NEW RESULTS IN PROBABILITY BOUNDING, A CONVEXITY STATEMENT AND UNIMODALITY OF MULTIVARIATE DISCRETE DISTRIBUTIONS

BY MUNEVVER MINE SUBASI

A dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Operations Research

Written under the direction of

Professor András Prékopa

and approved by

New Brunswick, New Jersey

May, 2008

ABSTRACT OF THE DISSERTATION

New Results in Probability Bounding, A Convexity Statement and Unimodality of Multivariate Discrete Distributions

by Munevver Mine Subasi

Dissertation Director: Professor András Prékopa

This report constitutes the Doctoral Dissertation for Munevver Mine Subasi and consists of three topics: sharp bounds for the probability of the union of events under unimodality condition, convexity theory in probabilistic constrained stochastic programming and strong unimodality of multivariate discrete distributions.

We formulate a linear programming problem for bounding the probability of the union of events, where the probability distribution of the occurrences is supposed to be unimodal with known mode and some of the binomial moments of the events are also known. Using a theorem on combinatorial determinants we fully describe the dual feasible bases of a relaxed problem. We present closed form lower and upper bounds for the probability of the union based on two (not necessarily consecutive) as well as first three binomial moments of the random variables involved. We also present upper bounds for the probability of the union based on first four binomial moments. We give a dual method to find customized algorithmic solution of the LP's involved. Numerical examples show that by the use of our bounding methodology, we obtain tighter bounds for the probability of the union.

Next we investigate the convexity theory of programming under probabilistic constraints. Prékopa [63, 74] has proved that if T is an $r \times n$ random matrix with independent, normally distributed rows such that their covariance matrices are constant multiples of each other, then the function $h(\mathbf{x}) = P(T\mathbf{x} \leq \mathbf{b})$ is quasi-concave in $\mathbb{R}^n$, where $\mathbf{b}$ is a constant vector. We prove that, under same condition, the converse is also true, a special quasi-concavity of $h(\mathbf{x})$ implies the above-mentioned property of the covariance matrices.

Finally we present sufficient conditions that ensure the strong unimodality of a multivariate discrete distribution and give an algorithm to find the maximum of a strongly unimodal multivariate discrete distribution. We also present examples of strongly unimodal multivariate discrete distributions.

Acknowledgements

Though only my name appears on the cover of this dissertation, I could never have reached the heights or explored the depths without the help, support, guidance and efforts of a lot of people. I owe my gratitude to all those who have made this dissertation possible and because of whom my graduate experience has been one that I will cherish forever.

First and foremost I would like to express my sincere gratitude to my Supervisor, Dr. András Prékopa for instilling in me the qualities of being a good scientist. His infectious enthusiasm has been one of the major driving forces through my graduate studies at Rutgers University. His encouragement, patience, guidance and support has helped me to overcome many crisis situations and finish this dissertation.

I am very grateful to him for introducing me the diversity of applications and the existence of endless supply of important problems in operations research. He has helped me to thrive on the joys of mathematics and to develop scientific reasoning powers and ability to think creatively. His devotion to his students and to his profession has inspired me to pursue an academic career. He not only helped me in my intellectual development, but set an excellent example for me to follow as a teacher.

I would like to thank other members of my dissertation committee, Dr. Tayfur Altiok, Dr. Endre Boros, Dr. Vladimir Gurvich, Dr. Sylvia Halasz, and Dr. Andrzej Ruszczyński, for providing many valuable comments and suggestions on the improvement of this dissertation.

I owe a special note of gratitude to Dr. Endre Boros for his various forms of support during my graduate study and for giving me the opportunity to work with him on exciting projects.

I also would like to thank the angels of RUTCOR, Lynn Agre, Katie D'Agosta, Terry Hart, and Clare Smietana, who have been always there to listen. I greatly value their genuine friendship.

I am very grateful to my parents. Their unwavering faith and confidence in my abilities and in me is what has shaped me to be the person I am today.

Finally, I would like to thank my husband, my colleague and my best friend, Ersoy, for his love, patience, endless support and encouragement. Having him as my colleague helped me to gain so much drive and an ability to tackle challenges head on. One of our best experiences was the birth of our son Sean Berkan, who provided an additional and joyful dimension to our life mission. They have been a constant source of love and strength throughout this endeavor.

Dedication

To my son Sean Berkan and my husband Ersoy.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Our work consists of three topics: sharp bounds for the probability of the union of events, convexity theory of probabilistic constrained stochastic programming problems and multivariate discrete unimodality.

The problem to find or approximate probabilities of Boolean functions of events has important applications in reliability theory, statistics, probability theory and stochastic programming. Since the problem is of practical importance it has extensively been studied in literature. Bonferroni [11] proved lower and upper bounds for the probability of the union of events in terms of binomial moments. For the case where only two binomial moments: S_1, S_2 are used, Dawson and Sankoff [22] proved a lower bound that is sharp in the sense that no better inequalities can be given if only S_1 and S_2 are known. Later on other sharp probability bounds have been proved using binomial moments. Prékopa [67, 68, 69, 70] has discovered that the sharp binomial moments based probability inequalities can be formulated as discrete moment problems. Earlier sharp Bonferroni inequalities have been recognized as special cases of binomial moment problems. The central results in this respect are those that concern the structure of the dual feasible bases. Any dual feasible basis provides us with a bound for the optimum value of the objective function. The bound is sharp if the basis is primal feasible as well. The combination of a dual feasible basis structure theorem and the dual method of linear programming is a powerful tool to find the sharp bound for the optimum value of the objective function. Boros and Prékopa [15] presented a variety of closed-form bounds, the derivations of which are based on the above-mentioned dual feasible basis structure theorems.

The paper by E. Subasi, M. Subasi and A. Prékopa [88] is the first, where sharp bounds are presented for the probability of the union under unimodality constraint for the probability distribution of the occurrences. In that paper the discrete moment problem was used, with the above-mentioned additional shape constraint of the distribution for the case where the first two power and binomial moments are known. Bounds for expectations as well as the probabilities of the unions of events were given by the use of formulas. In another paper by E. Subasi, M. Subasi and A. Prékopa, bounding formulas have been obtained for the probabilities that at least r and exactly r out of n events occur, under the same conditions (see [89]).

This dissertation is based on the papers [83, 82, 90]. Chapter 2 is devoted to derive a general theorem in connection with the binomial moment problem with unimodal distributions that characterizes the dual feasible bases of a relaxed version of the problem, further, to present closed form and algorithmic bounds for the probability of the union. The closed form bounds are based on any two binomial moments as well as the first three binomial moments. We also present upper bound for the probability of the union based on the knowledge of first four binomial moments of the random variables involved. For larger m values we give a specially designed dual algorithm that is applicable to cases with consecutive and non-consecutive moments. We finally present numerical examples to show that by the use of our bounding methodology, we obtain tighter bounds for the probability of the union of events.

In Chapter 3 we investigate the convexity theory of the probabilistic constrained stochastic programming problem. An important problem is the convexity of the set of feasible solutions, where in the stochastic (linear) constraints the coefficients are random variables, that is the convexity of the set $D = \{\mathbf{x} \mid h(\mathbf{x}) \geq p\}$, where $h(\mathbf{x}) = P(T\mathbf{x} \leq \xi)$, p, $(0 < p < 1)$, is a fixed probability, T a random matrix and ξ a random vector. The first results in this respect are due to Kataoka [42] and van de Panne and Popp [57] as well as those presented in Prékopa [74]. Prékopa [64] has proved that if T has independent, normally distributed rows such that their covariance matrices are convex multiples of each other, then the function $h(x)$ is quasi-concave. The next important results are due to Henrion [36]. Our contribution is the theorem where we prove that under some conditions the quasi concavity of the constraining function implies that the covariance matrices of the rows of T are constant multiples of a covariance matrix for the case where expectations of the random row vectors are supposed to be zero.

The last topic, multivariate discrete unimodality, is studied in Chapter 4. Unimodality of a continuous multivariate distribution means that its p.d.f. is quasi-concave. Here we are interested in unimodal multivariate discrete distributions. In both cases the distribution is called strongly unimodal if it is logconcave (see later the definitions). While continuous unimodal multivariate distributions enjoy a number of useful properties (see, e.g., [74]), many of them do not carry over to the discrete case. For example the convolution of two logconcave multivariate probability densities is logconcave, but the convolution of two logconcave multivariate probability functions is not logconcave in general.

Classical papers on unimodality for discrete distributions are Fekete [29] and Barndorff-Neilsen [5]. Pedersen [58] gave sufficient conditions for a bivariate discrete distribution to be strongly unimodal. He also proved that the trinomial distribution is logconcave and the convolution of any finite number of these distributions with possibly different parameter sets is also logconcave.

Favati and Tardella [28] introduced a notion of integer convexity. They analyzed some connections between the convexity of a function on $\mathbb{R}^n$ and the integer convexity of its restriction to $\mathbb{Z}^n$. They also presented a polynomial time algorithm to find the minimum of a submodular integrally convex function. A further paper in this respect is due to Murota [56]. He developed a theory of discrete convex analysis for integer-valued functions defined on integer lattice points.

The notion of discrete unimodality is of interest in connection with statistical physics where a typical problem is to find the maximum of a unimodal probability function.

In Section 4.2 we present sufficient conditions for a trivariate discrete distribution to be strongly unimodal. We then give a sufficient condition that ensures the strong unimodality of a multivariate discrete distribution and prove the strong unimodality of the negative multinomial distribution, the multivariate hypergeometric distribution, the multivariate negative hypergeometric distribution and the Dirichlet (or Beta) compound multinomial distribution. These theoretical investigations lead to some practical suggestion on how to find the maximum of a strongly unimodal multivariate discrete distribution. We use the results of the paper by Prékopa and Li [75] and present a dual type algorithm to find the maximum of a strongly unimodal multivariate discrete distribution.

The three concepts that we are concerned with in this dissertation are connected. Probability bounds are used in probabilistic constrained stochastic programming problems. Prékopa [78] showed how the probability bounds can be incorporated in probabilistic constrained stochastic programming models in order to obtain approximate solutions for them. Discrete logconcavity also plays important role in stochastic programming. To see this we refer to the papers on programming under probabilistic constraints with discrete random variables: Prékopa, Vizvári and Badics [77], Dentcheva, Prékopa and Ruszczyński [23, 24, 25], Vizvári [95], Dentcheva, Lai, Ruszczyński [26], Beraldi and Ruszczyński [8]. Many others have been published during the past few years.

Chapter 2

Bounds for the Probability of the Union of Events Under Unimodality Condition

2.1 Introduction

The problem to find or approximate the probability of the union, intersection and other Boolean functions of random events has important applications in reliability theory, statistics, probability theory and stochastic programming. Typical examples are the reliability evaluations of multistate networks such as oil and gas supply systems, communication systems, power generation and transmission systems, etc. The reliability of the system is simply the probability of the union of all events that yield system success. Similarly, the unreliability is the probability of the union of all events that yield system failure. In [72] Prékopa and Boros presented sharp lower and upper bounds for the probability that a feasible flow exists in a stochastic transportation network. Another application is due to Prékopa, Boros and Lih [73] where they presented bounds for the communication network reliability. Prékopa, Long and Szántai [80] used probability bounds in PERT problem. Subasi et al. [88] presented an application in reliability theory. In another paper they presented bounds for the probability distribution of the length of the critical path and bounds for the European call option price. Another interesting application was presented by Boros and Prékopa [16]. They showed that even a deterministic problem (maximum satisfiability) can be solved by the use of probability bounds. For further applications we refer to Barlow and Proschan [4], N. H. Roberts et al. [85], and F. Roberts et al. [84].

Since the problem is of practical importance, intensive research efforts have been made in this field. Boole [12, 13, 14] was the first to suggest algebraic methods to find the bounds for the probability of the union of events. The classical Boole inequality asserts that the probability of the union of a finite number of events is smaller than or equal to the sum of the probabilities of the individual events. For approximation purposes it is unsatisfactory in most practical applications. Hailperin [35] gave the modern linear programming formulation for the Boolean probability bounding of the union of events and showed that Boole's method is equivalent to Fourier-Motzkin elimination.

Given arbitrary events $A_1, ..., A_n$ in an arbitrary probability space Ω the kth binomial moment of them is designated by S_k and is defined by the equation:

$$S_k = \sum_{1 \le i_1 < ... < i_k \le n} P(A_{i_1} ... A_{i_k}) \ , \quad k = 1, ..., n \ .$$

Let $S_0 = 1$. It is well known that (see, e.g., Prékopa [74])

$$S_k = E\left[\binom{\nu}{k}\right] \ , \quad k = 0, ..., n \ , \tag{2.1.1}$$

where ν is the number of those events that occur.

If we introduce the notation $p_k = P(\nu = k), \ k = 0, ..., n$, then we can write (2.1.1) in the following more detailed form:

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

To compute the probability of the union of the events the inclusion-exclusion formula is available:

$$P(A_1 \cup ... \cup A_n) = S_1 - S_2 + ... + (-1)^{n-1} S_n \ .$$

For the history of this formula see Takács [91]. If n is large, we may not be able to compute all the binomial moments, still, we may be able to compute a few of them. Given that, and further information about the probability of the random variable ν, we can give lower and upper bounds for the probability of the union. The bounds may serve for approximation of that probability provided that they are close to each other.

The well-known bounds in this respect are due to Bonferroni [11]. He created the following lower and upper bounds for the probability of the union based on the knowledge of first m binomial moments, $S_1, ..., S_m$ for some $m < n$:

$$P(A_1 \cup ... \cup A_n) \ge S_1 - S_2 + ... + (-1)^{m-1} S_m \qquad \text{if } m \text{ is even,}$$

$$P(A_1 \cup ... \cup A_n) \le S_1 - S_2 + ... + (-1)^{m-1} S_m \qquad \text{if } m \text{ is odd .}$$

However, in general, Bonferroni bounds are not the best possible bounds.

Dawson and Sankoff [22] proposed a sharp lower bound for the probability of the union of events using the first two binomial moments of the occurrences and a linear programming formulation. For the case of $m \leq 3$, Kwerel [47] used linear programming to obtain sharp bounds for the probability of the union. He fully described the dual feasible bases of the problem, also in the cases, where we are bounding the probabilities that at least r and exactly r events occur, reproduced known formulas this way and gave special dual type algorithms to solve the problems. Prékopa [67, 68, 69] discovered that the sharp Bonferroni bounds can be formulated as discrete binomial moment problems and presented linear programming problems, with input data $S_1, ..., S_m$, the optimum values of which provides us with the bounds. As special cases he obtained the sharp Bonferroni inequalities of Dawson and Sankoff and some others bounds. He presented dual type algorithms for the solution of the general bounding problems. Boros and Prékopa [15] exploited the linear programming methodology and derived a variety of closed form bounds for the probability of the union as well as the probabilities that at least r and exactly r out of n events occur.

A classical upper bound is due to Hunter [38] which is based on a special graph structure called spanning tree. Bukszár and Prékopa [17] presented bounds of degree three by the use of graph structure called cherry-tree. Vizvari [96] generalized Hunter's upper bound using graph structures. Other papers presenting bounds along this line are Bukszár [19], Bukszár and Szántai [18], Veneziani [94].

Prékopa and Gao [81] used aggregation-disaggregation in linear programs to obtain bounds for the probability of the union. For other closed form probability bounds see Galambos and Mucci [31], Galambos and Simonelli [32], Alajaji, Kuai and Takahara [1].

In this chapter we assume that $S_{k_1}, ..., S_{k_m}$ are known for some $1 \leq k_1 < ... < k_m$, $m < n$. We call the attention that the binomial moments are not necessarily consecutive. We do not assume the knowledge of the probability distribution $\{p_i\}$ but we assume that it is unimodal, i.e., there exists an integer M $(0 \leq M \leq n)$ such that $p_0 \leq ... \leq p_M$ and $p_M \geq ... \geq p_n$. The number M may be equal to 0 or n, or satisfy $0 < M < n$. If we assume unimodality without specifying where the mode is, we solve a problem for each possible mode location and then optimize over the solutions.

To obtain lower and upper bounds for the probability of the union of events we formulate the LP:

$$\min(\max)\sum_{i=1}^{n} p_i$$

subject to

$$\sum_{i=0}^{n}\binom{i}{k_j}p_i = S_{k_j}\ ,\quad j=0,...,m \tag{2.1.2}$$

$$p_0 \le ... \le p_M$$

$$p_M \ge ... \ge p_n$$

$$p_i \ge 0\ ,\quad i=0,...,n\ ,$$

where $k_0 = 0$. In problem (2.1.2) the $p_0, ..., p_n$ are unknown variables. If $m < n$, then there are infinitely many probability distributions satisfying the constraints of problem (2.1.2). One of them is the true distribution of ν. This implies that the optimum value of the min (max) problem (2.1.2) is a lower (upper) bound for the probability of the union. These bounds have the property that, given $S_{k_1}, ..., S_{k_m}$ and the knowledge of the unimodality of $\{p_i\}$, no better bounds can be given for $P(A_1 \cup ... \cup A_n)$. In view of this fact, we call them sharp bounds. The binomial moment problem, without the unimodality constraint, has extensively been studied in [67, 68, 69, 70].

