ldots i_s} &= S_{\alpha_1 \ldots \alpha_s} \\
\alpha_j \ge 0, \ j = 1, \ldots, s, \ \alpha_1 + \ldots + \alpha_s \le m \\
\forall i_1, \ldots, i_s : x_{i_1 \ldots i_s} \ge 0.
\end{aligned} \tag{16}
$$

The $S_{\alpha_1 \ldots \alpha_s}$ $(\alpha_1 + \ldots + \alpha_s \le m)$ multivariate binomial moments can be computed from the probabilities $p_{i_1 \ldots i_k}$ $(1 \le i_1 < \ldots < i_k \le m)$. In order to simplify the rule how to do this, assume that $E_1 = \{1, \ldots, n_1\}, \ldots, E_s = \{n_1 + \ldots + n_{s-1} + 1, \ldots, n_1 + \ldots + n_s\}$. Then, we have the equality

$$
S_{\alpha_1 \ldots \alpha_s} = \sum p_{i_{11} \ldots i_{1\alpha_1} \ldots i_{s1} \ldots i_{s\alpha_s}},
$$

where the summation is extended over those indices which satisfy the relations

$$
\begin{gathered}
1 \le i_{11} < \ldots < i_{1\alpha_1} \le n_1 \\
\ldots \\
n_1 + \ldots + n_{s-1} + 1 \le i_{s1} < \ldots < i_{s\alpha_s} \le n_1 + \ldots + n_s.
\end{gathered}
$$

For example, if $n = 6$ and $E_1 = \{1, 2, 3\}$, $E_2 = \{4, 5, 6\}$, then

$$
S_{10} = p_1 + p_2 + p_3, \quad S_{01} = p_4 + p_5 + p_6, \quad S_{20} = p_{12} + p_{13} + p_{23}, \quad S_{02} = p_{45} + p_{46} + p_{56},
$$

$$
\begin{gathered}
S_{11} = p_{14} + p_{15} + p_{16} + p_{24} + p_{25} + p_{26} + p_{34} + p_{35} + p_{36}, \\
S_{21} = p_{124} + p_{125} + p_{126} + p_{134} + p_{135} + p_{136} + p_{234} + p_{235} + p_{236}, \\
S_{22} = p_{1245} + p_{1246} + p_{1256} + p_{1345} + p_{1346} + p_{1356} + p_{2345} + p_{2346} + p_{2356} \\
\text{etc.}
\end{gathered}
$$

We have yet to formulate suitable objective functions for problems (16). These depend on the type of bounds we want to create. Suppose that we want to

create bounds for two types of logical functions of events:
(i) at least r out of $A_1, ..., A_n$ occur, where $r \geq 1$;
(ii) exactly r out of $A_1, ..., A_n$ occur, where $0 \leq r \leq n$.

Then, we formulate the objective functions as follows. In case of (i):

$$\begin{aligned} f_{i_1 \dots i_s} &= 1, \text{ if } i_1 + \dots + i_s \geq r \\ f_{i_1 \dots i_s} &= 0, \text{ if } i_1 + \dots + i_s < r, \end{aligned} \tag{17}$$

and in case of (ii):

$$\begin{aligned} f_{i_1 \dots i_s} &= 1, \text{ if } i_1 + \dots + i_s = r \\ f_{i_1 \dots i_s} &= 0, \text{ if } i_1 + \dots + i_s \neq r. \end{aligned} \tag{18}$$

Problems (16) reduce to problems (2), if $s = 1$, and to problems (7), if $s = n$. Problems (16) are disaggregated counterparts of problems (2), and aggregated counterparts of problems (7). The objective functions (17), and (18) are counterparts of the objective functions (4), (5), and (12), (13), respectively.

Let us introduce the notations:
$P_{(r)}$ = probability that at least r out of $A_1, ..., A_n$ occur;
$P_{[r]}$ = probability that exactly r out of $A_1, ..., A_n$ occur.
Further notations are presented in the following tableau:

optimum value	Type, and problem	Objective function
$L_{(r)}$	min, (7)	(12)
$U_{(r)}$	max, (7)	(12)
$L_{[r]}$	min, (7)	(13)
$U_{[r]}$	max, (7)	(13)
$l_{(r)}$	min, (16)	(17)
$u_{(r)}$	max, (16)	(17)
$l_{[r]}$	min, (16)	(18)
$u_{[r]}$	max, (16)	(18)

By construction, we have the following inequalities:

$$l_{(r)} \leq L_{(r)} \leq P_{(r)} \leq U_{(r)} \leq u_{(r)} \tag{19}$$

$$l_{[r]} \leq L_{[r]} \leq P_{[r]} \leq U_{[r]} \leq u_{[r]}. \tag{20}$$

In fact, the problems with optimum values $l_{(r)}$, and $u_{(r)}$ ($l_{[r]}$, and $u_{[r]}$) are aggregations of problems with optimum values $L_{(r)}$, and $U_{(r)}$ ($L_{[r]}$, and $U_{[r]}$), respectively.