The paper by E. Subasi, M. Subasi and A. Prékopa [88] is the first, where sharp bounds are presented for the probability of the union under unimodality constraint for the distribution of the random variable ν. In that paper problem (2.1.2) was used for the case of $m = 2$, $k_1 = 1$, $k_2 = 2$ and bounds are given by the use of formulas as well as by the dual algorithm of linear programming. In another paper by E. Subasi, M. Subasi and A. Prékopa, bounding formulas have been obtained for the probability that at least r and exactly r out of n events occur, under the same conditions.

Our purpose is to derive a general theorem in connection with problem (2.1.2) that characterizes the dual feasible bases of a relaxed version of the problem, further, to present closed form and algorithmic bounds for the probability of the union.

As it is known in linear programming theory, the objective function value corresponding to any dual feasible basis in the minimization (maximization) problem provides us with a lower (upper) bound for the optimum value of the problem.

First we reformulate problem (2.1.2) by introducing new variables $v_0, ..., v_n$. This can be done in two different ways:

$$p_0 = v_0 \ , \quad p_1 = v_0 + v_1 \ , \ \dots \ , \quad p_M = v_0 + \dots + v_M$$

$$p_{M+1} = v_{M+1} + \dots + v_n \ , \ p_{M+2} = v_{M+2} + \dots + v_n \ , \ \dots \ , \quad p_n = v_n \ , \tag{2.1.3}$$

and

$$p_0 = v_0 \ , \quad p_1 = v_0 + v_1 \ , \ \dots \ , \quad p_{M-1} = v_0 + \dots + v_{M-1}$$

$$p_M = v_M + \dots + v_n \ , \ p_{M+1} = v_{M+1} + \dots + v_n \ , \ \dots \ , \quad p_n = v_n \ . \tag{2.1.4}$$

The case $M = n$ is included in (2.1.3) and the case $M = 0$ included in (2.1.4).

If we use representation (2.1.3) in problem (2.1.2), we obtain the following problem:

$$min(max) \ \{Mv_0 + \sum_{i=1}^{M}(M - i + 1)v_i + \sum_{i=M+1}^{n}(i - M)v_i\}$$

subject to

$$\sum_{i=0}^{M}(M - i + 1)v_i + \sum_{i=M+1}^{n}(i - M)v_i = 1 \tag{2.1.5}$$

$$\sum_{i=0}^{M}\left[\binom{i}{k_j} + \dots + \binom{M}{k_j}\right]v_i + \sum_{i=M+1}^{n}\left[\binom{M+1}{k_j} + \dots + \binom{i}{k_j}\right]v_i = S_{k_j} \ , \qquad j = 1, ..., m$$

$$v_0 + \dots + v_M - v_{M+1} - \dots - v_n \geq 0 \tag{2.1.5a}$$

$$v_i \geq 0 \ , \quad i = 0, ..., n \ .$$

In case of representation (2.1.4) the problem can be formulated as follows:

$$min(max) \ \{(M - 1)v_0 + \sum_{i=1}^{M-1}(M - i)v_i + \sum_{i=M}^{n}(i - M + 1)v_i\}$$

subject to

$$\sum_{i=0}^{M-1}(M - i)v_i + \sum_{i=M}^{n}(i - M + 1)v_i = 1 \tag{2.1.6}$$

$$\sum_{i=0}^{M-1}\left[\binom{i}{k_j}+...+\binom{M-1}{k_j}\right]v_i+\sum_{i=M}^{n}\left[\binom{M}{k_j}+...+\binom{i}{k_j}\right]v_i=S_{k_j}\ ,$$
$$j=1,...,m$$

$$v_M+...+v_n-v_0-...-v_{M-1}\geq 0 \qquad (2.1.6a)$$

$$v_i\geq 0\ ,\quad i=0,...,n\ .$$

Problem (2.1.5) without constraint (2.1.5a) and problem (2.1.6) without (2.1.6a) will be called relaxed problems. For both relaxed problems $A=(a_0,...,a_n)$ will designate the matrix of the equality constraints, b the right hand side vector and c the vector of coefficients of the objective function.

The organization of this chapter is as follows. In Section 2.2 we characterize the dual feasible bases of the relaxed problem. In Section 2.3 bounding formulas are derived for the probability of the union, for the case of $m=2$ and general k_1, k_2 $(1\leq k_1<k_2\leq n)$. In Section 2.4 we present closed form bounds for the case of $m=3$ and $k_1=1$, $k_2=2$, $k_3=3$. In Section 2.5 upper bound formulas are derived for the probability of the union for the case of $m=4$ and $k_1=1$, $k_2=2$, $k_3=3$, $k_4=4$. In Section 2.6 general algorithms are presented to obtain algorithmic bounds. Finally, numerical examples are presented in Section 2.7.

2.2 Characterization of Dual Feasible Bases of The Relaxed Problem

In what follows we make use of a general theorem for the Pascal matrix, i.e., the matrix P consisting of binomial coefficients:

$$P=\begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 & 1 \\ & 1 & 2 & 3 & \cdots & n-1 & n \\ & & 1 & \binom{3}{2} & \cdots & \binom{n-1}{2} & \binom{n}{2} \\ & & & & \ddots & \vdots & \vdots \\ & & & & & 1 & \binom{n}{n-1} \\ & & & & & & 1 \end{pmatrix},$$

where there are zeros in the unfilled positions.

The term "minor" of a matrix will be used in the following sense: it is the determinant of a submatrix crossed out arbitrarily by the same number of rows and columns.

Theorem 1. [34, 67] *Any minor of P that has all positive entries in its main diagonal, is positive.*

In the next theorem we characterize the dual feasible bases of the relaxed version of problems (2.1.5), (2.1.6). For basic notions, facts and algorithms in connection with linear programming the reader is referred to the paper by Prékopa [76].

Theorem 2. *Any dual feasible basis of any of the relaxed problems* (2.1.5), (2.1.6) *has one of the following structures, presented in terms of the subscripts:*

	$m+1$ *even*	$m+1$ *odd*
min problem	$\{0, i, i+1, ..., j, j+1, n\}$	$\{0, i, i+1, ..., j, j+1\}$
max problem	$\{0, 1, i, i+1, ..., j, j+1\}$ *or* $I_B \subset \{1, ..., n\}$	$\{0, 1, i, i+1, ..., j, j+1, n\}$ *or* $I_B \subset \{1, ..., n\}$

Table 2.1: Dual feasible bases of problems (2.1.5), (2.1.6)

where I_B is the set of subscripts of the vectors that are in the basis B. In addition all dual feasible bases are dual nondegenerate, except for those with $I_B \subset \{1, ..., n\}$ which are dual degenerate.

Proof. We carry out the proof for the relaxed problem (2.1.5). The proof of the assertion for problem (2.1.6) is the same. For the sake of simplicity we prove the assertion for the case of $k_j = j, \ j = 1, ..., m$. The reasoning is, however applicable for the general case.

Let us write up in detailed form the matrix A of the equality constraints of problem (2.1.5), with the objective function coefficients on top of it:

	0	1	2	$\dots$	$M-1$	M	$M+1$	$M+2$	$\dots$	n
c^T	M	M	$M-1$	$\dots$	2	1	1	2	$\dots$	$n-M$
0	$M+1$	M	$M-1$	$\dots$	2	1	1	2	$\dots$	$n-M$
1	$0+\dots+M$	$1+\dots+M$	$2+\dots+M$	$\dots$	$M-1+M$	M	$M+1$	$M+1+M+2$	$\dots$	$M+1+\dots+n$
2	$\binom{0}{2}+\dots+\binom{M}{2}$	$\binom{1}{2}+\dots+\binom{M}{2}$	$\binom{2}{2}+\dots+\binom{M}{2}$	$\dots$	$\binom{M-1}{2}+\binom{M}{2}$	$\binom{M}{2}$	$\binom{M+1}{2}$	$\binom{M+1}{2}+\binom{M+2}{2}$	$\dots$	$\binom{M+1}{2}+\dots+\binom{n}{2}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\dots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\dots$	$\vdots$
m	$\binom{0}{m}+\dots+\binom{M}{m}$	$\binom{1}{m}+\dots+\binom{M}{m}$	$\binom{2}{m}+\dots+\binom{M}{m}$	$\dots$	$\binom{M-1}{m}+\binom{M}{m}$	$\binom{M}{m}$	$\binom{M+1}{m}$	$\binom{M+1}{m}+\binom{M+2}{m}$	$\dots$	$\binom{M+1}{m}+\dots+\binom{n}{m}$

Table 2.2: Matrix of equality constraints of problem (2.1.5), with the objective function coefficients on top

If $M = n$, then the columns below $M+1, ..., n$ do not exist. A basis B in the minimization problem (2.1.5) is dual feasible if the following inequalities hold:

$$c_B^T B^{-1} a_p \le c_p \quad \text{for any nonbasic } p \ .$$

For the maximization problem the dual feasibility of a basis is defined by the reversed inequalities. A basis B is dual degenerate if there is at least one nonbasic p such that $c_p - c_B^T B^{-1} a_p = 0$. Since we have

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

the first component of the solution of this equation can be expressed as

$$c_p - c_B^T B^{-1} a_p = \frac{1}{|B|} \begin{vmatrix} c_p & c_B^T \\ a_p & B \end{vmatrix} .$$

We are interested in the sign of $|B|$ and $\begin{vmatrix} c_p & c_B^T \\ a_p & B \end{vmatrix}$. In connection with them we prove the following.

Lemma. We have the inequality $|B| > 0$ and if $0 \in I_B$, then the determinant that comes out of $\begin{vmatrix} c_p & c_B^T \\ a_p & B \end{vmatrix}$, if we put $\begin{pmatrix} c_p \\ a_p \end{pmatrix}$ in its right place (the column subscripts are in increasing order), is also positive, where p is a nonbasic subscript.

Proof of the Lemma. Since B is a basis, it follows that $|B| \neq 0$. We prove that this value is positive.

The entries in the first row can be written up as sum of 1's so that the number of terms in any position in that row is equal to the number of terms in any entry in its column. Then we apply a column subtraction procedure, further, split the obtained determinant into a sum of determinants. Any determinant in the obtained sum is either zero, or positive, by Theorem 1, because they are minors, crossed out of the matrix P. At least one term must be positive because $|B| \neq 0$. It follows that $|B| > 0$.

Now we prove the second assertion. If $\begin{pmatrix} c_p \\ a_p \end{pmatrix}$ is put in its right place, then the first column of $\begin{pmatrix} c^T \\ A \end{pmatrix}$ will be the first column of the new determinant. If we subtract the first row from the second row in the determinant, then the first entry in the second row becomes -1 and the others 0. If we develop the determinant according to the second row, then, due to the special structure of the determinant, we obtain a minor of order $m+1$ crossed out of the matrix A. If $I_B = \{0, i_1, ..., i_m\}$, where $1 \leq i_1 < ... < i_m$, then the subscript set of the columns of the minors is $\{i_1, ..., p, ..., i_m\}$, where $i_1 < ... < p < ... < i_m$. Thus, 0 is removed from I_B and p is included. It is not difficult to see that the positivity of $|B|$ implies the positivity of the minor. We have proved the Lemma.

Returning to the proof of the Theorem 2, consider first the case $0 \notin I_B$. Then $|B| > 0$ and

$$\begin{vmatrix} c_p & c_B^T \\ a_p & B \end{vmatrix} = \begin{cases} 0 & \text{if } \; p \neq 0 \\ < 0 & \text{if } \; p = 0 \; . \end{cases}$$

Hence, B is a dual feasible basis in the maximization problem.

If, on the other hand, $0 \in I_B$, then still $|B| > 0$ and by the Lemma, the determinant $\begin{vmatrix} c_p & c_B^T \\ a_p & B \end{vmatrix}$ is equal to the $(m+1) \times (m+1)$ minor taken from A, corresponding to the columns $\{i_1, ..., p, ..., i_m\}$, multiplied by $(-1)^{h(p)}$, where $h(p)$ is the number of subscripts in B that are smaller than p. The minor is positive by the Lemma. We want to ensure the positivity of $\begin{vmatrix} c_p & c_B^T \\ a_p & B \end{vmatrix}$ for any nonbasic p. Now, if it is a minimization problem, then $h(p)$ must be even for any nonbasic p which implies that $\{i_1, ..., i_m\} = \{i, i+1, ..., j, j+1\}$ if m is even ($m+1$ is odd) and $\{i_1, ..., i_m\} = \{i, i+1, ..., j, j+1, n\}$ if m is odd ($m+1$ is even). If it is a maximization problem, then $h(p)$ must be odd for any nonbasic p which implies that $\{i_1, ..., i_m\} = \{1, i, i+1, ..., j, j+1, n\}$ if m is even ($m+1$ is odd) and $\{i_1, ..., i_m\} = \{1, i, i+1, ..., j, j+1\}$ if m is odd ($m+1$ is even). This proves the theorem. $\square$

Remark. If $k_j = j, \; j = 1, ..., m$, then all $(m+1) \times (m+1)$ submatrices of A are nonsingular. This is, however, not necessarily the case if $\{k_1, ..., k_m\} \neq \{1, ..., m\}$. Thus, when picking a dual feasible basis satisfying the structure in Theorem 2, we have to check on their independence as well.

2.3 Closed Form Bounds for The Probability of The Union Based on S_{k_1}, S_{k_2}

Let $m = 2$ and assume that the binomial moments S_{k_1}, S_{k_2}, $1 \leq k_1 < k_2 \leq n$, are known. If $C(n,k) = \binom{n}{k}$, then we have the following recurrence relation known as Pascal's rule:

$$C(n+1, k+1) = C(n,k) + C(n, k+1) \ .$$

By the use of these the coefficients of the equality constraints in the relaxed problems can be given as follows:

$$\sum_{s=i}^{j} \binom{s}{k} = \binom{j+1}{k+1} - \binom{i}{k+1} \ . \tag{2.3.1}$$

In view of (2.3.1) the relaxed version of problem (2.1.5) can be written in the form:

$$min(max) \ \{Mv_0 + \sum_{i=1}^{M}(M-i+1)v_i + \sum_{i=M+1}^{n}(i-M)v_i\}$$

subject to

$$\sum_{i=0}^{M}(M-i+1)v_i + \sum_{i=M+1}^{n}(i-M)v_i = 1 \tag{2.3.2}$$

$$\sum_{i=0}^{M}\left[\binom{M+1}{k_1+1} - \binom{i}{k_1+1}\right]v_i + \sum_{i=M+1}^{n}\left[\binom{i+1}{k_1+1} - \binom{M+1}{k_1+1}\right]v_i = S_{k_1}$$

$$\sum_{i=0}^{M}\left[\binom{M+1}{k_2+1} - \binom{i}{k_2+1}\right]v_i + \sum_{i=M+1}^{n}\left[\binom{i+1}{k_2+1} - \binom{M+1}{k_2+1}\right]v_i = S_{k_2}$$

$$v_i \geq 0 \ , \quad i = 0, ..., n \ .$$

Theorem 2 provides us with the following dual feasible bases for the above problem:

$$B_{min} = \{0, i, i+1\} \ , \quad 1 \leq i \leq n-1 \ ,$$

$$B_{max} = \{0, 1, n\} \quad \text{or} \quad B_{max} \subset \{1, ..., n\} \ .$$

In order to present our formulas in compact forms we introduce the notations:

$$\Sigma^r_{i,j} = (j-i+1)\binom{i-1}{r} - \binom{j+1}{r+1} + \binom{i}{r+1}$$

$$\Sigma^{r,t}_{i,j} = \binom{i-1}{r}\left[\binom{j+1}{t+1} - \binom{i}{t+1}\right] - \binom{i-1}{t}\left[\binom{j+1}{r+1} - \binom{i}{r+1}\right]$$

$$\gamma^r_{i,j} = i\left[\binom{j+1}{r+1} - \binom{i}{r+1}\right] - (j-i+1)\binom{i}{r+1}$$

$$\gamma^{r,t}_{i,j} = \binom{i}{r+1}\left[\binom{j+1}{t+1} - \binom{i}{t+1}\right] - \binom{i}{t+1}\left[\binom{j+1}{r+1} - \binom{i}{r+1}\right]$$

$$\beta^r_{i,j} = (j+1)\binom{j+1}{r} - \binom{j+1}{r+1} + \binom{i}{r+1}$$

$$\beta^{r,t}_{i,j} = \binom{j+1}{r}\left[\binom{j+1}{t+1} - \binom{i}{t+1}\right] - \binom{j+1}{t}\left[\binom{j+1}{r+1} - \binom{i}{r+1}\right]$$

$$\alpha^r_{i,j} = (j-i+1)\binom{j+1}{r} - \binom{j+1}{r+1} + \binom{i}{r+1}$$

$$\alpha^{r,t}_{i,j} = \binom{j+1}{r}\left[\binom{j+1}{t+1} - \binom{i}{t+1}\right] - \binom{j+1}{t}\left[\binom{j+1}{r+1} - \binom{i}{r+1}\right]$$

$$\delta^r_{i,j} = (i-1)\left[\binom{j+1}{r+1} - \binom{i}{r+1}\right] - (j-i)\binom{i}{r+1}$$

Table 2.3: Notations

We use problem (2.3.2) to present lower and upper bounds for $P(\nu \geq 1)$. To do this we find the optimal bases for the minimization and maximization problems, respectively.