The duals of problems (16) are the following:

$$\max(\min) \sum_{\substack{\alpha_j \geq 0,\ j=1,...,s \\ 1 \leq \alpha_1+...+\alpha_s \leq m}} y_{\alpha_1...\alpha_s} S_{\alpha_1...\alpha_s}$$

$$\text{subject to} \tag{21}$$

$$\sum_{\substack{\alpha_j \geq 0,\ j=1,...,s \\ 1 \leq \alpha_1+...+\alpha_s \leq m}} y_{\alpha_1...\alpha_s} \binom{i_1}{\alpha_1} \cdots \binom{i_s}{\alpha_s} \leq (\geq)\ f_{i_1...i_s}$$

$$0 \ \leq\ i_j \ \leq\ n_j,\ j=1,...,s$$
$$i_1+...+i_s \ \geq\ 1.$$

In the left-hand sides of the constraints of problems (21) there are values of a polynomial of the variables $i_1, ..., i_s$, defined on the lattice points of the set $\times_{j=1}^{s}[0, n_j]$. Replacing z_j for i_j, the m-degree polynomial takes the form

$$P(z_1, ..., z_s) \ = \sum_{\substack{\alpha_j \geq 0,\ j=1,...,s \\ \alpha_1+...+\alpha_s \leq m}} y_{\alpha_1...\alpha_s} \binom{z_1}{\alpha_1} \cdots \binom{z_s}{\alpha_s}. \tag{22}$$

Problems (16) provide us with a method to construct polynomials $P(z_1, ..., z_s)$ for one sided approximation of the function $f_{z_1...z_s}$ which we will also designate by $f(z_1, ..., z_s)$. Any polynomial can be used to create bound, provided that it runs entirely below or above the function $f(z_1, ..., z_s)$.

If this latter condition holds, then the bound can be obtained in such a way that we write up the polynomial in the form of (22), subdivide the set $\{1, ..., n\}$ into pairwise disjoint, nonempty subsets $E_1, ..., E_s$, define the $S_{\alpha_1...\alpha_s}$ accordingly, and then make the following assignment: the constant term, if different from zero, is assigned to $S_{0...0} = 1$ in problem (16), $y_{\alpha_1...\alpha_s}$ is assigned to $S_{\alpha_1...\alpha_s}$ in problem (16) for every $\alpha_1, ..., \alpha_s$ for which $\alpha_j \geq 0,\ j = 1, ..., s,\ \alpha_1 + ... + \alpha_s \leq m$. Then, we form the products of the assigned quantities and the $S_{\alpha_1...\alpha_s}$; the sum of the products is the bound.

4 Construction of Polynomials for One-Sided Approximations

In this section we describe a general method to construct polynomials of the type (22) which satisfy (21).

The method consists of construction of dual feasible bases to problem (16). Each dual feasible basis of problem (16) determines a dual vector satisfying the inequalities (21), hence it determines a polynomial (22), which approximates the function f in a one-sided manner.

First we make a general remark. Suppose that the matrix A of the linear programming problem: $\min c^T x$, subject to $Ax = b,\ x \geq 0$, has rank equal to its number of rows m. Let T be an $m \times m$ nonsingular matrix and formulate the problem: $\min c^T x$, subject to $(TA)x = Tb,\ x \geq 0$. Then a basis is primal (dual) feasible in one of these two problems if and only if it is primal (dual) feasible in the other one. In fact, if $A = (a_1, ..., a_n)$, then we have the relations

$$(TB)^{-1}Tb \ = \ B^{-1}b$$

$$c_k - c_B^T(TB)^{-1}Ta_k \ = \ c_k - c_B^T B^{-1}a_k,$$

which imply the assertion.

Let us associate with problem (16) a multivariate power moment problem in such a way that we replace $i_1^{\alpha_1}, ..., i_s^{\alpha_s}$ for $\binom{i_1}{\alpha_1} ... \binom{i_s}{\alpha_s}$ and the power moment $\mu_{\alpha_1 ... \alpha_s}$ for the binomial moment $S_{\alpha_1 ... \alpha_s}$ on the right-hand side. A single linear transformation takes the column vector in (16): $\left(\binom{i_1}{\alpha_1} ... \binom{i_s}{\alpha_s} :\ \alpha_j \geq 0,\ j = 1, ..., s;\ \alpha_1 + ... + \alpha_s \leq m\right)$, into the vector $(i_1^{\alpha_1} ... i_s^{\alpha_s} :\ \alpha_j \geq 0,\ j = 1, ..., s;\ \alpha_1 + ... + \alpha_s \leq m)$ The same transformation applies to the right-hand sides. The matrix of this transformation is nonsingular (it is also triangular). Thus, the above remark applies, and therefore a basis in the multivariate binomial moment problem is primal (dual) feasible if and only if the corresponding basis in the power moment problem primal (dual) feasible.

In view of the above fact the construction of dual feasible bases can be carried out by the use of the results concerning the multivariate power moment problems. In this respect we will make use of the results of the paper by Prékopa (1993), which is entirely devoted to this problem. We recall a few facts from that paper.

Let us associate the lattice point $(i_1, ..., i_s) \in R^s$ with the vector

$$\left(\begin{pmatrix} i_1 \\ \alpha_1 \end{pmatrix} \cdots \begin{pmatrix} i_s \\ \alpha_s \end{pmatrix} : \ \alpha_j \geq 0, \ j = 1, ..., s; \ \alpha_1 + ... + \alpha_s \leq m \right)$$

of the matrix of the equality constraints of problem (16). Let B_Δ and B^Δ designate the sets of vectors corresponding to the sets of lattice points

$$\{(i_1, ..., i_s) \mid i_j \geq 0, \ j = 1, ..., s; \ i_1 + ... + i_s \leq m\}, \tag{23}$$

and

$$\{(n_1 - i_1, ..., n_s - i_s) \mid i_j \geq 0, \ j = 1, ..., s; \ i_1 + ... + i_s \leq m\}, \tag{24}$$

respectively. In (23), and (24) we assume that $m \leq n_j, \ j = 1, ..., s$. It is shown in Prékopa (1993) that both B_Δ and B^Δ are bases in problem (16). The following theorem summarizes the results in Theorems 5.1 and 5.2 of Prékopa (1993).

Theorem 4.1 *The bases B_Δ and B^Δ are dual feasible bases in the following types of problems (16):*

Case 1: all divided differences of f of total order $m+1$ are nonnegative

	$m+1$ *even*	$m+1$ *odd*
B_Δ	*min*	*min*
B^Δ	*min*	*max.*

Case 2: all divided differences of f of total order $m+1$ are non-positive

	$m+1$ *even*	$m+1$ *odd*
B_Δ	*max*	*max*
B^Δ	*max*	*min.*

□

Any dual feasible basis produces a one-sided approximation for f, hence also a bound. A dual feasible basis in a maximization (minimization) problem produces an upper (lower) bound. If a bound of this type is not satisfactory (e.g. a lower bound may be negative, an upper bound may be greater than 1, or a lower (upper) bound is not close enough to a known upper (lower) bound), then we regard the basis as an initial dual feasible basis, and carry out the solution of the problem by the dual method. This way we obtain the best possible bound, at least for a given subdivision $E_1, ..., E_s$ of the set $\{1, ..., n\}$.

Note that problem (7) has $1 + n + \binom{n}{2} + ... + \binom{n}{m}$ equality constraints and 2^n variables, whereas problem (16) has $\binom{s+m}{m}$ constraints and $(n_1+1)...(n_s+1)$ variables. Thus, problem (16) has a much smaller size than problem (7). For example, if $n = 20,\ s = 2,\ n_1 = n_2 = 10,\ m = 3$, then problem (7) has sizes 1351 and 1,048,576, whereas problem (16) has sizes 10 and 121.

To obtain the best possible bound which can be given by our method, one has to maximize (minimize) the lower (upper) bound with respect to all subdivisions $E_1, ..., E_s$ of the set $\{1, ..., n\}$. In practice we use only a few trial subdivisions, and choose that one which provides us with the best bound.

Next, we consider the objective function (17) for the cases of $r = 1$ and $r = n$. If $r = n$, then the function (17) is the same as the function (18). Thus, if $r = 1$, then we look at

$$f_{i_1,...,i_s} = \begin{cases} 0 & \text{if } (i_1, ..., i_s) = (0, ..., 0) \\ \\ 1 & otherwise, \end{cases} \tag{25}$$

and if $r = n$, then we look at

$$f_{i_1,...,i_s} = \begin{cases} 1 & \text{if } (i_1, ..., i_s) = (n_1, ..., n_s) \\ \\ 0 & otherwise. \end{cases} \tag{26}$$

It is easy to check that all divided differences of any order of the function (26) are nonnegative, and if $m+1$ is even (odd), then all divided differences of the function (25) of total order $m+1$ are non-positive (nonnegative). Combining this with Theorem 4.1, we obtain

Theorem 4.2 *The bases B_Δ and B^Δ are dual feasible bases in the following types of problems (16):*

the objective function is given by (25)

	$m+1$ *even*	$m+1$ *odd*
B_Δ	*max*	*min*
B^Δ	*max*	*max.*

the objective function is given by (26);

	$m+1$ *even*	$m+1$ *odd*
B_Δ	*min*	*min*
B^Δ	*min*	*max.*

$\square$

Note that the problems with objective functions (25), and (26) can be transformed into each other. The optimum value of the problem with objective function (25) is equal to 1-(optimum value of the problem with objective function (26), and $S_{\alpha_1 \dots \alpha_s}$ replaced by $\bar{S}_{\alpha_1 \dots \alpha_s}$). The binomial moments $\bar{S}_{\alpha_1 \dots \alpha_s}$ correspond to the complementary events $\bar{A}_1, ..., \bar{A}_n$ in the same way as $S_{\alpha_1 \dots \alpha_s}$ correspond to $A_1, ..., A_n$.

The polynomials determined by the bases B_Δ, and B^Δ can be taken from Prékopa (1993). They are multivariate Lagrange interpolation polynomials with base points (23) and (24), respectively. We designate them by $L_\Delta(z_1, ..., z_s)$, and $L^\Delta(z_1, ..., z_s)$, respectively, and present them here in Newton's form:

$$L_\Delta(z_1, ..., z_s) = \sum_{\substack{i_1 + ... + i_s \le m \\ 0 \le i_j \le n_j, j=1,...,s}} [0, ..., i_1; ...0, ..., i_s; f] \prod_{j=1}^{s} \prod_{h=0}^{i_j - 1} (z_j - h) \tag{27}$$

and

$$L^\Delta(z_1, ..., z_s) = \sum_{\substack{i_1 + ... + i_s \le m \\ 0 \le i_j \le n_j, j=1,...,s}} [n_1 - i_1, ..., n_1; ...n_s - i_s, ..., n_s; f] \prod_{j=1}^{s} \prod_{h=n_j - i_j + 1}^{n_j - 1} (z_j - h). \tag{28}$$