We already have a full description of the dual feasible bases. What we need is to find those (one for the min problem and one for the max problem) that are also primal feasible. Three cases will be considered.

Closed Form Lower Bounds

Case 1. Let $1 \leq i \leq M-1$. The primal feasibility conditions for B_{min} are given below:

$$S_{k_1}\Sigma_{i+1,M}^{k_2} - S_{k_2}\Sigma_{i+1,M}^{k_1} + \Sigma_{i+1,M}^{k_1,k_2} \geq 0 \ ,$$

$$S_{k_1}\gamma_{i+1,M}^{k_2} - S_{k_2}\gamma_{i+1,M}^{k_1} - \gamma_{i+1,M}^{k_1,k_2} \geq 0 \ ,$$

$$S_{k_1}\gamma_{i,M}^{k_2} - S_{k_2}\gamma_{i,M}^{k_1} + \gamma_{i,M}^{k_1,k_2} \leq 0 \ .$$

In this case the closed form lower bound for $P(\nu \geq 1)$ is expressed by

$$1 - \frac{S_{k_1}\Sigma_{i+1,M}^{k_2} - S_{k_2}\Sigma_{i+1,M}^{k_1} + \Sigma_{i+1,k}^{k_1,k_2}}{i\ \Sigma_{i+1,M}^{k_1,k_2} + \binom{i}{k_1+1}\Sigma_{i+1,M}^{k_2} - \binom{i}{k_2+1}\Sigma_{i+1,M}^{k_1}} \leq P(\nu \geq 1) \ , \qquad (2.3.3)$$

where $\Sigma_{i,j}^{r}, \Sigma_{i,j}^{r,t}, \gamma_{i,j}^{r}, \gamma_{i,j}^{r,t}$ are given in Table 2.3.

Case 2. Let $i = M$. The conditions that ensure the primal feasibility of $B_{min} = \{0, M, M+1\}$ are as follows:

$$S_{k_1}\binom{M}{k_2-1} - S_{k_2}\binom{M}{k_1-1} - \frac{k_2-k_1}{M-k_2+1}\binom{M+1}{k_1}\binom{M}{k_2} \geq 0 \ ,$$

$$S_{k_1}\beta_{1,M}^{k_2} - S_{k_2}\beta_{1,M}^{k_1} + \beta_{1,M}^{k_1,k_2} \geq 0 \ ,$$

$$S_{k_1}\beta_{1,M-1}^{k_2} - S_{k_2}\beta_{1,M-1}^{k_1} + \beta_{1,M-1}^{k_1,k_2} \geq 0 \ .$$

The corresponding closed form lower bound for $P(\nu \geq 1)$ is given by

$$1 \;-\; \frac{S_{k_1}\binom{M}{k_2-1} - S_{k_2}\binom{M}{k_1-1} \;-\; \frac{k_2-k_1}{M-k_2+1}\binom{M+1}{k_1}\binom{M}{k_2}}{\beta^{k_1,k_2}_{1,M-1} \;-\; \frac{M\;(k_2-k_1)}{M-k_2+1}\binom{M+1}{k_1}\binom{M}{k_2}} \;\le\; P(\nu \ge 1)\;, \tag{2.3.4}$$

where $\beta^{r}_{i,j}, \beta^{r,t}_{i,j}$ are given in Table 2.3.

Case 3. Let $M+1 \le i \le n-1$. B_{min} is primal feasible if and only if i is determined by the following conditions:

$$S_{k_1}\alpha^{k_2}_{M+1,i} - S_{k_2}\alpha^{k_1}_{M+1,i} + \alpha^{k_1,k_2}_{M+1,i} \ge 0\;,$$

$$S_{k_1}\gamma^{k_2}_{M+1,i+1} - S_{k_2}\gamma^{k_1}_{M+1,i+1} - \gamma^{k_1,k_2}_{M+1,i+1} \ge 0\;,$$

$$S_{k_1}\gamma^{k_2}_{M+1,i} - S_{k_2}\gamma^{k_1}_{M+1,i} - \gamma^{k_1,k_2}_{M+1,i} \ge 0\;.$$

Then the closed form lower bound is the following:

$$1 \;-\; \frac{S_{k_1}\alpha^{k_2}_{M+1,i} - S_{k_2}\alpha^{k_1}_{M+1,i} + \alpha^{k_1,k_2}_{M+1,i}}{\binom{M+1}{k_1+1}\,\alpha^{k_2}_{M+1,i} - \binom{M+1}{k_2+1}\alpha^{k_1}_{M+1,i} + \;(M+1)\;\alpha^{k_1,k_2}_{M+1,i}} \;\le\; P(\nu \ge 1)\;, \tag{2.3.5}$$

where $\alpha^{r}_{i,j}, \alpha^{r,t}_{i,j}, \gamma^{r}_{i,j}, \gamma^{r,t}_{i,j}$ are given in Table 2.3.

Closed Form Upper Bounds

If $B_{max} \subset \{1,...,n\}$ is primal feasible in the relaxed version of the maximization problem (2.3.2), then the upper bound for the probability of the union is equal to 1.

The basis $B_{max} = \{0,1,n\}$ is primal feasible if and only if the following conditions hold:

$$S_{k_1}\delta^{k_2}_{M+1,n} - S_{k_2}\delta^{k_1}_{M+1,n} - \gamma^{k_1,k_2}_{M+1,n} \le 0\;,$$

$$S_{k_1}\gamma^{k_2}_{M+1,n} - S_{k_2}\gamma^{k_1}_{M+1,n} - \gamma^{k_1,k_2}_{M+1,n} \geq 0 \ ,$$
$$S_{k_1}\begin{pmatrix} M+1 \\ k_2+1 \end{pmatrix} \ \leq \ S_{k_2}\begin{pmatrix} M+1 \\ k_1+1 \end{pmatrix} .$$

The corresponding closed form upper bound for $P(\nu \geq 1)$ is given below:

$$P(\nu \geq 1) \ \leq \ \frac{S_{k_1}\delta^{k_2}_{M+1,n} - S_{k_2}\delta^{k_1}_{M+1,n}}{\gamma^{k_1,k_2}_{M+1,n}} \ , \tag{2.3.6}$$

where $\delta^r_{i,j}, \gamma^r_{i,j}, \gamma^{r,t}_{i,j}$ are given in Table 2.3.

If we use the relaxed version of problem (2.1.6), rather than that of problem (2.1.5), then the lower and upper bounds change in such a way that we have to replace $M-1$ for M in the formulas of Section 2.3.

2.4 Closed Form Bounds for The Probability of The Union Based on S_1, S_2, S_3

We look at the relaxed versions of problems (2.1.5), (2.1.6) and create bounds for the probability of the union, based on the knowledge of the binomial moments S_1, S_2, S_3. Since $m+1$ is even, then by the use of Theorem 2, we derive that any dual feasible basis B_{min} of the relaxed version of the minimization problem (2.1.5) has the form:

$$B_{min} = \{0, i, i+1, n\} \ , \quad i = 1, ..., n-2 \ .$$

Similarly, any dual feasible basis B_{max} of relaxed version of the maximization problem has the form:

$$B_{max} = \{0, 1, i, i+1\}, \ i = 2, ..., n-1, \quad \text{or} \quad B_{max} \subset \{1, ..., n\} \ .$$

Below we present conditions that ensure the primal feasibility of B_{min} as well as the corresponding lower bounds for $P(\nu \geq 1)$, i.e., the probability of the union of the events.

<u>Closed Form Lower Bounds</u>

Case 1. Let $1 \le i \le M-1$. B_{min} is primal feasible if and only if i is determined by the conditions

$$2[iM+(n-1)(i+M-1)]S_1 - 6(n+i+M-3)S_2 + 24S_3 \ge Mni \ ,$$

$$2[M(i-1)+(n-1)(i+M-2)]S_1 - 6(n+i+M-4)S_2 + 24S_3 \le Mn(i-1) \ ,$$

$$2(i-1)(i+2M-2)S_1 - 6(2i+M-4)S_2 + 24S_3 \ge Mi(i-1) \ ,$$

$$2[i(n+2M+i)+(n-1)(i+M-1)]S_1 - 6(n+2i+M-3)S_2 + 24S_3$$

$$\le \ i[M(2n+i+1)+(i+1)(n+1)] \ .$$

In this case the lower bound for $P(\nu \ge 1)$ is obtained as follows:

$$\frac{2[i(n+2M+i)+(n-1)(i+M-1)]S_1 - 6(n+2i+M-3)S_2 + 24S_3}{(n+1)(M+1)(i+1)i} + \frac{Mn(i-1)}{(n+1)(M+1)(i+1)} \ \le \ P(\nu \ge 1) \ . \tag{2.4.1}$$

Case 2. Let $i = M$. Basis $B_{min} = \{0, M, M+1, n\}$ is primal feasible if and only if the following conditions are satisfied:

$$2M(2n+M-1)S_1 - 6(n+2M-2)S_2 + 24S_3 \ge M(M+1)n \ ,$$

$$2(M-1)(2n+M-2)S_1 - 6(n+2M-4)S_2 + 24S_3 \le (M-1)Mn \ ,$$

$$6M(M-1)S_1 - 18S_2 + 24S_3 \ge (M-1)M(M+1) \ ,$$

$$6M(n+M)S_1 - 6(n+3M-2)S_2 + 24S_3 \le (M+1)(3n+M+2) \ .$$

The corresponding lower bound for $P(\nu \ge 1)$ is given below:

$$\frac{6M(n+M)S_1 - 6(n+3M-2)S_2 + 24S_3}{M(M+1)(M+2)(n+1)} + \frac{n(M-1)}{(M+2)(n+1)} \ \le \ P(\nu \ge 1) \ . \tag{2.4.2}$$

Case 3. Let $M+1 \le i \le n-2$. B_{min} is primal feasible if and only if i satisfies the following conditions:

$$2[nM + i(n+M-1)]S_1 - 6(n+i+M-2)S_2 + 24S_3 \ge Mn(i+1) \ ,$$

$$2[iM + (n-1)(i+M-1)]S_1 - 6(n+i+M-3)S_2 + 24S_3 \le Mni \ ,$$

$$2i(i+2M-1)S_1 - 6(2i+M-2)S_2 + 24S_3 \ge i(M+1)M \ ,$$

$$2[i(n+2M+i) + (n+1)(i+M+1)]S_1 - 6(n+2i+M-1)S_2 + 24S_3$$
$$\le \ (i+1)[iM + (n+1)(i+2M+2)] \ .$$

In this case the lower bound is obtained as follows:

$$\frac{2[i(n+2M+i) + (n+1)(i+M+1)]S_1 - 6(n+2i+M-1)S_2 + 24S_3}{(i+1)(i+2)(M+1)(n+1)}$$
$$+\frac{niM}{(i+2)(M+1)(n+1)} \ \le \ P(\nu \ge 1) \ . \qquad (2.4.3)$$

<u>Closed Form Upper Bounds</u>

In order to obtain an upper bound for $P(\nu \ge 1)$ we consider the relaxed version of the maximization problem (2.1.5). Note that if the dual feasible basis $B_{max} \subset \{1, ..., n\}$ is also primal feasible, then the optimum value of the maximization problem, i.e., the upper bound for the probability of the union, is equal to 1. As before, we have three cases for the choice of i.

Case 1. Let $2 \le i \le M-1$. The primal feasibility conditions for the basis $B_{max} = \{0, 1, i, i+1\}$ are as follows:

$$2(i-1)(i+2M-2)S_1 - 6(2i+M-4)S_2 + 24S_3 \ge M(i-1)i \ ,$$

$$2(i-1)(M-1)S_1 - 6(i+M-3)S_2 + 24S_3 \le 0 \ ,$$

$$2(i-2)(M-1)S_1 - 6(i+M-4)S_2 + 24S_3 \ge 0 \ ,$$

$$2[i(i+M) + (i-1)(M-1)]S_1 - 6(2i+M-3)S_2 + 24S_3 \le i(i+1)(M+1) \ .$$

The corresponding upper bound for $P(\nu \ge 1)$ is presented below:

$$P(\nu \geq 1) \ \leq \ \frac{2[i(i+M)+(i-1)(M-1)]S_1 - 6(2i+M-3)S_2 + 24S_3}{i(i+1)(M+1)} \ . \qquad (2.4.4)$$

Case 2. Let $i = M$. The basis $B_{max} = \{0, 1, M, M+1\}$ is primal feasible if and only if

$$6M(M-1)S_1 - 18(M-1)S_2 + 24S_3 \geq (M-1)M(M+1) \ ,$$

$$2M(M-1)S_1 - 12(M-1)S_2 + 24S_3 \leq 0 \ ,$$

$$2(M-1)(M-2)S_1 - 12(M-2)S_2 + 24S_3 \geq 0 \ ,$$

$$6M^2S_1 - 6(3M-2)S_2 + 24S_3 \leq M(M-1)(M+1) \ .$$

The corresponding upper bound for $P(\nu \geq 1)$ is

$$P(\nu \geq 1) \ \leq \ \frac{6M^2S_1 - 6(3M-2)S_2 + 24S_3}{M(M+1)(M+2)} \ . \qquad (2.4.5)$$

Case 3. Let $M + 1 \leq i \leq n - 1$. The basis B_{max} is primal feasible if and only if i is determined by the following conditions:

$$2i(i+2M-1)S_1 - 6(2i+M-2)S_2 + 24S_3 \geq i(i+1)M \ ,$$

$$2i(M-1)S_1 - 6(i+M-2)S_2 + 24S_3 \leq 0 \ ,$$

$$2(i-1)(M-1)S_1 - 6(i+M-3)S_2 + 24S_3 \geq 0 \ ,$$

$$2[i(i+M)+(i+1)(M+1)]S_1 - 6(2i+M-1)S_2 + 24S_3 \leq (i+1)(i+2)(M+1) \ .$$

With i satisfying these inequalities we have the upper bound given by:

$$P(\nu \geq 1) \ \leq \ \frac{2[i(i+M)+(i+1)(M+1)]S_1 - 6(2i+M-1)S_2 + 24S_3}{(i+1)(i+2)(M+1)} \ . \qquad (2.4.6)$$

If we replace $M - 1$ for M in the formulas of Section 2.4, then we obtain the closed form bounds that come out of the relaxed version of problem (2.1.6).

2.5 Closed Form Upper Bounds for The Probability of The Union Based on S_1, S_2, S_3, S_4

In this section we present upper bound formulas for the probability of the union of events based on the first four binomial moments.

Since $m+1$ is odd, then by Theorem 2, any dual feasible basis B_{max} of the relaxed version of the maximization problem (2.1.5) or (2.1.6) is of the form:

$$B_{max} = \{0, 1, i, i+1, n\},\ i = 2, ..., n-1, \quad \text{or} \quad B_{max} \subset \{1, ..., n\}\ .$$

If $B_{max} \subset \{1, ..., n\}$, then the upper bound for $P(\nu \geq 1)$ is 1. In order to determine the index i that ensures the primal feasibility of the basis of the form $B_{max} = \{0, 1, i, i+1, n\}$ we consider the following cases.