In case of function (25) we have $L^\Delta(z_1, ..., z_s) \equiv 1$, and

$$L_\Delta(z_1, ..., z_s) = \sum_{1 \le i_1 + ... + i_s \le m} (-1)^{i_1 + ... + i_s - 1} \binom{z_1}{i_1} \dots \binom{z_s}{i_s}. \tag{29}$$

In case of function (26) we have $L_\Delta(z_1, ..., z_s) \equiv 0$, and

$$L^\Delta(z_1, ..., z_s) = 1 + \sum_{\substack{1 \le i_1 + ... + i_s \le m \\ 0 \le i_j \le n_j,\ j=1,...,s}} (-1)^{i_1 + ... + i_s} \binom{n_1 - z_1}{i_1} \dots \binom{n_s - z_s}{i_s}. \tag{30}$$

Theorem 4.2 tells us the following. If f is the function (25) and $L_\Delta(z_1, ..., z_s)$ is the polynomial (29), then

$$L_\Delta(z_1, ..., z_s) \ \geq (\leq) \ f(z_1, ..., z_s), \tag{31}$$

if $m+1$ is even (odd); if $L^\Delta(z_1, ..., z_s)$ is the polynomial (30), then

$$L^\Delta(z_1, ..., z_s) \ \geq \ f(z_1, ..., z_s), \tag{32}$$

no matter if $m+1$ is even, or odd. If f is the function (26), then we have the inequalities

$$L_\Delta(z_1, ..., z_s) \ \leq \ f(z_1, ..., z_s), \tag{33}$$

no matter if $m+1$ is even, or odd, and

$$L^\Delta(z_1, ..., z_s) \ \leq (\geq) \ f(z_1, ..., z_s), \tag{34}$$

if $m+1$ is even (odd).

5 Numerical Examples

Example 1. Let $n = 20,\ n_1 = n_2 = 10,\ m = 3$, and assume that we have obtained the following numbers:

$$S_{01} = S_{10} = 4.5, \quad S_{02} = S_{20} = 12, \quad S_{11} = 20.25, \quad S_{03} = S_{30} = 21, \quad S_{12} = S_{21} = 54.$$

The polynomial (29) takes the form

$$\begin{aligned} L_\Delta(z_1, z_2) \ \equiv \ & z_1 - \binom{z_1}{2} + \binom{z_1}{3} + z_2 - z_1 z_2 \\ & + \binom{z_1}{2} z_2 - \binom{z_2}{2} + z_1 \binom{z_2}{2} + \binom{z_2}{3}, \end{aligned} \tag{35}$$

hence the dual vector corresponding to the basis B_Δ equals:

$$y \ = \ (0\ 1\ \ -1\ 1\ 1\ \ -1\ 1\ \ -1\ 1\ 1)^T. \tag{36}$$

The polynomial (30) takes the form

$$L^\Delta(z_1, z_2) \ \equiv \ 1, \tag{37}$$

hence the dual vector corresponding to the basis B^Δ equals:

$$y \ = \ (1\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0)^T. \tag{38}$$

Note that B_Δ, and B^Δ correspond to the lattice points $\{(0,0), (0,1), (0,2), (0,3), (1,0), (1,1), (1,2), (2,0), (2,1), (3,0)\}$, and $\{(10,7), (10,8), (10,9), (10,10), (9,8), (9,9), (9,10), (8,9), (8,10), (7,10)\}$, respectively.

By (32) we have that $L^\Delta(z_1, z_2) \geq f(z_1, z_2)$, which is a trivial inequality in view of (37). Since $m+1=4$ is even, by (31) we have that $L_\Delta(z_1, z_2) \geq f(z_1, z_2)$ for all (z_1, z_2). Thus, both B^Δ, and B_Δ are dual feasible bases in the maximization problem (16).

The dual vector (36) produces the trivial upper bound $y^T S = 114.75$, where

$$S = (S_{00}, S_{10}, S_{20}, S_{30}, S_{01}, S_{11}, S_{21}, S_{02}, S_{12}, S_{03})^T.$$

The dual vector (38) produces the upper bound $y^T S = 1$, which is at the same time the optimum value of the maximization problem (16), and the sharp upper bound for $P(\cup_{i=1}^{20} A_i)$.

The sharp lower bound is obtained by the solution of the minimization problem (16). We obtained the following optimal solution.

$$x_{00} = 0.11, \quad x_{90} = 0.055556, \quad x_{72} = 0.160714, \quad x_{73} = 0, \quad x_{55} = 0.33,$$

$$x_{36} = 0.208333, \quad x_{28} = 0, \quad x_{09} = 0.075397, \quad x_{10,9} = 0, \quad x_{10,10} = 0.06.$$

This provides us with the lower bound:

$$P(\cup_{i=1}^{20} A_i) \geq 1 - x_{00} = 0.89.$$

The dual vector corresponding to the optimal basis is:

$$y = (0, 0.28, -0.0577, 0.0066, 0.2, -0.04, 0.0044, -0.0222, 0.0022, 0).$$

This determines the polynomial

$$\begin{aligned} L_\Delta(z_1, z_2) &= 0.28 z_1 - 0.0577 \binom{z_1}{2} + 0.0066 \binom{z_1}{3} + 0.2 z_2 - 0.04 z_1 z_2 + \\ &\qquad 0.0044 \binom{z_1}{2} z_2 - 0.0222 \binom{z_2}{2} + 0.0022 z_1 \binom{z_2}{2}, \end{aligned}$$

which satisfies $L(z_1, z_2) \leq f(z_1, z_2)$ for all (z_1, z_2).

Example 2. In this example we consider 40 events for which all binomial moments of order up to 11 have been computed. The 40 events have been subdivided into two 20-element groups and all bivariate binomial moments of total order at most 6 have been computed.

Lower and upper bounds for the probability that at least one out of the 40 events occurs have been computed based on the two sets of data. The bounds are displayed for all lower order binomial moments, too. Thus, we have two sequences of bounds. The bounds in the first sequence are optimum values of problems (2), where the objective function is (4) and $r = 1$. The bounds in the second sequence are optimum values of problems (16), where the objective function is (17) and $r = 1$. The latter problem is a partially disaggregated problem, as compared to problem (2).

The results show that much better bounds can be obtained in the latter case. The bounds obtained from the partially disaggregated problem for $m = 6$ are better than those obtained from the aggregated problem for $m = 11$. The data and the bounds are presented below.

Univariate binomial moments, 40 events

S_0	1.000
S_1	8.164
S_2	54.025
S_3	290.574
S_4	1435.025
S_5	7115.369
S_6	34884.230
S_7	158338.877
S_8	637735.541
S_9	2249527.156
S_{10}	6955762.090
S_{11}	18955303.836

Bivariate binomial moments when the 40 events are subdivided into two 20-element groups

first	second group						
group	0	1	2	3	4	5	6
0	1.00	1.93	4.70	12.19	41.05	127.37	317.72
1	6.23	3.28	31.15	186.89	794.26	2541.64	
2	46.04	31.15	295.90	1775.41	7545.49		
3	216.09	186.89	1775.41	10652.46			
4	724.30	794.26	7545.49				
5	1848.66	2541.64					
6	3739.79						

Bounds based on univariate binomial moments

m	lower bound	upper bound
1	0.20410	1.00000
2	0.57400	1.00000
3	0.63452	1.00000
4	0.67613	1.00000
5	0.77875	1.00000
6	0.78559	0.97028
7	0.79960	0.92088
8	0.80000	0.81438
9	0.80156	0.81185
10	0.80191	0.80671
11	0.80299	0.80638

Bounds based on bivariate binomial moments

m	lower bound	upper bound
1	0.31147	1.00000
2	0.66045	1.00000
3	0.79552	0.91272
4	0.80255	0.83071
5	0.80275	0.80583
6	0.80325	0.80410

Example 3. In this example we consider 40 events (different from those of Example 2) which we subdivide into two 20-element groups. We have computed the

univariate binomial moments of order up to 16 and the bivariate binomial moments of total order up to 8. Based on these, two sequences of bounds have been computed. The bounds in the first sequence are optimum values of problems (2) with objective function(4), where $r = 3$. The bounds in the second sequence are optimum values of problem (16) with objective function (17), where $r = 3$. The results show the usefulness of using problem (16) rather than problem (2) to create bounds. The data and the bounds are presented below.

Univariate binomial moments, 40 events

S_0	1.000
S_1	13.714
S_2	110.413
S_3	603.262
S_4	2658.333
S_5	10803.206
S_6	43678.754
S_7	174426.944
S_8	656045.333
S_9	2238906.635
S_{10}	6817994.468
S_{11}	18451870.302
S_{12}	44444753.675
S_{13}	95592963.786
S_{14}	184250604.611
S_{15}	319293071.452
S_{16}	498850545.349

Bivariate binomial moments when the 40 events are subdivided into two 20-element groups

first			second group						
group	0	1	2	3	4	5	6	7	8
0	1.0	4.4	13.1	36.3	107.1	291.6	671.8	1285.6	2040.5
1	9.3	48.7	142.7	385.1	1106.6	2965.7	6776.9	12911.5	
2	48.5	263.6	865.8	2769.6	9029.3	25931.6	61585.5		
3	160.7	916.2	3546.9	13607.6	49160.0	148050.8			
4	384.1	2360.4	11254.7	50856.2	197398.0				
5	728.1	4966.1	29560.7	151037.0					
6	1183.6	9113.2	65687.4						
7	1742.9	14978.8							
8	2355.9								

Bounds based on univariate binomial moments

m	lower bound	upper bound
1	0.09360	1.00000
2	0.13693	0.99709
3	0.43143	0.86736
4	0.63421	0.84467
5	0.66969	0.84276
6	0.75414	0.84128
7	0.76038	0.84127
8	0.77968	0.84126
9	0.78129	0.84120
10	0.78748	0.83873
11	0.78845	0.83831
12	0.78885	0.83823
13	0.78899	0.83042

Bounds based on bivariate binomial moments

m	lower bound	upper bound
1	0.09360	1.00000
2	0.13693	0.96005
3	0.49703	0.85782
4	0.67984	0.84248
5	0.70577	0.84164
6	0.77074	0.84118
7	0.81049	0.83394
8	0.82896	0.82896