Case 1. Let $2 \leq i \leq M-1$. The primal feasibility conditions are:

$$2[i(i-1)(n+M) - (M-1)(n-1) + 2i(nM+1)]S_1$$

$$-6(3n+3M-nM-i^2+5i-2iM-2ni-7)S_2 + 24(n+2i+M-6)S_3 - 120S_4$$

$$\leq\ i(i+1)(n+1)(M+1)\ ,$$

$$2(i-1)(ni+iM+2nM-2n-i-2M+2)S_1 - 6[nM+(i-2)(2n+2M+i-5)]S_2$$

$$+24(n+2i+M-7)S_3 - 120S_4\ \geq\ (i-1)inM\ ,$$

$$2(M-1)(n-1)(i-1)S_1 - 6(ni+iM+nM-3n-3i-3M+7)S_2$$

$$+24(n+i+M-6)S_3 - 120S_4\ \leq\ 0\ ,$$

$$2(M-1)(n-1)(i-2)S_1 - 6(ni+iM+nM-4n-3i-4M+10)S_2$$

$$+24(n+i+M-7)S_3 - 120S_4\ \geq\ 0\ ,$$

$$2(M-1)(i-1)(i-2)S_1 - 6(i-2)(i+2M-5)S_2$$

$$+24(2i+M-7)S_3 - 120S_4\ \leq\ 0\ .$$

Under these conditions the corresponding upper bound is given below:

$$
\begin{aligned}
P(\nu \geq 1) \ \leq \ & \frac{2[i(i-1)(n+M)-(M-1)(n-1)+2i(nM+1)]S_1}{(M+1)i(i+1)(n+1)} \\
& - \frac{6(3n+3M-nM-i^2+5i-2iM-2ni-7)S_2}{(M+1)i(i+1)(n+1)} \\
& + \frac{24(n+2i+M-6)S_3-120S_4}{(M+1)i(i+1)(n+1)} \ .
\end{aligned}
\tag{2.5.1}
$$

Case 2. Let $i = M$. The basis $B_{max} = \{0, 1, M, M+1, n\}$ is primal feasible if and only if

$$
2M(3nM+M^2+2)S_1 - 6(3M^2+(3M-2)(n-2))S_2 + 24(n+3M-5)S_3 - 120S_4
$$

$$
\leq \ M(M+1)(M+2)(n+1) \ ,
$$

$$
2M(M-1)(3n+M-2)S_1 - 18(M-1)(n+M-2)S_2 + 24(n+3M-6)S_3 - 120S_4
$$

$$
\geq \ M(M-1)(M+1)n \ ,
$$

$$
2M(M-1)(n-1)S_1 - 6(M-1)(2n+M-4)S_2 + 24(n+2M-5)S_3 - 120S_4 \ \leq \ 0 \ ,
$$

$$
2(M-1)(M-2)(n-1)S_1 - 6(M-2)(2n+M-5)S_2 + 24(n+2M-7)S_3 - 120S_4 \ \geq \ 0 \ ,
$$

$$
2M(M-1)(M-2)S_1 - 18(M-1)(M-2)S_2 + 72(M-2)S_3 - 120S_4 \ \leq \ 0 \ .
$$

The closed form upper bound for $P(\nu \geq 1)$ is given by

$$
\begin{aligned}
P(\nu \geq 1) \ \leq \ & \frac{2M(3nM+M^2+2)S_1 - 6(3M^2+(3M-2)(n-2))S_2}{M(M+1)(M+2)(n+1)} \\
& + \frac{24(n+3M-5)S_3-120S_4}{M(M+1)(M+2)(n+1)} \ .
\end{aligned}
\tag{2.5.2}
$$

Case 3. Let $M+1 \leq i \leq n-2$. B_{max} is primal feasible if and only if i satisfies the conditions:

$$
2[(i+1)(ni+nM+iM+1)+n+i+M]S_1 - 6[(i-1)(i-2)+2i(n+M)+(n-1)(M-1)]S_2
$$

$$
+24(n+2i+M-4)S_3 - 120S_4 \ \leq \ (i+1)(i+2)(M+1)(n+1) \ ,
$$

$$2i[n(i+M)+(M-1)(n+i-1)]S_1 - 6[(i-1)(2n+2M+i-4)+nM]S_2$$

$$+24(n+2i+M-5)S_3 - 120S_4 \ \geq \ ni(i+1)M \ ,$$

$$2i(M-1)(n-1)S_1 - 6[(n-2)(i+M-2)+i(M-1)]S_2$$

$$+24(n+i+M-5)S_3 - 120S_4 \ \leq \ 0 \ ,$$

$$2(i-1)(M-1)(n-1)S_1 - 6[(n-3)(i+M-1)+iM-2n+4]S_2$$

$$+24(n_i+M-6)S_3 - 120S_4 \ \geq \ 0 \ ,$$

$$2i(i-1)(M-1)S_1 - 6(i-1)(i+2M-4)S_2 + 24(2i+M-5)S_3 - 120S_4 \ \leq \ 0 \ .$$

The corresponding upper bound for $P(\nu \geq 1)$ is given by

$$\begin{aligned} P(\nu \geq 1) \ \leq \ & \frac{2[(i+1)(ni+nM+iM+1)+n+i+M]S_1}{(i+1)(i+2)(n+1)(M+1)} \\ & - \frac{6[(i-1)(i-2)+2i(n+M)+(n-1)(M-1)]S_2}{(i+1)(i+2)(n+1)(M+1)} \\ & + \frac{24(n+2i+M-4)S_3 - 120S_4}{(i+1)(i+2)(n+1)(M+1)} \ . \end{aligned} \tag{2.5.3}$$

As before, if we apply our bounding technique on the relaxed problem (2.1.6), rather than (2.1.5), then the just derived formulas provide us with the upper bounds if we replace $M-1$ for M.

2.6 Algorithmic Bounds

In Sections 2.3, 2.4 and 2.5 we have derived closed form bounds for the probability of the union, by the use of the relaxed problems (2.1.5), (2.1.6) for the cases of $m = 2, 3, 4$. For larger m values the solution of the relaxed problems can be obtained by specially designed dual algorithms of linear programming. Once an algorithm of this kind terminates, the solutions for the non-relaxed problem can be continued again by the dual algorithm. In fact, as it is well known in linear programming, the dual algorithm can efficiently be used, as a reoptimization technique, whenever the optimal basis has already been found but a further constraint is introduced into the problem.

The algorithm presented below works in this way and is applicable to cases with consecutive and non-consecutive moments. We remark that it is more practical to carry out the algorithms to obtain the bound, rather than to apply a complicated closed form formula.

***Algorithmic solutions of problems* (2.1.5), (2.1.6)**

Step 0. Find an initial dual feasible basis B to the relaxed problem. Any basis that has the structure presented in Theorem 2 is suitable.

Step 1. Check for primal feasibility. If $B^{-1}b \geq 0$, then the solution of the relaxed problem terminates. Go to Step 4. Otherwise go to Step 2.

Step 2. If $(B^{-1}b)_j < 0$, then the jth vector in B (not necessarily equal to a_j) is a candidate to leave the basis. Choose arbitrarily among the candidates to leave the basis. Go to Step 3.

Step 3. Include the vector a_l into the basis that restores the dual feasible basis structure. Go to Step 1.

Step 4. If the additional constraint $v_0 + ... + v_M \geq v_{M+1} + ... + v_n$ (or $v_M + ... + v_n \geq v_0 + ... + v_{M-1}$) is satisfied, then the solution of problem (2.1.5) (or (2.1.6)) terminates. Otherwise go to Step 5.

Step 5. Reoptimize the problem with the additional constraint (2.1.5a) or (2.1.6a): introduce slack variable into the additional inequality constraint, prescribe nonnegativity relation for the slack variable, set up the new dual tableau and carry out the dual method.

If the sequence of probabilities $p_0, ..., p_n$ is increasing or decreasing, i.e., if $M = n$ or $M = 0$, then the solution of problem (2.1.5) or (2.1.6) terminates with Step 3. No reoptimization is needed. The relaxed problem is equivalent to the original problem (2.1.5) or (2.1.6).

Suppose that the additional constraint (2.1.5a) is added to the problem. If the original solution to the original problem (2.1.5) satisfies the added constraint, it is then obvious that the point is also an optimal solution of the new problem. If, on the other hand, the point does not satisfy constraint (2.1.5a), we can use the dual simplex method to find the new optimal solution.

Let B be the optimal basis before the constraint

$$v_0 + ... + v_M \geq v_{M+1} + ... + v_n \ ,$$

or equivalently

$$v_0 + ... + v_M - v_{M+1} - ... - v_n - v_{n+1} = 0 \ , \tag{2.6.1}$$

where v_{n+1} is slack variable, is added. The corresponding canonical system is as follows:

$$z + (c_B B^{-1} N - c_N) v_N = c_B B^{-1} b \tag{2.6.2}$$

$$v_B + B^{-1} N v_N = B^{-1} b \ . \tag{2.6.3}$$

Let $\mathbf{a}$ be the coefficient vector of the new constraint (2.6.1), i.e., $\mathbf{a} = (1, ..., 1, -1, ..., -1, -1)$ is an $(n+1)$-vector. Then $\mathbf{a}$ can be decomposed into $(\mathbf{a}_B, \mathbf{a}_N)$. Therefore equation (2.6.1) can be written as

$$\mathbf{a}_B v_B + \mathbf{a}_N v_N + v_{n+1} = 0 \ .$$

Multiplying equation (2.6.3) by $\mathbf{a}_B$ and subtracting from the new constraint gives the following system:

$$z + (c_B B^{-1} N - c_N) v_N = c_B B^{-1} b$$

$$v_B + B^{-1} N v_N = B^{-1} b$$

$$(\mathbf{a}_N - \mathbf{a}_B B^{-1} N) v_N + v_{n+1} = -\mathbf{a}_B B^{-1} b \ .$$

These equations give us a basic solution of the new system. The only possible violation of optimality of the new problem is the sign of $-\mathbf{a}_B B^{-1} b$. So if, in case of minimization problem, $-\mathbf{a}_B B^{-1} b \geq 0$, then the current solution is optimal. Otherwise, if $-\mathbf{a}_B B^{-1} b < 0$, then the dual simplex method is used to restore feasibility. Similar modifications can be done for the case of maximization problem. Note that if constraint (2.1.6a) is added to the problem, then $\mathbf{a} = (-1, ..., -1, 1, ..., 1, -1)$ and the rest of the calculations remains the same. For more information about reoptimization using dual simplex method the reader is referred to, e.g., [76].

2.7 Numerical Examples

We present numerical examples to show that if the probability distribution is unimodal with known mode, M, then by the use of our bounding methodology, we can obtain tighter bounds for the probability of the union, $P(\nu \geq 1)$. In the following examples LB and UB stand for lower and upper bounds, respectively.

Example 1. We give an illustration of the algorithm that we have presented in Section 2.6. Assume that the probability distribution is unimodal and its mode is 5. Let $n = 10$, $S_1 = 5.3568245$, $S_2 = 16.2332237$, $S_3 = 32.377332$.

We consider the relaxed version of the minimization problem (2.1.6) and choose the initial basis $B = \{0, 2, 3, 10\}$, which is dual feasible by Theorem 2.

Iteration 1

Step 0. Initial dual feasible basis: $B = \{0, 2, 3, 10\}$.

Step 1. Since

$$B^{-1}b = \begin{pmatrix} 0.070527273 \\ -0.024297844 \\ 0.06646205 \\ 0.097888845 \end{pmatrix} \ngeq 0 \ ,$$

it follows that B is not primal feasible.

Step 2. The second vector in B, that is a_2, leaves the basis since $(B^{-1}b)_2 \ < \ 0$.

Step 3. The vector a_4 restores the dual feasible basis structure, hence it enters the basis. We proceed to the second iteration with the updated basis, $B = \{0, 4, 3, 10\}$.

Iteration 2

Step 1. We have

$$B^{-1}b = \begin{pmatrix} 0.068539267 \\ 0.046860129 \\ 0.0117919 \\ 0.097809956 \end{pmatrix} \ > \ 0 \ .$$

Thus B is optimal and the optimum value of the relaxed problem (2.1.6) is 0.931460733. The solution of the relaxed problem terminates.

Step 4. The additional constraint (2.1.6a) is equivalent to

$$v_5 + ... + v_{10} - v_0 - ... - v_4 = -0.02938134 \ < \ 0 \ .$$

The optimal solution to the relaxed problem does not satisfy constraint (2.1.6a).

Step 5. In order to ensure the mode of the distribution is 5 we prescribe (2.1.6a) as an additional constraint:

$$v_5 + ... + v_{10} - v_0 - ... - v_4 \ \geq \ 0 \ .$$

Let us rewrite the constraint in the form

$$v_5 + ... + v_{10} - v_0 - ... - v_4 - v_{11} \ = \ 0 \ ,$$

where $v_{11} \geq 0$ is slack variable. We use the dual method to reoptimize the problem (see, e.g., [76]) After applying the dual method to the new problem, we obtain the optimal basis and the optimum value of problem (2.1.6), i.e., the lower bound for the probability of the union as given below:

$$\begin{pmatrix} v_0 \\ v_3 \\ v_4 \\ v_5 \\ v_{10} \end{pmatrix} = \begin{pmatrix} 0.0685393 \\ 0.0117919 \\ 0.0468601 \\ 0.0097938 \\ 0.09780996 \end{pmatrix} \qquad \text{and} \qquad 0.931905905 \ \leq \ P(\nu \geq 1) \ .$$

Example 2. We assume that the first m binomial moments of the events are known. In Table 2.4 we present bounds for $P(\nu \geq 1)$ with and without unimodality condition.

The bounds for $P(\nu \geq 1)$, obtained by the use of the relaxed problems (2.1.5) and (2.1.6), are presented in Table 2.5.

		with unimodality						without unimodality					
		$m = 2$		$m = 3$		$m = 4$		$m = 2$		$m = 3$		$m = 4$	
n	M	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB
10	3	0.9172	0.9404	0.9205	0.9326	0.9227	0.9281	0.7317	1	0.8190	1	0.8662	1
10	5	0.9255	0.9562	0.9319	0.9470	0.9334	0.9401	0.7594	1	0.8423	1	0.8815	1
10	7	0.9310	0.9735	0.9389	0.9603	0.9403	0.9475	0.7832	1	0.8580	1	0.8923	1
10	3	0.9093	0.9690	0.9151	0.9380	0.9173	0.9280	0.7225	1	0.8129	1	0.8618	1
10	5	0.9248	0.9753	0.9312	0.9544	0.9341	0.9447	0.7617	1	0.8442	1	0.8843	1
10	7	0.9339	0.9884	0.9424	0.9687	0.9446	0.9540	0.7935	1	0.8662	1	0.8998	1
10	3	0.9096	0.9446	0.9128	0.9259	0.9140	0.9201	0.7197	1	0.8078	1	0.8559	1
10	5	0.9187	0.9483	0.9225	0.9361	0.9244	0.9308	0.7428	1	0.8267	1	0.8697	1
10	7	0.9241	0.9570	0.9296	0.9459	0.9311	0.9371	0.7622	1	0.8405	1	0.8787	1
10	3	0.9114	0.9516	0.9156	0.9322	0.9174	0.9255	0.7236	1	0.8126	1	0.8609	1
10	5	0.9230	0.9622	0.9286	0.9470	0.9309	0.9394	0.7544	1	0.8381	1	0.8790	1
10	7	0.9300	0.9755	0.9375	0.9600	0.9394	0.9473	0.7812	1	0.8563	1	0.8911	1
10	3	0.9046	0.9364	0.9074	0.9185	0.9084	0.9132	0.7099	1	0.7990	1	0.8480	1
10	5	0.9132	0.9367	0.9161	0.9276	0.9176	0.9227	0.7301	1	0.8144	1	0.8604	1
10	7	0.9185	0.9447	0.9227	0.9352	0.9185	0.9286	0.7456	1	0.8270	1	0.7456	1

Table 2.4: Bounds for the probability of the union of events with and without unimodality

		Relaxed Problem (2.1.5)						Relaxed Problem (2.1.6)					
		$m = 2$		$m = 3$		$m = 4$		$m = 2$		$m = 3$		$m = 4$	
n	M	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB	LB	UB
10	3	0.9156	0.9404	0.9205	0.9326	0.9227	0.9288	0.9172	0.9460	0.9205	0.9333	0.9227	0.9281
10	5	0.9238	0.9592	0.9319	0.9500	0.9334	0.9401	0.9255	0.9562	0.9315	0.9470	0.9334	0.9401
10	7	0.9299	0.9805	0.9386	0.9641	0.9401	0.9475	0.9310	0.9735	0.9389	0.9603	0.9403	0.9477
10	3	0.9086	0.9690	0.9139	0.9380	0.9173	0.9280	0.9093	0.9986	0.9151	0.9455	0.9173	0.9280
10	5	0.9230	0.9753	0.9312	0.9567	0.9341	0.9450	0.9248	0.9788	0.9312	0.9544	0.9339	0.9447
10	7	0.9328	0.9949	0.9424	0.9731	0.9444	0.9540	0.9339	0.9884	0.9422	0.9687	0.9446	0.9551
10	3	0.9091	0.9446	0.9120	0.9259	0.9140	0.9201	0.9096	0.9618	0.9128	0.9300	0.9140	0.9201
10	5	0.9176	0.9483	0.9225	0.9376	0.9244	0.9310	0.9187	0.9500	0.9225	0.9361	0.9242	0.9308
10	7	0.9232	0.9609	0.9296	0.9485	0.9309	0.9371	0.9241	0.9570	0.9295	0.9459	0.9311	0.9377
10	3	0.9105	0.9516	0.9147	0.9322	0.9174	0.9256	0.9114	0.9700	0.9156	0.9368	0.9174	0.9255
10	5	0.9214	0.9622	0.9286	0.9491	0.9309	0.9394	0.9230	0.9635	0.9285	0.9470	0.9307	0.9394
10	7	0.9290	0.9811	0.9375	0.9636	0.9392	0.9473	0.9300	0.9755	0.9375	0.9600	0.9394	0.9480
10	3	0.9046	0.9364	0.9069	0.9185	0.9084	0.9132	0.9042	0.9562	0.9074	0.9243	0.9084	0.9140
10	5	0.9126	0.9367	0.9161	0.9277	0.9176	0.9230	0.9132	0.9412	0.9161	0.9276	0.9175	0.9227
10	7	0.9179	0.9459	0.9227	0.9368	0.9238	0.9286	0.9185	0.9447	0.9225	0.9352	0.9238	0.9292