6 Upper Bounds Based on Graph Structures

The bounds given in Section 4 can be interpreted as bounds based on special hypergraphs. Let

$$\sum_{\substack{i_1+...+i_s \le m \\ 0 \le i_j \le n_j, j=1,...,s}} a_{i_1...i_s} \binom{z_1}{i_1} \cdots \binom{z_s}{i_s}$$

be any polynomial. Let $N = \{1,...,n\}$, and $N_j \subset N, \;\; j = 1,...,m$ with $N = \cup_{j=1}^m N_j$, $| N_j | \le n_j$, and $\forall j_1 \ne j_2 : \;\; N_{j_1} \cap N_{j_2} = \emptyset$. Let $E_{i_1...i_s}$ be the set of all subsets of N

containing exactly i_j elements from N_j, $j = 1, ..., s$. Then we define the hypergraph as follows:

$$H = (N, E),$$

where

$$E = \bigcup_{\substack{i_1 + ... + i_s \leq m \\ 0 \leq i_j \leq n_j, j = 1, ..., s}} E_{i_1 ... i_s}.$$

All hyperedges lying in $E_{i_1 ... i_s}$ are weighted by $a_{i_1 ... i_s}$. These weights form a dual feasible vector of problem (6). The scalar product of that and the right-hand side vector of problem (8) provides us with the lower or upper bounds. If $m = 2$, then only nodes and pairs of nodes have weights. To each node we assign the weight 1.

In what follows, the components of an $\binom{n}{2}$-vector are indexed by $11, 12, ..., (n-1)n$ or $1, ..., \binom{n}{2}$ depending on which notation is more convenient. The following lemma is very simple, the proof is omitted.

Lemma 6.1 *The* $\binom{n}{2}$*-component vector* $(1, 1, ..., 1, -w_{12}, ..., -w_{n-1,n})$ *is feasible in the minimization problem (8) if and only if for all* $S \subset N$ *containing at least two elements the inequality* $\sum_{i,j \in S, i<j} w_{ij} \leq \mid S \mid -1$ *holds.* $\square$

Remark. It is easy to see, that any feasible solution to the problem (8) has $w_1, ..., w_n \leq 1$. Lemma 6.1 implies that if $w_1 = ... = w_n \leq 1$ then $\forall 1 \leq i, j \leq n, i \neq j : w_{ij} \leq 1$.

The above lemma can be applied in the following way. Let $G^1(N, E^1)$ and $G^2(N, E^2)$ be two graphs on the vertex set N. Assume that to each $\{i, j\}$, $i, j \in N$, $i \neq j$ a real number w_{ij} is assigned and the following conditions are satisfied:
(i) $E^1 \cap E^2 = \emptyset$,
(ii) if $\{i, j\} \in E^1$ then $w_{ij} = 1$,
(iii) if $\{i, j\} \in E^2$ then $w_{ij} \leq 0$,
(iv) if $\{i, j\} \notin E^1 \cup E^2$ then $w_{ij} = 0$,
(v) is $S \subset N, \mid S \mid \geq 2$, then $\sum_{i,j \in S, i<j} w_{ij} \leq \mid S \mid -1$.

The first bound which can be discussed in the framework of the above lemma is Hunter's bound (see Hunter (1976)). Let $G^1 = T(N, E)$ be any tree, and $G^2 = (N, \emptyset)$. Let

$$w_{ij} = \begin{cases} 1 & \text{if } \{i, j\} \in E \\ 0 & \text{otherwise.} \end{cases}$$

Let $S \subset N, \mid S \mid \geq 2$ As any induced subgraph of a tree is a forest, it follows that $\sum_{i,j \in S, i<j} w_{ij} \leq \mid S \mid -1$. Thus, the conditions of the lemma are satisfied. This means that any tree determines an upper bound, and Hunter's bound is the best among them. Thus, Lemma 6.1 generalizes that bound.

Lemma 6.2 *If* $n \geq 3$ *then*

$$P(A_1 \cup \ldots \cup A_n) \ \leq \ S_1 - \max_{1 \leq k < l \leq n} \sum_{i \neq k,l} (p_{ik} + p_{li}) \ + \ (n-3) p_{kl}. \tag{39}$$

Proof For a fixed k and l let G^1 be the complete bipartite graph connecting k and l with all other vertices, G^2 the edge $\{k, l\}$, and $w_{kl} = 3 - n$. Thus, G^1 has $2n - 4$ edges. Let $S \subset N$ be any subset containing at least two elements. If $k, l \notin S$, then the subgraph of G^1 induced by S has no edge; thus,

$$\sum_{i,j \in S, i<j} w_{ij} \ = \ 0.$$

If S contains only one of k, and l, then

$$\sum_{i,j \in S, i<j} w_{ij} \ = \mid S \mid -1.$$

Finally, if S contains both k, and l, then

$$\sum_{i,j \in S, i<j} w_{ij} \ = \ 2 \mid S \mid \ -4 \ + \ 3\, n \quad \leq \mid S \mid \ - \ 1.$$

Thus, the conditions of Lemma 6.1 are satisfied in all cases. □

As the structure of the graph obeys the above mentioned hypergraph scheme, the polynomial

$$g(z_1, z_2) \ = \ z_1 \ + \ z_2 \ - \ z_1 z_2 \ + \ (n-3) \binom{z_2}{2}$$

satisfies the condition:

$$\forall z_1, z_2 \in Z_+ : \ (z_1, z_2) \neq (0,0), \ z_1 \leq n-2, \ z_2 \leq 2 \text{ implies } g(z_1, z_2) \ \geq \ 1.$$

Lemma 6.3 *Assume that* $n \geq 4$. *Let* $G^1(N, E)$ *be any 1-tree and* $C = \{\{u_1, u_2\}, \{u_2, u_3\}, \ldots, \{u_{k-1}, u_k\}, \{u_k, u_1\}\}$ *be the unique simple circuit contained in* G^1. *Assume that* $k \geq 4$. *Let* s, t *be positive integers with* $1 \leq s < t \leq k$, *and* $t - s \not\equiv \pm 1 (\text{mod } k)$. *Let* G^2 *be a graph containing a single edge such that* $E = \{\{u_s, u_t\}\}$. *Finally, let*

$$w_{ij} = \begin{cases} 1 & \text{if } i < j \text{ and } \{i,j\} \in E^1 \\ -1 & \text{if } i < j \text{ and } \{i,j\} \in E^2 \\ 0 & \text{otherwise.} \end{cases}$$

Then we have the inequality:

$$P(A_1 \cup \ldots \cup A_n) \leq S_1 - \sum_{\{i,j\} \in E^1, i<j} p_{ij} + p_{u_s u_t}. \tag{40}$$

Proof Any subgraph of G^1 induced by a set $S \subset N$ contains at most $\mid S \mid -1$ edges, except C, which contains as many edges as vertices. But even in this case the necessary inequality, given in Lemma 6.1, holds because of the presence of the (-1)-valued edge $\{u_s, u_t\}$, and thus, the conditions of Lemma 6.1 are satisfied. □

The following method is an approximation algorithm for determining the best bound of this type. The algorithm works on the complete graph $K_n(N, E)$. The edge $\{i, j\}$ of K_n is weighted by p_{ij}.

STEP 1: Find a maximum weight spanning tree of K_n, designate it by $T(N, E_T)$.

STEP 2: For any edge $\{i, j\} \in E \setminus E_T$ let $C_{ij} = \{\{u_1^{ij}, u_2^{ij}\}, \ldots, \{u_{l_{ij}-1}^{ij}, u_{l_{ij}}^{ij}\}, \{u_{l_{ij}}^{ij}, u_1^{ij}\}\}$ be the unique simple circuit of the graph $T_{ij}(N, E_T \cup \{i, j\})$, where l_{ij} is the length of C_{ij}. Then, let

$$(i^*, j^*, s^*, t^*) = \text{argmax } \{p_{ij} - p_{st} : l_{ij} \geq 4, 1 \leq s < t \leq l_{ij}, t - s \not\equiv \pm 1 \text{ mod } l_{ij}\}. \tag{41}$$

If $p_{i^*j^*} - p_{s^*t^*} > 0$, then the resulting bound based on the graphs $G = T_{i^*j^*}$ and $G^2(N, \{s^*, t^*\})$ is an improvement on Hunter's bound. The order of the algorithm is $O(n^4)$. In (41) the number of pairs $\{i, j\}$ to be considered is $O(n^2)$. The determination of C_{ij} is equivalent with finding the unique simple path going from i to j in T which can be done in $O(n)$ steps as the sum of the degrees of the vertices in T is $2n - 2$. Then, the selection of the best possible pair $\{s, t\}$ takes $O(n^2)$ operations.

A special case of this type of upper bound is obtained by restricting G^1 to be a Hamiltonian circuit. Let $\mathcal{H}$ be the set of all Hamiltonian circuits. In this way the following upper bound can be obtained:

$$P(A_1 \cup \ldots \cup A_n) \leq S_1 - \max_{H \in \mathcal{H}} \sum_{\{i,j\} \in H, i<j} p_{ij} + \min_{\{s,t\} \notin H} p_{st}. \tag{42}$$

The second term of the right-hand side is equivalent to a traveling salesman problem which is known to be NP-hard. But plenty of good and fast heuristics are available to generate approximate solutions.

7 Comparison with the Aggregated Upper Bound

The optimal value of the maximization problem (2) with objective function (4) and $r = 1$, and $m = 2$ is

$$S_1 - \frac{2}{n} S_2$$

as it is shown by Kwerel (1975a), Sathe, Pradhan and Shah (1980), and Boros-Prékopa (1989). In this section a general lemma is proved, which makes it easy to prove, for a wide class of upper bounds, that they are at least as good as the corresponding aggregated ones.

Lemma 7.1 *Let* $N^1 = \{\{i,j\} \mid \ 1 \leq i,j \leq n; i \neq j\}$, $N^2 = \{1,2,...,r\}$, *where* $r = \left(\begin{array}{c} n \\ 2 \end{array} \right)$, *and assume that the function* $\rho : N^1 \to N^2$ *defines a one-to-one correspondence between the two sets. Let* $w_1, ..., w_r$ *be any real numbers satisfying the equation*

$$\sum_{j=1}^{r} w_j \ = \ n-1.$$

Finally, let π *be any permutation of the set* $\{ 1, ..., n \}$. *Then we have the inequality:*

$$\max_{\pi:} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} w_{\rho(i,j)} p_{\rho(\pi(i),\pi(j))} \ \geq \ \frac{2}{n} S_2. \tag{43}$$

Proof The left-hand side of the inequality is the maximum of some numbers. The average of the same numbers is

$$\frac{1}{n!} \sum_{\pi:} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} w_{\rho(i,j)} p_{\rho(\pi(i),\pi(j))}.$$

The symmetricity of the expression implies that all p's must have the same coefficient in the sum, which is

$$\frac{\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} w_{\rho(i,j)}}{\left(\begin{array}{c} n \\ 2 \end{array} \right)} = \frac{2}{n},$$

as there are $n!$ permutations, the number of w's is $\binom{n}{2}$, and their sum is $n-1$. Thus, the above average is equal to the right-hand side of the inequality. Hence the statement follows immediately. $\square$

Remark The proof does not use any property of the p's, hence the statement holds for any vector $p \in R^{\frac{n(n-1)}{2}}$, and $S_2 = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} p_{ij}$.