Table 2.5: Bounds for the probability of the union of events for the case of problems (2.1.5), (2.1.6)

Example 3. Table 2.7 and Table 2.6 show bounds for $P(\nu \geq 1)$ based on $S_{k_1},\ S_{k_2}$.

		without unimodality									
		k_1=1 k_2=3		k_1=1 k_2=4		k_1=2 k_2=3		k_1=2 k_2=4		k_1=3 k_2=4	
n	*M*	*LB*	*UB*	LB	UB	LB	UB	LB	UB	LB	UB
10	3	0.6850	1	0.6480	1	0.5470	1	0.5080	1	0.4330	1
10	5	0.7150	1	0.6780	1	0.5830	1	0.5440	1	0.4670	1
10	7	0.7410	1	0.7060	1	0.6210	1	0.5840	1	0.5100	1
10	3	0.6680	1	0.6320	1	0.5270	1	0.4790	1	0.4030	1
10	5	0.7150	1	0.6790	1	0.5830	1	0.5350	1	0.4570	1
10	7	0.7530	1	0.7160	1	0.6320	1	0.5950	1	0.5200	1
10	3	0.6690	1	0.6330	1	0.5320	1	0.4880	1	0.4140	1
10	5	0.6970	1	0.6610	1	0.5640	1	0.5220	1	0.4460	1
10	7	0.7200	1	0.6830	1	0.5930	1	0.5570	1	0.4830	1
10	3	0.6730	1	0.6360	1	0.5340	1	0.4900	1	0.4150	1
10	5	0.7090	1	0.6720	1	0.5760	1	0.5330	1	0.4550	1
10	7	0.7390	1	0.7030	1	0.6170	1	0.5790	1	0.5050	1
10	3	0.6590	1	0.6230	1	0.5210	1	0.4780	1	0.4050	1
10	5	0.6820	1	0.6460	1	0.5480	1	0.5040	1	0.4300	1
10	7	0.7020	1	0.6660	1	0.5720	1	0.5340	1	0.4610	1

Table 2.6: Bounds based on any two binomial moments, without unimodality condition

		with unimodality									
		k_1=1 k_2=3		k_1=1 k_2=4		k_1=2 k_2=3		k_1=2 k_2=4		k_1=3 k_2=4	
n	*M*	*LB*	*UB*	LB	UB	LB	UB	LB	UB	LB	UB
10	3	0.9160	0.9480	0.9160	0.9540	0.9140	0.9970	0.9130	1	0.9120	1
10	5	0.9240	0.9650	0.9240	0.9700	0.9220	0.9970	0.9210	1	0.9200	1
10	7	0.9300	0.9840	0.9290	0.9900	0.9270	1	0.9260	1	0.9240	1
10	3	0.9070	1	0.9060	1	0.9030	1	0.9020	1	0.9000	1
10	5	0.9230	0.9910	0.9220	1	0.9200	1	0.9190	1	0.9170	1
10	7	0.9320	1	0.9310	1	0.9290	1	0.9280	1	0.9260	1
10	3	0.9080	0.9630	0.9080	0.9800	0.9060	1	0.9050	1	0.9040	1
10	5	0.9180	0.9570	0.9170	0.9640	0.9160	0.9930	0.9150	1	0.9140	1
10	7	0.9230	0.9650	0.9220	0.9720	0.9210	0.9960	0.9200	1	0.9190	1
10	3	0.9100	0.9710	0.9090	0.9880	0.9070	1	0.9060	1	0.9040	1
10	5	0.9220	0.9730	0.9210	0.9810	0.9190	1	0.9180	1	0.9170	1
10	7	0.9290	0.9870	0.9280	0.9960	0.9260	1	0.9250	1	0.9230	1
10	3	0.9040	0.9540	0.9030	0.9710	0.9020	1	0.9020	1	0.9000	1
10	5	0.9130	0.9440	0.9120	0.9510	0.9110	0.9740	0.9100	0.9920	0.9090	1
10	7	0.9180	0.9520	0.9170	0.9570	0.9160	0.9750	0.9160	0.9830	0.9150	1

Table 2.7: Bounds based on any two binomial moments, under unimodality condition

Chapter 3

Programming Under Probabilistic Constraints: A Convexity Theorem

3.1 Introduction

Probabilistic constrained programming belongs to the major approaches for dealing with random parameters in optimization problems. Typical areas of application are economic planning, engineering design, finance, where uncertainties like product demand, meteorological or demographic conditions, currency exchange rates, stock prices etc. enter the inequalities describing the proper working of a system under consideration. The main difficulty of such models is due to (optimal) decisions that have to be taken prior to the observation of random parameters. In this situation, one can hardly find any decision which would definitely exclude later constraint violation caused by unexpected random effects.

Programming under probabilistic constraints was initiated by Charnes, Cooper and Symonds [21] under the title of "chance constrained programming". They imposed individual probabilistic constraint on each stochastic constraint. Joint probabilistic constraints for independent random variables were used first by Miller and Wagner [53]. The general probabilistic constrained stochastic programming model was introduced and studied by Prékopa [60, 62]. The stochastically dependent random elements created challenging mathematical and computational problems.

The basic problems can be formulated in the following manner:

$$\begin{aligned} \min \quad & h(x) \\ \text{subject to} \quad & \\ & h_0(x) = P(g_1(x,\xi) \geq 0, ..., g_r(x,\xi) \geq 0) \geq p \\ & h_1(x) \geq p_1, ..., h_m(x) \geq p_m \ , \end{aligned} \tag{3.1.1}$$

and

$$\begin{aligned} \max \quad & h_0(x) = P(g_1(x,\xi) \geq 0, ..., g_r(x,\xi) \geq 0) \geq p \\ \text{subject to} \quad & \\ & h_1(x) \geq p_1, ..., h_m(x) \geq p_m \ , \end{aligned} \tag{3.1.2}$$

where $g_1(x,y)$,...,$g_r(x,y)$ are functions of $x \in \mathbb{R}^n$, $y \in \mathbb{R}^q$; $h(x)$, $h_1(x), ..., h_m(x)$ are functions of $x \in \mathbb{R}^n$; $p, p_1, ..., p_m$ are constants, $0 < p < 1$ and $\xi = (\xi_1, ..., \xi_q)^T$ is a random vector.

An important special case of problem (3.1.1) is

$$\begin{aligned} \min \quad & c^T\mathbf{x} \\ \text{subject to} \quad & \\ & P(T\mathbf{x} \geq \xi) \geq p \\ & A\mathbf{x} \geq b \\ & \mathbf{x} \geq 0 \ . \end{aligned} \tag{3.1.3}$$

In [60, 62, 74, 79] convexity theorems and algorithms have been presented for the solution of problem (3.1.3) and the companion problem

$$\begin{aligned} \max \quad & P(T\mathbf{x} \geq \xi) \geq p \\ \text{subject to} \quad & \\ & Ax \geq b \\ & x \geq 0 \ . \end{aligned} \tag{3.1.4}$$

In the above problem formulations we allow T and ξ to be random.

This chapter is devoted to programming under probabilistic constraints with random coefficient matrix. The diet problem is of this type and is to find the cheapest combination of foods that will satisfy all the daily nutritional requirements of a person (see Balintfy and Amstrong [3]).

In the theory of programming under probabilistic constraints an important problem is the convexity of the set $\{x \mid h_0(x) \geq p\}$. More specifically we are interested in the behavior of the function $h_0(x)$ from the point of view of the concavity or quasi-concavity. Before we formulate our statement we need some definitions.

3.2 Preliminary Notions and Theorems

A real function f, defined on a convex set $A \subset \mathbb{R}^n$ is said to be *quasi-concave*, if for every pair $x, y \in A$ and $0 \leq \lambda \leq 1$, we have

$$f(\lambda x + (1-\lambda)y) \geq \min[f(x), f(y)] \ . \tag{3.2.1}$$

If the inequality in (3.2.1) is reversed, then f is said to be *quasi-convex*.

One special case is logconcavity. A nonnegative function f defined on a convex subset A of the space $\mathbb{R}^n$ is said to be *logconcave* if for any pair $x, y \in A$ and $0 < \lambda < 1$ we have the inequality

$$f(\lambda x + (1-\lambda)y) \geq [f(x)]^{\lambda}[f(y)]^{1-\lambda} \ . \tag{3.2.2}$$

If f is positive valued, then this means that $\log f$ is a concave function on A. If the inequality in (4.1.1) is reversed, then f is said to be *logconvex* on A.

Any logconcave (logconvex) function is quasi-concave (quasi-convex) on the same set. A function is quasi-concave (quasi-convex) on A if and only if all sets of the type

$$\{x \mid f(x) \geq b\} \quad (\{x \mid f(x) \leq b\})$$

are convex, where $-\infty < b < \infty$.

The following theorem proved by Prékopa [61, 63] is the following.

Theorem 3. *If $g_1(\mathbf{x}, \mathbf{y}), ..., g_r(\mathbf{x}, \mathbf{y})$ are concave functions in $\mathbb{R}^{n+q}$, where $\mathbf{x} \in \mathbb{R}^n, \mathbf{y} \in \mathbb{R}^q$ and the random vector $\xi \in \mathbb{R}^q$ has logconcave probability distribution, then the function*

$$h_0(x) = P(g_1(\mathbf{x}, \xi) \geq 0, ..., g_r(\mathbf{x}, \xi) \geq 0)$$

is logconcave.

Tamm [92, 93] has observed that the concavity property can be replaced by the weaker quasi-concavity.

Now consider the set

$$D = \{x \mid h(x) \geq p\} \ ,$$

where

$$h(x) = P(T\mathbf{x} \leq \xi) \ ,$$

p is a fixed probability ($0 < p < 1$), T a random matrix and ξ a random vector.

The theorems mentioned above guarantee the quasi-concavity of $h_0(x)$ in problem (3.1.1). However, having a probabilistic constraint of type

$$h(\mathbf{x}) = P(T\mathbf{x} \leq \xi) \geq p$$

with random ξ and T, the above theorems do not apply in a direct manner because the functions $T_1\mathbf{x} - \xi_1$,...,$T_r\mathbf{x} - \xi_r$ are not quasi-concave.

For constant T and continuously distributed random ξ with logconcave p.d.f. Prékopa [61, 63] has proved that $h(\mathbf{x})$ is also a logconcave function. This fact clearly implies the convexity of the set D.

For the case of a random technology matrix a few convexity theorems are also known. Let $r = 1$ and consider the function

$$h(\mathbf{x}) = P(T\mathbf{x} \leq \mathbf{b}) \ ,$$

where T is a random vector and b is a constant. The following theorem was first proved by Kataoka [42] and van de Panne and Popp [57].

Theorem 4. *If T has normal distribution, then the function $h(x)$ is quasi-concave on the set*

$$\left\{\mathbf{x} \mid \ P(T\mathbf{x} \leq \mathbf{b}) \geq \frac{1}{2}\right\} \ .$$

The next theorem and its corollary have been obtained by Prékopa [64].

Theorem 5. *Let ξ be constant and*

$$T = \begin{pmatrix} T_1 \\ \vdots \\ T_r \end{pmatrix}$$

a random matrix with independent, normally distributed rows such that their covariance matrices are constant multiples of each other. Then $h(\mathbf{x})$ is a quasi-concave function in $\mathbb{R}^n$.

Corollary 1. *If just one column of T is random and its elements have a joint normal distribution, then the set of x vectors satisfying*

$$P(T\mathbf{x} \leq 0) \geq p \tag{3.2.3}$$

is convex, provided that $p \geq 1/2$.

Burkauskas [20] has observed that the independence assumption is superfluous.

Note that the set of x vectors satisfying

$$P(T\mathbf{x} \leq \xi) \geq p \tag{3.2.4}$$

is the same as those, satisfying

$$P\left((T, -\xi) \left(\begin{pmatrix} \mathbf{x} \\ x_{n+1} \end{pmatrix} \right) \leq 0 \right) \geq p \ ,$$

where $x_{n+1} = 1$. Hence a statement for the constraint (3.2.3) can be carried over to the constraint (3.2.4).

The quasi-concavity of $h(\mathbf{x})$ implies the convexity of the set D. Similar theorem holds if the columns of T satisfy the condition that their covariance matrices are constant multiples of each other. The right hand side vector ξ may also be random. In this case we assume that it is independent of T and its covariance matrix is a constant multiple of any of the other covariance matrices. For more detailed information the reader is referred to [74].

Recently Henrion [36] proved the convexity of the set D for independent $T_1, ..., T_r$ under same condition.

Theorem 6. *Assume that the rows T_i of T are pair-wise independent and normally distributed with mean μ_i and variance Σ_i. Then D is convex for*

$$p > \Phi(\max\{\sqrt{3}, u^*\}) \ ,$$

where Φ is the one-dimensional standard normal distribution function,

$$u^* = \max_{i=1,\dots,m} 4\lambda_{max}^{(i)}[\lambda_{min}^{(i)}]^{-3/2}||\mu_i|| \ ,$$

and $\lambda_{max}^{(i)}$, $\lambda_{min}^{(i)}$ refer to the largest and smallest eigenvalue of Σ_i, respectively.

In Section 2.3 we show that under same conditions the converse of Theorem 3 is also true.

While the sum of concave functions is also concave, the same is not true, in general, for quasi-concave functions. However, we can define a special class of quasi-concave functions such that the sums and products, within the class, are also quasi-concave.

Definition. Let $h_1(x), ..., h_r(x)$ be quasi-concave functions in a convex set E. We say that they are *uniformly quasi-concave* if for any $x, y \in E$ either

$$min(h_i(x), h_i(y)) = h_i(x) \ , \quad i = 1, ..., r$$

or

$$min(h_i(x), h_i(y)) = h_i(y) \ , \quad i = 1, ..., r \ .$$

Obviously, the sum of uniformly quasi-concave functions, on the same set, is also quasi-concave and if the functions are also nonnegative, then the same holds for their product as well. The latter property is used in the next section, where we prove our main result.

3.3 The Main Theorem

Consider the set

$$D = \{\mathbf{x}|P(T\mathbf{x} \leq b)\} \ , \tag{3.3.1}$$

where T is a random matrix with independent, normally distributed rows and b is a constant vector.

Let r be an arbitrary positive integer and introduce the function:

$$h_i(\mathbf{x}) = P(T_i\mathbf{x} \leq b_i) \ , \quad i = 1, ..., r \ .$$

If $\mu_i = E(T_i) = 0 \ , \ \ i = 1, ..., r$ and $p \geq 1/2$, then the inequality

$$P(T_i\mathbf{x} \leq b_i) \geq P(T\mathbf{x} \leq \mathbf{b}) \geq p$$

shows that we have to assume $b_i \geq 0 \ , \ \ i = 1, ..., r$, otherwise the set (3.3.1) is empty. So, let $b_i \geq 0 \ , \ \ i = 1, ..., r$.

Let Φ designate the c.d.f. of the $N(0,1)$-distribution. We also see that, under the same condition, for any $\mathbf{x}$ that satisfies $\mathbf{x}^T C\mathbf{x} > 0$

$$\begin{aligned} A &= \bigcap_{i=1}^{r} \{\mathbf{x} \mid P(T_i\mathbf{x} \leq b_i) \geq p\} \\ &= \bigcap_{i=1}^{r} \left\{\mathbf{x} \mid P\left(\frac{T_i\mathbf{x}}{\mathbf{x}^T C_i \mathbf{x}} \leq \frac{b_i}{\mathbf{x}^T C_i \mathbf{x}}\right) \geq p\right\} \\ &= \bigcap_{i=1}^{r} \left\{\mathbf{x} \mid \Phi\left(\frac{b_i}{\mathbf{x}^T C_i \mathbf{x}}\right) \geq p\right\} \\ &= \bigcap_{i=1}^{r} \left\{x \mid \sqrt{\mathbf{x}^T C_i \mathbf{x}} \leq \frac{b_i}{\Phi^{-1}(p)}\right\} \ , \end{aligned} \tag{3.3.2}$$

where C_i is covariance matrix of T_i, $i = 1, ..., r$. Hence

$$h_i(\mathbf{x}) = \Phi\left(\frac{b_i}{\sqrt{\mathbf{x}^T C\mathbf{x}}}\right) \ , \quad i = 1, ..., r$$

and

$$\Phi^{-1}(p)\sqrt{\mathbf{x}^T C_i \mathbf{x}} \leq b_i \ , \quad i = 1, ..., r \ . \tag{3.3.3}$$

Since $\Phi^{-1}(p) \geq 0$ and $\sqrt{\mathbf{x}^T C\mathbf{x}}$ is a convex function in $\mathbb{R}^n$, it follows that inequality (3.3.3) determines a convex set. Thus each function h_i is quasi-concave on the set (3.3.2).