In the statement the w's represent a fixed structure and the permutation of the p's ensures that the best sample is chosen which is isomorphic with the fixed structure. For example, the statement that Hunter's bound is at least as good as the aggregated bound, follows from the lemma in two steps. First, the vector w is fixed in such a way that it represents a certain tree structure. The best tree is selected which is isomorphic with this structure. Then, we look at all tree structures and the best of bests gives Hunter's bound. But it follows from the lemma that the best of any tree structure is at least as good as the aggregated bound.

Assume that if the vector $(1, 1, ..., 1, -w_1, ..., -w_r^T) \in R^{r+n}$ represents a dual feasible solution to problem (6). Then Lemma 7.1 is applicable and

$$S_1 - \max_{\pi:} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} w_{\rho(i,j)} p_{\rho(\pi(i),\pi(j))} \ \leq \ S_1 \ - \ \frac{2}{n} S_2.$$

8 Some Special Problem Classes

In this section we show that all upper bounds mentioned in Section 6 have at least one problem class containing for every n a problem such that the upper bound coincides with the actual value of $P(A_1 \cup ... \cup A_n)$.

Lemma 8.1 *If the vector $(1, ..., 1, -w_{12}, ..., -w_{n-1,n})^T \in R^{\frac{n(n+1)}{2}}$ is a feasible solution of the dual of the maximization problem, and for every j the inequality $w_j > 0$ implies that $w_j = 1$, then there is a problem instance such that the upper bound is equal to $P(A_1 \cup ... \cup A_n)$.*

Proof The upper bound is

$$S_1 - \sum_{j=i+1}^{n} w_{ij} p_{ij}.$$

If $A_1, ..., A_n$ are events such that $p_i = 1/n$ $(1 \le i \le n)$ and if $i \ne j$ then

$$p_{ij} = \begin{cases} \frac{1}{n^2} & \text{if } w_{ij} = 1 \\ 0 & \text{if } w_{ij} \le 0, \end{cases} \tag{44}$$

then the following equations hold

$$P(A_1 \cup ... \cup A_n) = 1 - \frac{\sum_{(i,j):i<j, w_{ij}=1} 1}{n^2} = S_1 - \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} w_{ij} p_{ij},$$

i.e., the statement of the lemma is true. These events $A_1, ..., A_n$ can be constructed in the following way. Let $\omega_1, ..., \omega_{n^2}$ be n^2 mutually exclusive events. Let the probability of each ω_i be $1/n^2$. Let $A_1 = \omega_1 \cup ... \cup \omega_n$. If $w_{12} \le 0$, then we define $A_2 = \omega_{n+1} \cup ... \cup \omega_{2n}$, otherwise let $A_2 = \omega_1 \cup \omega_{n+1} \cup \; ... \cup \omega_{2n-1}$. Assume that $A_1, ..., A_{i-1}$ are determined and $A_j = \omega_{k_{j1}} \cup ... \cup \omega_{k_{jn}}$ if $1 \le j \le i-1$, i.e. the set of the indices of ω_l's contained in the composite event A_j is $\{ k_{j1}, ..., k_{jn} \}$. Let $\{ l_1, ..., l_t \} = \{ j : 1 \le j \le i-1,\ w_{ji} = 1 \}$. Then, let $A_i = \omega_{k_{l_1 i}} \cup ... \cup \omega_{k_{l_t i}} \cup \omega_{k_{i-1,n}+1} \cup ... \cup \omega_{k_{i-1,n}+n-t}$. Thus, A_i and A_j are mutually exclusive if $w_{ij} \le 0$, otherwise $p_{ij} = 1/n^2$. □

References

[1] Bonferroni, C.E. (1937), Teoria statistica delle classi e calcolo delle probabilitá. Volume in onore di Riccardo Dalla Volta, Universitá di Firenze, pp. 1-62.

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

[3] Boole, G. (1868), Of Propositions Numerically Definite, in *Transactions of Cambridge Philosophical Society, Part II, XI*, reprinted as Study IV in the next reference.

[4] Boole, G. (1952), Collected Logical Works, *Vol. I. Studies in Logic and Probability*, R. Rhees (ed.), Open Court Publ. Co., LaSalle, Ill.

[5] 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.

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

[7] Galambos, J. (1977), Bonferoni Inequalities, *Annals of Probability*, **5**, 577-581.

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

[9] Kounias, S., and J. Marin (1976), Best Linear Bonferoni Bounds, *SIAM J. on Applied Mathematics*, **30**, 307-323.

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

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

[12] Lemke, C.E. (1954), The Dual Method for Solving the Linear Programming Problem, *Naval Research Logistic Quarterly*, **22**, 978-981.

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

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

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

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

[17] Prékopa, A. (1993), Bounds on Probabilities and Expectations Using Multivariate Moments of Discrete Distributions, *Rutcor Research Report* 34-93.

[18] Sathe, Y.S., M. Pradhan and S.P. Shah (1980), Inequalities for the Probability of the Occurrence of at Least m out of n Events, *Journal of Applied Probability*, **17**, 1127-1132.