Theorem 7. *Let $b_i > 0 \ , \ \ i = 1, ..., r$ and $E(T_i) = 0 \ , \ \ i = 1, ..., r$. If the functions $h_1, ...h_r$ are uniformly quasi-concave on the set A, defined by (3.3.2), and $C_i \neq 0 \ , \ \ i = 1, ..., r$, then each C_i is a constant multiple of a covariance matrix C.*

Proof. We already know that the functions $h_1, ..., h_r$ are all quasi-concave on the set (3.3.2) and that

$$h_i(\mathbf{x}) = \Phi\left(\frac{b_i}{\sqrt{\mathbf{x}^T C\mathbf{x}}}\right) \ , \quad i = 1, ..., r \ .$$

It is enough to show that if we take two functions $h_i, h_j,\ i \neq j$, then the corresponding covariance matrices C_i, C_j are constant multiples of each other. Let h_1 and h_2 be the two functions.

Since h_1 and h_2 are uniformly quasi-concave on A, it follows that for any two vectors $y, z \in E$, the inequality

$$\Phi\left(\frac{b_1}{\sqrt{\mathbf{y}^T C_1 \mathbf{y}}}\right) \geq \Phi\left(\frac{b_1}{\sqrt{\mathbf{z}^T C_1 \mathbf{z}}}\right)$$

implies that

$$\Phi\left(\frac{b_2}{\sqrt{\mathbf{y}^T C_2 \mathbf{y}}}\right) \geq \Phi\left(\frac{b_2}{\sqrt{\mathbf{z}^T C_2 \mathbf{z}}}\right).$$

An equivalent form of the statement is that the inequality

$$\mathbf{z}^T C_1 \mathbf{z} \geq \mathbf{y}^T C_1 \mathbf{y} \tag{3.3.4}$$

implies that

$$\mathbf{z}^T C_2 \mathbf{z} \geq \mathbf{y}^T C_2 \mathbf{y}\ . \tag{3.3.5}$$

From linear algebra we know that for any two quadratic forms, in the same variables, there exists a basis such that both quadratic forms are sums of squares if the variables are expressed in that basis. In our case this means that there exist linearly independent vectors $a_1, ..., a_n$ such that the transformation

$$\mathbf{x} = a_1 u_1 + ... + a_n u_n = A\mathbf{u}\ ,$$

applied to the quadratic forms $\mathbf{x}^T C_1 \mathbf{x}$ and $\mathbf{x}^T C_2 \mathbf{x}$, takes them to the forms

$$\mathbf{x}^T C_1 \mathbf{x} = \mathbf{u}^T A^T C_1 A \mathbf{u} = \lambda_1 u_1^2 + ... + \lambda_n u_n^2$$
$$\mathbf{x}^T C_2 \mathbf{x} = \mathbf{u}^T A^T C_2 A \mathbf{u} = \gamma_1 u_1^2 + ... + \gamma_n u_n^2\ ,$$

where $\lambda_i, \gamma_i \geq 0\ ,\quad i = 1, ..., n$ and $\lambda_1 + ... + \lambda_n > 0\ ,\quad \gamma_1 + ... + \gamma_n > 0$.

The transformation $\mathbf{x} = A\mathbf{u}$ transforms the set A into a set H that is the intersection of ellipsoids with centers in the origin and main axes lying in the coordinate axes.

If

$$\mathbf{u} = A^{-1}\mathbf{y} \text{ and } \mathbf{v} = A^{-1}\mathbf{z} ,$$

then the statement that (3.3.4) implies (3.3.5) can be formulated in such a way that if $\mathbf{u}, \mathbf{v} \in H$, then

$$\lambda_1 v_1^2 + ... + \lambda_n v_n^2 \geq \lambda_1 u_1^2 + ... + \lambda_n u_n^2$$

implies that

$$\gamma_1 v_1^2 + ... + \gamma_n v_n^2 \geq \gamma_1 u_1^2 + ... + \gamma_n u_n^2 .$$

This, in turn, is the same as the statement:

$$\lambda_1(v_1^2 - u_1^2) + ... + \lambda_n(v_n^2 - u_n^2) \geq 0$$

implies that

$$\gamma_1(v_1^2 - u_1^2) + ... + \gamma_n(v_n^2 - u_n^2) \geq 0 .$$

Let us introduce the notation $w_i = v_i^2 - u_i^2 \;, \quad i = 1, ..., n$. Then a further form of the statement is:

$$\lambda_1 w_1 + ... + \lambda_n w_n \geq 0 \tag{3.3.6}$$

implies that

$$\gamma_1 w_1 + ... + \gamma_n w_n \geq 0 . \tag{3.3.7}$$

The above implication is true for any $w = (w_1, ..., w_n)$ in an open convex set around the origin in $\mathbb{R}^n$. It follows that it is also true without any limitation for the variables $w_1, ..., w_n$.

By Farkas' theorem there exists a nonnegative number α such that

$$\gamma = (\gamma_1, ..., \gamma_n) = \alpha(\lambda_1, ..., \lambda_n) = \alpha\lambda . \tag{3.3.8}$$

Since $\gamma \neq 0$, $\lambda \neq 0$, the number α must be positive. The relation can be written in matrix form:

$$\begin{pmatrix} \gamma_1 & & 0 \\ & \ddots & \\ 0 & & \gamma_n \end{pmatrix} = \alpha \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} . \tag{3.3.9}$$

If we take into account that

$$\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = A^T C_1 A$$

$$\begin{pmatrix} \gamma_1 & & 0 \\ & \ddots & \\ 0 & & \gamma_n \end{pmatrix} = A^T C_2 A \ ,$$

and combine it with (3.3.9), we can derive the equation

$$A^T C_2 A = \alpha \ A^T C_1 A \ .$$

Since A is a nonsingular matrix, we conclude that

$$C_2 = \alpha \ C_1 \ .$$

This proves the theorem. $\square$

Chapter 4

Strong Unimodality of Multivariate Discrete Distributions

4.1 Introduction

A probability measure P, defined on $\mathbb{R}^n$, is said to be *logconcave* if for every pair of nonempty convex sets $A, B \subset \mathbb{R}^n$ (any convex set is Borel measurable) and we have the inequality

$$P(\lambda A + (1-\lambda)B) \geq [P(A)]^{\lambda}[P(B)]^{(1-\lambda)},$$

where the $+$ sign refers to Minkowski addition of sets, i.e.,

$$\lambda A + (1-\lambda)B = \{\lambda \mathbf{x} + (1-\lambda)\mathbf{y} | \mathbf{x} \in A, \mathbf{y} \in B\}.$$

The above notion generalizes in a natural way to nonnegative valued measures. In this case we require the logconcavity inequality to hold for finite $P(A), P(B)$. The notion of a logconcave probability measure was introduced in [61, 63].

In 1912 Fekete [29] introduced the notion of an r-times positive sequence. The sequence of nonnegative elements $\ldots, a_{-2}, a_{-1}, a_0, \ldots$ is said to be *r-times positive* if the matrix

$$A = \begin{bmatrix} \ddots & \ddots & \ddots & & \\ \ddots & a_0 & a_1 & a_2 & \\ \ddots & a_{-1} & a_0 & a_1 & \ddots \\ & a_{-2} & a_{-1} & a_0 & \ddots \\ & & \ddots & \ddots & \ddots \end{bmatrix}$$

has no negative minor of order smaller than or equal to r.

Twice-positive sequences are those for which we have

$$\begin{vmatrix} a_i & a_j \\ a_{i-t} & a_{j-t} \end{vmatrix} = a_i a_{j-t} - a_j a_{i-t} \geq 0. \tag{4.1.1}$$

for every $i < j$ and $t \geq 1$. This holds if and only if

$$a_i^2 \geq a_{i-1}a_{i+1} \ .$$

Fekete [29] also proved that the convolution of two r-times positive sequences is r-times positive. Twice-positive sequences are also called *logconcave sequences*. For this, Fekete's theorem states that the convolution of two logconcave sequences is logconcave.

A discrete probability distribution, defined on the real line, is said to be logconcave if the corresponding probability function is logconcave.

In what follows we present our results in terms of probability functions. They generalize in a straightforward manner for more general logconcave functions.

Let $\mathbb{Z}^n$ designate the set of lattice points in the space. The convolution of two logconcave distributions on $\mathbb{Z}^n$ is no longer logconcave in general, if $n \geq 2$.

Another notion is strong unimodality. Following Barndorff-Nielsen [5] a discrete probability function $p(\mathbf{z}), \mathbf{z} \in \mathbb{Z}^n$ is called strongly unimodal if there exists a convex function $f(\mathbf{x}),\ \mathbf{x} \in \mathbb{R}^n$ such that $f(\mathbf{x}) = -\log p(\mathbf{x})$ if $\mathbf{x} \in \mathbb{Z}^n$. If $p(\mathbf{z}) = 0$, then by definition $f(\mathbf{z}) = \infty$. This notion is not a direct generalization of that of the one-dimensional case, i.e., of formula 4.1.1. However in case of $n = 1$ the two notions are the same (see, e.g., [74]). It is trivial that if p is strongly unimodal, then it is logconcave. The joint probability function of a finite number of mutually independent discrete random variables, where each has a logconcave probability function is strongly unimodal.

Pedersen [58] gave the following two sufficient conditions for a bivariate discrete distribution to be strongly unimodal. Let p be a discrete probability function on $\mathbb{Z}^2$ and p_{ij} denote the value of p on $(i, j) \in \mathbb{Z}^2$. It is sufficient for p to be strongly unimodal if it satisfies one of the following conditions (a) and (b):

$$\begin{aligned}
(a)\quad & p_{i-1,j}p_{i,j-1} \geq p_{ij}p_{i-1,j-1} \ , & (b)\quad & p_{ij}p_{i-1,j-1} \geq p_{i-1,j}p_{i,j-1} \ , \\
& p_{i-1,j}p_{ij} \geq p_{i,j-1}p_{i-1,j+1} \ , & & p_{ij}p_{i-1,j} \geq p_{i,j+1}p_{i-1,j-1} \ , \\
& p_{ij}p_{i,j-1} \geq p_{i-1,j}p_{i+1,j-1} \ , & & p_{ij}p_{i,j-1} \geq p_{i+1,j}p_{i-1,j-1} \ .
\end{aligned}$$

Pedersen [58] also proved that the trinomial probability function is logconcave and the convolution of any finite number of these distributions with possibly different parameter sets is also logconcave.

The notion of discrete unimodality is of interest in connection with statistical physics where a typical problem is to find the maximum of a unimodal probability function.

A function $f(\mathbf{z}), \mathbf{z} \in \mathbb{R}^n$ is said to be polyhedral (simplicial) on the bounded convex polyhedron $K \subseteq \mathbb{R}^n$ if there exists a subdivision of K into n-dimensional convex polyhedra (simplices), with pairwise disjoint interiors such that f is continuous on K and linear on each subdividing polyhedron (simplex). Prékopa and Li [75] presented a dual method to solve a linearly constrained optimization problem with convex, polyhedral objective function, along with a fast bounding technique, for the optimum value. Any $f(x)$, defined by the use of a strongly unimodal probability function $p(x)$, is a simplicial function and can be used in the above-mentioned methodology. In an earlier paper [71] Prékopa developed a dual type method for the solution of a one-stage stochastic programming problems. The method was improved and implemented in [27].

In Section 4.2 we give sufficient conditions for a trivariate discrete distribution to be strongly unimodal. In Section 4.3 we present a sufficient condition that ensures the strong unimodality of a multivariate discrete distribution. In Section 4.4 we use the results of the paper by Prékopa and Li [75] and present a dual type algorithm to find the maximum of a strongly unimodal multivariate discrete distribution. In Section 4.5 we present four multivariate discrete distributions that are strongly unimodal.

4.2 Sufficient Conditions for a Trivariate Discrete Distribution to be Strongly Unimodal

In this section we give sufficient conditions for a discrete probability function defined on $\mathbb{Z}^3$ that ensure its strong unimodality. The function f defined on $\mathbb{R}^3$ that we fit to the values of $-\log p(.)$ is piecewise linear. We accomplish the job in such a way that we subdivide $\mathbb{R}^3$ into simplices with disjoint interiors such that the function $f(\mathbf{x})$ is linear on each of them. First we subdivide $\mathbb{R}^3$ into unit cubes and then subdivide each cube into six simplices with disjoint interiors. In each cube the same type of subdivision is used. On each simplex we define $f(\mathbf{x})$ by the equation of the hyperplane determined by the values of $-\log p(\mathbf{x})$ at the vertices. Next we ensure that $f(\mathbf{x})$ is convex on any two neighboring simplices (two simplices having a common facet). The resulting function $f(\mathbf{x})$ is convex on the entire space.

Any cube in $\mathbb{R}^3$ can be subdivided into simplices with disjoint interiors (such that the vertices of the simplices are those of the cube) in six different ways as in Figures 4.6-**??**. In the following we present vertices of the subdividing simplices.

Subdivision 1. Let $T_{1c}(i,j,k),\ c = 1,...,6$ be the simplices in $\mathbb{R}^3$ defined by

$$T_{11}(i,j,k) = conv\{(i,j,k),(i+1,j,k),(i+1,j+1,k),(i+1,j+1,k+1)\},$$
$$T_{12}(i,j,k) = conv\{(i,j,k),(i+1,j,k),(i+1,j,k+1),(i+1,j+1,k+1)\},$$
$$T_{13}(i,j,k) = conv\{(i,j,k),(i,j+1,k),(i+1,j+1,k),(i+1,j+1,k+1)\},$$
$$T_{14}(i,j,k) = conv\{(i,j,k),(i,j+1,k),(i,j+1,k+1),(i+1,j+1,k+1)\},$$
$$T_{15}(i,j,k) = conv\{(i,j,k),(i,j,k+1),(i+1,j,k+1),(i+1,j+1,k+1)\},$$
$$T_{16}(i,j,k) = conv\{(i,j,k),(i,j,k+1),(i,j+1,k+1),(i+1,j+1,k+1)\}.$$

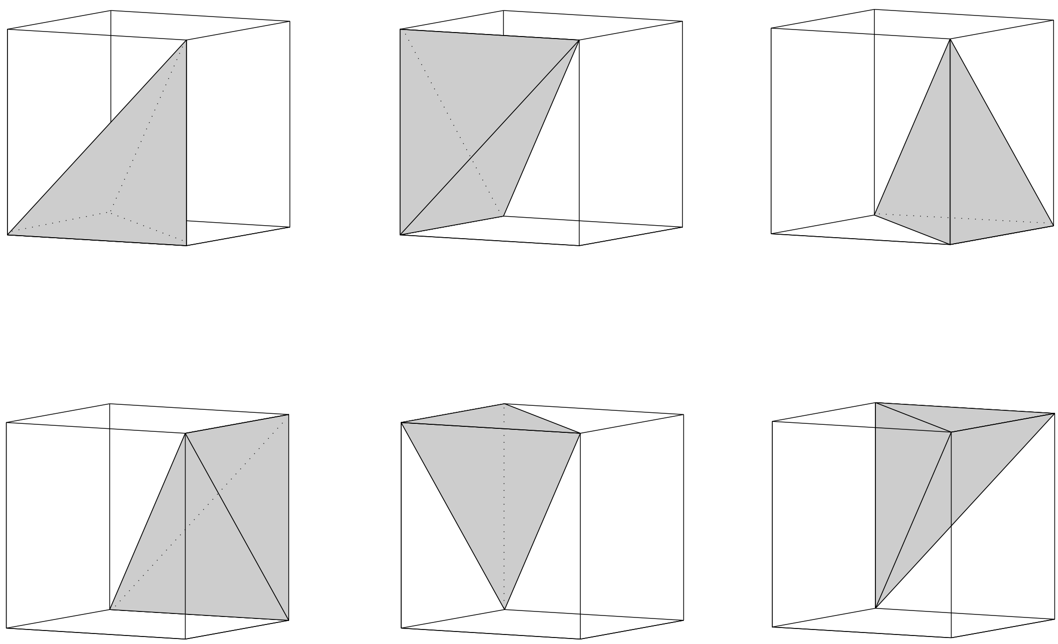

Figure 4.1: Subdivision 1

Subdivision 2. Let $T_{2c}(i,j,k),\ c=1,...,6$ be the simplices in $\mathbb{R}^3$ defined by

$$T_{21}(i,j,k) = conv\{(i,j,k),(i,j,k+1),(i+1,j,k+1),(i+1,j+1,k+1)\},$$

$$T_{22}(i,j,k) = conv\{(i,j,k),(i,j,k+1),(i,j+1,k+1),(i+1,j+1,k+1)\},$$

$$T_{23}(i,j,k) = conv\{(i,j,k),(i,j+1,k),(i,j+1,k+1),(i+1,j+1,k+1)\},$$

$$T_{24}(i,j,k) = conv\{(i,j,k),(i+1,j,k),(i+1,j,k+1),(i+1,j+1,k+1)\},$$

$$T_{25}(i,j,k) = conv\{(i,j,k),(i+1,j,k),(i,j+1,k),(i+1,j+1,k+1)\},$$

$$T_{26}(i,j,k) = conv\{(i+1,j,k),(i,j+1,k),(i+1,j+1,k),(i+1,j+1,k+1)\}.$$

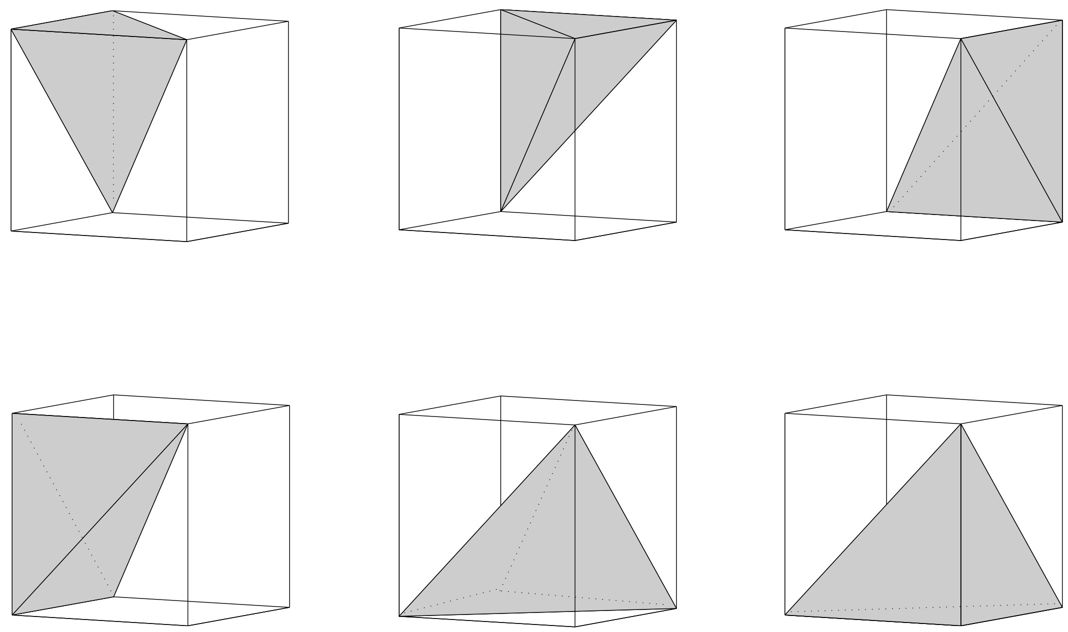

Figure 4.2: Subdivision 2

Subdivision 3. Let $T_{3c}(i,j,k),\ c = 1,...,6$ be the simplices in $\mathbb{R}^3$ defined by

$$T_{31}(i,j,k) = conv\{(i,j,k),(i,j,k+1),(i+1,j,k+1),(i,j+1,k+1)\},$$

$$T_{32}(i,j,k) = conv\{(i,j,k),(i+1,j,k+1),(i,j+1,k+1),(i+1,j+1,k+1)\},$$

$$T_{33}(i,j,k) = conv\{(i,j,k),(i+1,j+1,k),(i,j+1,k+1),(i+1,j+1,k+1)\},$$

$$T_{34}(i,j,k) = conv\{(i,j,k),(i,j+1,k),(i+1,j+1,k),(i,j+1,k+1)\},$$

$$T_{35}(i,j,k) = conv\{(i,j,k),(i+1,j,k+1),(i+1,j,k),(i+1,j+1,k+1)\},$$

$$T_{36}(i,j,k) = conv\{(i,j,k),(i+1,j+1,k),(i+1,j,k),(i+1,j+1,k+1)\}.$$

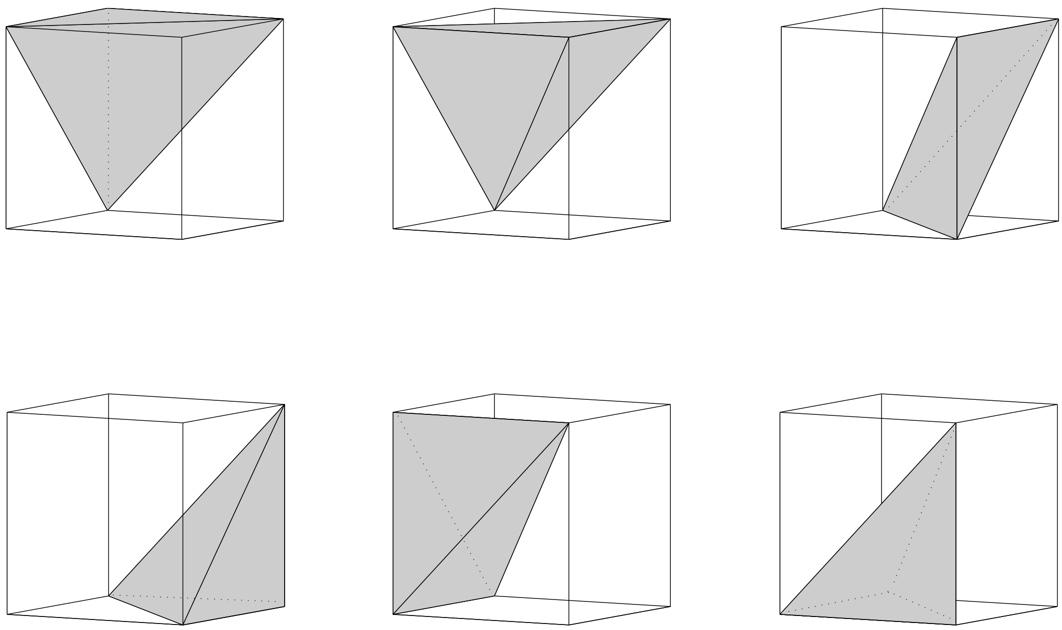

Figure 4.3: Subdivision 3

Subdivision 4. Let $T_{4c}(i,j,k),\ c = 1,...,6$ be the simplices in $\mathbb{R}^3$ defined by

$$T_{41}(i,j,k) = conv\{(i,j,k),(i,j,k+1),(i+1,j,k+1),(i,j+1,k+1)\},$$

$$T_{42}(i,j,k) = conv\{(i,j,k),(i+1,j,k+1),(i,j+1,k+1),(i+1,j+1,k+1)\},$$

$$T_{43}(i,j,k) = conv\{(i,j,k),(i,j+1,k),(i,j+1,k+1),(i+1,j+1,k+1)\},$$

$$T_{44}(i,j,k) = conv\{(i,j,k),(i,j+1,k),(i+1,j+1,k),(i+1,j+1,k+1)\},$$

$$T_{45}(i,j,k) = conv\{(i,j,k),(i+1,j,k),(i+1,j,k+1),(i+1,j+1,k+1)\},$$

$$T_{46}(i,j,k) = conv\{(i,j,k),(i+1,j,k+1),(i+1,j,k),(i+1,j+1,k+1)\}.$$

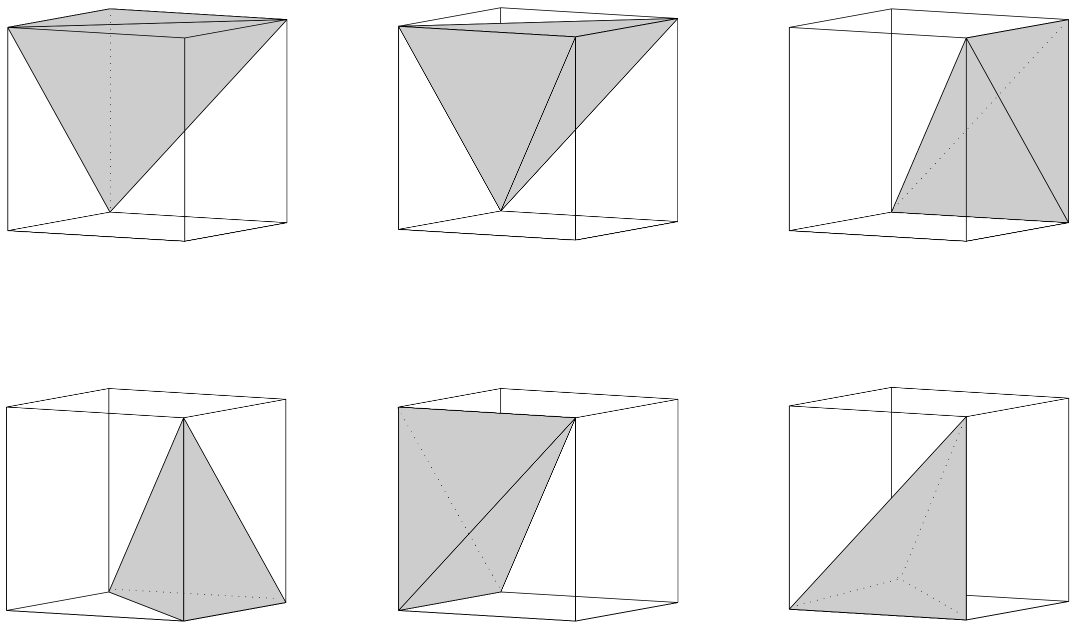

Figure 4.4: Subdivision 4

Subdivision 5. Let $T_{5c}(i,j,k),\ c = 1, ..., 6$ be the simplices in $\mathbb{R}^3$ defined by

$$T_{51}(i,j,k) = conv\{(i,j,k),(i,j,k+1),(i+1,j,k+1),(i,j+1,k+1)\},$$

$$T_{52}(i,j,k) = conv\{(i,j,k),(i+1,j,k+1),(i,j+1,k+1),(i+1,j+1,k+1)\},$$

$$T_{53}(i,j,k) = conv\{(i,j,k),(i+1,j+1,k),(i,j+1,k+1),(i+1,j+1,k+1)\},$$

$$T_{54}(i,j,k) = conv\{(i,j,k),(i,j+1,k),(i,j+1,k+1),(i+1,j+1,k)\},$$

$$T_{55}(i,j,k) = conv\{(i,j,k),(i+1,j+1,k),(i+1,j,k+1),(i+1,j+1,k+1)\},$$

$$T_{56}(i,j,k) = conv\{(i,j,k),(i+1,j,k),(i+1,j+1,k),(i+1,j,k+1)\}.$$

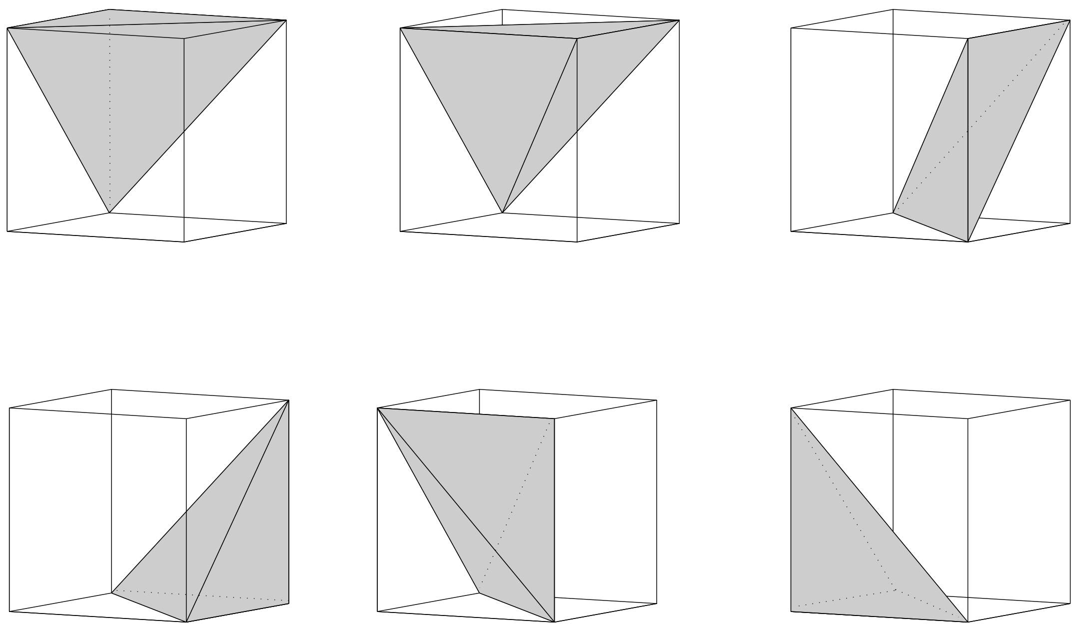

Figure 4.5: Subdivision 5

Subdivision 6. Let $T_{6c}(i,j,k),\ c = 1, ..., 6$ be the simplices in $\mathbb{R}^3$ defined by

$$T_{61}(i,j,k) = conv\{(i,j,k),(i,j,k+1),(i+1,j,k+1),(i,j+1,k+1)\},$$

$$T_{62}(i,j,k) = conv\{(i,j,k),(i+1,j,k+1),(i,j+1,k+1),(i+1,j+1,k+1)\},$$

$$T_{63}(i,j,k) = conv\{(i,j,k),(i,j+1,k),(i,j+1,k+1),(i+1,j+1,k+1)\},$$

$$T_{64}(i,j,k) = conv\{(i,j,k),(i+1,j,k),(i+1,j,k+1),(i+1,j+1,k+1)\},$$

$$T_{65}(i,j,k) = conv\{(i,j,k),(i+1,j,k),(i,j+1,k),(i+1,j+1,k+1)\},$$

$$T_{66}(i,j,k) = conv\{(i+1,j,k),(i+1,j+1,k),(i,j+1,k),(i+1,j+1,k+1)\}.$$

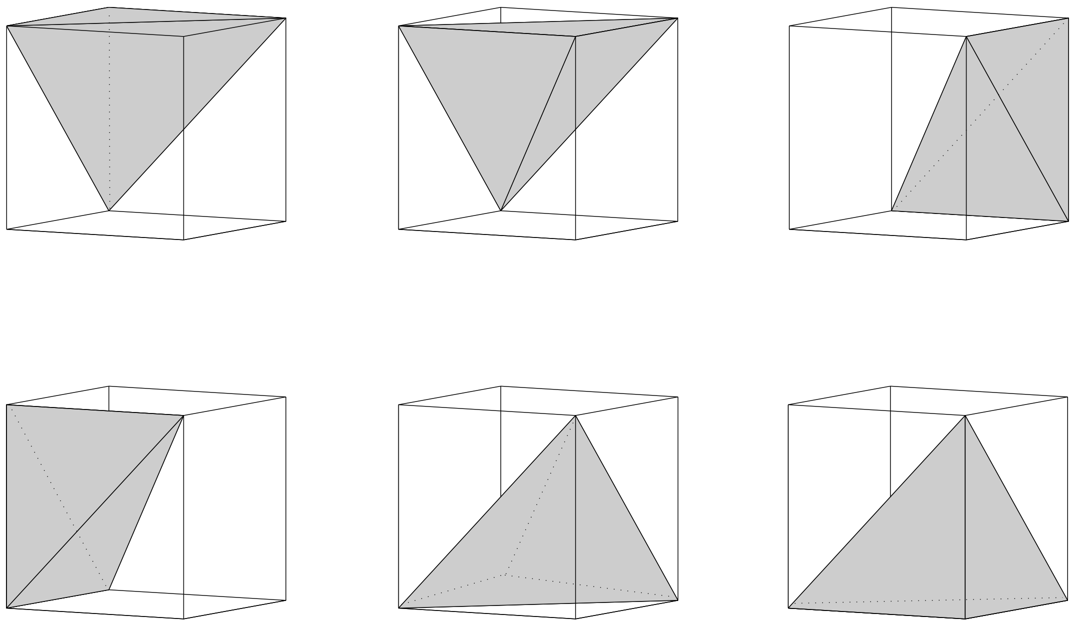

Figure 4.6: Subdivision 6

Let p be the probability function of a discrete probability distribution defined on $\mathbb{R}^3$ and p_{ijk} the value of p at $(i,j,k) \in \mathbb{Z}^3$. Let S denote the support of p. Define $C_t,\ t = 1, ..., 6$ as the collection of the simplices $T_{tc}(i,j,k),\ c = 1, ..., 6,\ (i,j,k) \in \mathbb{Z}^3$, all vertices of which belong to S.

Theorem 8. *If p satisfies one of the following conditions (a), (b), (c), (d), (e), (f) for all $(i,j,k) \in \mathbb{Z}^3$, then it is strongly unimodal.*

(a) *C_1 is the collection of the simplices $T_{1c}(i,j,k), c = 1, 2, ..., 6$ and*

(1) $p_{i+1,jk}p_{i,j+1,k} \leq p_{ijk}p_{i+1,j+1,k}$,

(2) $p_{i+1,jk}p_{ij,k+1} \leq p_{ijk}p_{i+1,j,k+1}$,

(3) $p_{i,j+1,k}p_{ij,k+1} \leq p_{ijk}p_{i,j+1,k+1}$,

(4) $p_{i+1,j+1,k}p_{i+1,j,k+1} \leq p_{i+1,jk}p_{i+1,j+1,k+1}$,

(5) $p_{i+1,j+1,k}p_{i,j+1,k+1} \leq p_{i,j+1,k}p_{i+1,j+1,k+1}$,

(6) $p_{i+1,j,k+1}p_{i,j+1,k+1} \leq p_{ij,k+1}p_{i+1,j+1,k+1}$,

(7) $p_{i-1,jk}p_{i+1,j+1,k+1} \leq p_{ijk}p_{i,j+1,k+1}$,

(8) $p_{i,j-1,k}p_{i+1,j+1,k+1} \leq p_{ijk}p_{i+1,j,k+1}$,

(9) $p_{ij,k-1}p_{i+1,j+1,k+1} \leq p_{ijk}p_{i+1,j+1,k}$,

(10) $p_{ijk}p_{i+2,j+1,k+1} \leq p_{i+1,jk}p_{i+1,j+1,k+1}$,

(11) $p_{ijk}p_{i+1,j+2,k+1} \leq p_{i,j+1,k}p_{i+1,j+1,k+1}$,

(12) $p_{ijk}p_{i+1,j+1,k+2} \leq p_{ij,k+1}p_{i+1,j+1,k+1}$.

(b) *C_2 is the collection of the simplices $T_{2c}(i,j,k), c = 1, 2, ..., 6$ and*

(13) $p_{i+1,j,k+1}p_{i,j+1,k+1} \leq p_{ij,k+1}p_{i+1,j+1,k+1}$,

(14) $p_{i+1,jk}p_{ij,k+1} \leq p_{ijk}p_{i+1,j,k+1}$,

(15) $p_{ij,k+1}p_{i,j+1,k} \leq p_{ijk}p_{i,j+1,k+1}$,

(16) $p_{i,j+1,k+1}p_{i+1,jk} \leq p_{ijk}p_{i+1,j+1,k+1}$,

(17) $p_{i+1,j,k+1}p_{i,j+1,k} \leq p_{ijk}p_{i+1,j+1,k+1}$,

(18) $p_{i+1,j+1,k}p_{ijk} \leq p_{i+1,jk}p_{i,j+1,k}$,

(19) $p_{i,j-1,k}p_{i+1,j+1,k+1} \leq p_{ijk}p_{i+1,j,k+1}$,

(20) $p_{i+1,j+2,k+1}p_{ijk} \leq p_{i,j+1,k}p_{i+1,j+1,k+1}$,

(21) $p_{i-1,j+1,k}p_{i+1,j+1,k+1} \leq p_{i,j+1,k}p_{i,j+1,k+1}$,

(22) $p_{i+2,j+1,k+1}p_{ijk} \le p_{i+1,jk}p_{i+1,j+1,k+1}$,

(23) $p_{i+1,j-1,k}p_{i+1,j+1,k+1} \le p_{i+1,jk}p_{i+1,j,k+1}$,

(24) $p_{i-1,jk}p_{i+1,j+1,k+1} \le p_{ijk}p_{i,j+1,k+1}$,

(25) $p_{i+1,j+2,k+1}p_{i+1,jk} \le p_{i+1,j+1,k}p_{i+1,j+1,k+1}$,

(26) $p_{i+2,j+1,k+1}p_{i,j+1,k} \le p_{i+1,j+1,k}p_{i+1,j+1,k+1}$.

(c) C_3 *is the collection of the simplices* $T_{3c}(i,j,k), c = 1, 2, ..., 6$ *and*

(27) $p_{ij,k+1}p_{i+1,j+1,k+1} \le p_{i,j+1,k+1}p_{i+1,j,k+1}$,

(28) $p_{i+1,j+1,k}p_{i+1,j,k+1}p_{i,j+1,k+1} \le p_{ijk}p^2_{i+1,j+1,k+1}$,

(29) $p_{i,j+1,k}p_{i+1,j+1,k+1} \le p_{i+1,j+1,k}p_{i,j+1,k+1}$,

(30) $p_{i+1,jk}p_{i,j+1,k+1} \le p_{ijk}p_{i+1,j+1,k+1}$,

(31) $p_{i+1,j+1,k}p_{i+1,j,k+1} \le p_{i+1,jk}p_{i+1,j+1,k+1}$,

(32) $p_{i-1,jk}p_{i+1,j,k+1} \le p_{ijk}p_{ij,k+1}$,

(33) $p_{i+2,j,k+1}p_{ijk} \le p_{i+1,jk}p_{i+1,j,k+1}$,

(34) $p_{i+2,j+1,k}p_{ijk} \le p_{i+1,jk}p_{i+1,j+1,k}$,

(35) $p_{i-1,jk}p_{i+1,j+1,k} \le p_{ijk}p_{i,j+1,k}$.

(d) C_4 *is the collection of the simplices* $T_{4c}(i,j,k), c = 1, 2, ..., 6$ *and*

(36) $p_{ij,k+1}p_{i+1,j+1,k+1} \le p_{i,j+1,k+1}p_{i+1,j,k+1}$,

(37) $p_{i+1,j,k+1}p_{i,j+1,k} \le p_{ijk}p_{i+1,j+1,k+1}$,

(38) $p_{i+1,j+1,k}p_{i,j+1,k+1} \le p_{i,j+1,k}p_{i+1,j+1,k+1}$,

(39) $p_{i+1,jk}p_{i,j+1,k+1} \le p_{ijk}p_{i+1,j+1,k+1}$,

(40) $p_{i+1,jk}p_{i,j+1,k} \le p_{ijk}p_{i+1,j+1,k}$,

(41) $p_{i+1,j+1,k}p_{i+1,j,k+1} \le p_{i+1,jk}p_{i+1,j+1,k+1}$,

(42) $p_{i-1,jk}p_{i+1,j,k+1} \le p_{ijk}p_{ij,k+1}$,

(43) $p_{i,j+2,k+1}p_{ijk} \le p_{i,j+1,k}p_{i,j+1,k+1}$,

(44) $p_{i-1,jk}p_{i+1,j+1,k+1} \le p_{ijk}p_{i,j+1,k+1}$,

(45) $p_{ijk}p_{i+1,j+2,k+1} \leq p_{i,j+1,k}p_{i+1,j+1,k+1}$,

(46) $p_{ijk}p_{i+2,j,k+1} \leq p_{i+1,jk}p_{i+1,j,k+1}$,

(47) $p_{i,j-1,k}p_{i+1,j+1,k+1} \leq p_{ijk}p_{i+1,j,k+1}$,

(48) $p_{ijk}p_{i+2,j+1,k+1} \leq p_{i+1,jk}p_{i+1,j+1,k+1}$.

(e) C_5 *is the collection of the simplices* $T_{5c}(i,j,k), c = 1, 2, ..., 6$ *and*

(49) $p_{ij,k+1}p_{i+1,j+1,k+1} \leq p_{i,j+1,k+1}p_{i+1,j,k+1}$,

(50) $p_{i+1,j+1,k}p_{i+1,j,k+1}p_{i,j+1,k+1} \leq p_{ijk}p^2_{i+1,j+1,k+1}$,

(51) $p_{i+1,j+1,k+1}p_{i,j+1,k} \leq p_{i,j+1,k+1}p_{i+1,j+1,k}$,

(52) $p_{i+1,jk}p_{i+1,j+1,k+1} \leq p_{i+1,j+1,k}p_{i+1,j,k+1}$.

(f) C_6 *is the collection of the simplices* $T_{6c}(i,j,k), c = 1, 2, ..., 6$ *and*

(53) $p_{ij,k+1}p_{i+1,j+1,k+1} \leq p_{i,j+1,k+1}p_{i+1,j,k+1}$,

(54) $p_{i,j+1,k}p_{i+1,j,k+1} \leq p_{ijk}p_{i+1,j+1,k+1}$,

(55) $p_{i+1,jk}p_{i,j+1,k+1} \leq p_{ijk}p_{i+1,j+1,k+1}$,

(56) $p_{i+1,j+1,k}p_{ijk} \leq p_{i+1,jk}p_{i,j+1,k}$,

(57) $p_{ijk}p_{i+1,j+1,k+2} \leq p_{i+1,j,k+1}p_{i,j+1,k+1}$,

(58) $p_{i-1,jk}p_{i+1,j,k+1} \leq p_{ijk}p_{i,j+1,k+1}$,

(59) $p_{i,j-1,k}p_{i,j+1,k+1} \leq p_{ijk}p_{ij,k+1}$,

(60) $p_{ijk}p_{i,j+2,k+1} \leq p_{i,j+1,k}p_{i,j+1,k+1}$,

(61) $p_{i-1,j+1,k}p_{i+1,j+1,k+1} \leq p_{i,j+1,k}p_{i,j+1,k+1}$,

(62) $p_{ijk}p_{i+2,j,k+1} \leq p_{i+1,jk}p_{i+1,j,k+1}$,

(63) $p_{i+1,j-1,k}p_{i+1,j+1,k+1} \leq p_{i+1,jk}p_{i+1,j,k+1}$,

(64) $p_{ij,k-1}p_{i+1,j+1,k+1} \leq p_{i+1,jk}p_{i,j+1,k}$,

(65) $p_{i,j+1,k}p_{i+2,j+1,k+1} \leq p_{i+1,j+1,k}p_{i+1,j+1,k+1}$,

(66) $p_{ijk}p_{i+1,j+2,k+1} \leq p_{i+1,j+1,k}p_{i+1,j+1,k+1}$.

Proof. We prove the sufficiency of (a). We subdivide $\mathbb{R}^3$ into unit cubes and then subdivide each cube into six simplices with disjoint interiors. We assume that in each cube Subdivision 1 is used.

Let $L(c,i,j,k),\ (i,j,k) \in \mathbb{Z}^3,\ c = 1,2,...,6$ designate the linear function on $\mathbb{R}^3$ that coincides with $-\log p(.)$ on the vertices of $T_{1c}(i,j,k)$ and define

$$f(\mathbf{y}) = \begin{cases} L(c,i,j,k) & \text{if} \quad \mathbf{y} \in T_{1c}(i,j,k), (i,j,k) \in \mathbb{Z}^3 \\ \infty & \text{if} \quad \mathbf{y} \notin C_1 \end{cases}$$

Obviously, f coincides on $\mathbb{Z}^3$ with $-\log p(.)$.

Claim: Conditions $(1),...,(12)$ ensure that the restriction of f to any two simplices in $\mathbb{Z}^3$, having a facet in common, is convex.

Proof of the claim: We prove that for any two neighboring simplices f satisfies the convexity property. On each simplex we define a linear function. In case of any simplex a linear piece is determined by the vertices of the simplex and the corresponding values of $-\log p(.)$. The collection of these linear pieces form the function f.

The function f is convex on any two neighboring simplices with a common facet if for any lattice points

$$\mathbf{z}_i = (z_{i1}, z_{i2}, z_{i3}),\ i = 0,1,2,3 \ \text{ and } \ \mathbf{y}_0 = (y_1, y_2, y_3)$$

such that $\mathbf{z}_i$ are the vertices of a simplex in Subdivision 1 and y is the vertex of a neighboring simplex which does not belong to the current one, we have the relation

$$\frac{\begin{vmatrix} f(\mathbf{y}) & f(\mathbf{z}_0) & f(\mathbf{z}_1) & f(\mathbf{z}_2) & f(\mathbf{z}_3) \\ 1 & 1 & 1 & 1 & 1 \\ y_1 & z_{01} & z_{11} & z_{21} & z_{31} \\ y_2 & z_{02} & z_{12} & z_{22} & z_{32} \\ y_3 & z_{03} & z_{13} & z_{23} & z_{33} \end{vmatrix}}{\begin{vmatrix} 1 & 1 & 1 & 1 \\ z_{01} & z_{11} & z_{21} & z_{31} \\ z_{02} & z_{12} & z_{22} & z_{32} \\ z_{03} & z_{13} & z_{23} & z_{33} \end{vmatrix}} \geq 0\ . \tag{4.2.1}$$

First note that any simplex of type $T_{1c},\ c = 1, ..., 6$ has four neighbors: two in the same cube and two in different cubes. We prove that inequalities $(1), ..., (6)$ ensure the convexity of f within a cube and inequalities $(7), ..., (12)$ within two neighboring simplices that are in different cubes.

We consider simplex $T_{11}(i, j, k)$ whose neighbors are:

$$conv\{(i,j,k),(i+1,j,k),(i+1,j,k+1),(i+1,j+1,k+1)\},$$
$$conv\{(i,j,k),(i,j+1,k),(i+1,j+1,k),(i+1,j+1,k+1)\},$$
$$conv\{(i+2,j+1,k+1),(i+1,j,k),(i+1,j+1,k),(i+1,j+1,k+1)\},$$
$$conv\{(i,j,k-1),(i,j,k),(i+1,j,k),(i+1,j+1,k)\}.$$

Note that the first two simplices and $T_{11}(i, j, k)$ are in the same cube whereas the last two are in two different cubes.

Let $\mathbf{z}_i, i = 0, 1, 2, 3$ be the vertices of simplex T_{11} and $\mathbf{y}$ the vertex of its first neighbor that does not belong to T_{11}. In this case inequality (4.2.1) can be written as

$$\frac{\begin{vmatrix} f(\mathbf{y}) & f(\mathbf{z}_0) & f(\mathbf{z}_1) & f(\mathbf{z}_2) & f(\mathbf{z}_3) \\ 1 & 1 & 1 & 1 & 1 \\ i+1 & i & i+1 & i+1 & i+1 \\ j & j & j & j+1 & j+1 \\ k+1 & k & k & k & k+1 \end{vmatrix}}{\begin{vmatrix} 1 & 1 & 1 & 1 \\ i & i+1 & i+1 & i+1 \\ j & j & j+1 & j+1 \\ k & k & k & k+1 \end{vmatrix}} \geq 0\ , \tag{4.2.2}$$

where $f = -\log p(.)$. Since the denominator in (4.2.2) is equal to 1 the convexity of f is satisfied if the numerator is nonnegative. Therefore we need to ensure that

$$\begin{vmatrix} f(\mathbf{y}) & f(\mathbf{z}_0) & f(\mathbf{z}_1) & f(\mathbf{z}_2) & f(\mathbf{z}_3) \\ 1 & 1 & 1 & 1 & 1 \\ i+1 & i & i+1 & i+1 & i+1 \\ j & j & j & j+1 & j+1 \\ k+1 & k & k & k & k+1 \end{vmatrix} =$$

$$\begin{vmatrix} f(\mathbf{y})-f(\mathbf{z}_0) & f(\mathbf{z}_0) & f(\mathbf{z}_1)-f(\mathbf{z}_0) & f(\mathbf{z}_2)-f(\mathbf{z}_1) & f(\mathbf{z}_3)-f(\mathbf{z}_2) \\ 0 & 1 & 0 & 0 & 0 \\ 1 & i & 1 & 0 & 0 \\ 0 & j & 0 & 1 & 0 \\ 1 & k & 0 & 0 & 1 \end{vmatrix} \geq 0 \ .$$

It follows that $f(\mathbf{y})+f(\mathbf{z}_2) \geq f(\mathbf{z}_1)+f(\mathbf{z}_3)$. This is, however, the same as

$$p(\mathbf{y})p(\mathbf{z}_2) \leq p(\mathbf{z}_1)p(\mathbf{z}_3) \quad \text{or} \quad p_{i+1,j,k+1}p_{i+1,j+1,k} \leq p_{i+1,jk}p_{i+1,j+1,k+1}$$

which is inequality (4) of condition (a). Taking $\mathbf{y}=(i,j+1,k)$ we obtain inequality (1) of condition (a). Inequalities (2), (3), (5), (6) can be obtained by considering any two neighboring simplices having a common facet in the same cube.

Now, let $\mathbf{y}=(i+2,j+1,k+1)$ and $\mathbf{z}_i, i=0,1,2,3$ defined as before. In this case (4.2.1) provides us with the following inequality:

$$p_{ijk}p_{i+2,j+1,k+1} \leq p_{i+1,jk}p_{i+1,j+1,k+1} \ ,$$

that is inequality (10). If we take $\mathbf{y}=(i,j,k-1)$, then we obtain inequality (9).

We remark that in case of Subdivision 1, there are 12 possible layouts of neighboring simplices in different cubes. However, they only provide us with six additional conditions, i.e., inequalities (7), ..., (12). For example, inequality (10) is also obtained if we consider the following two simplices

$$conv\{(i,j,k),(i+1,j,k),(i+1,j,k+1),(i+1,j+1,k+1)\},$$
$$conv\{(i,j,k),(i+1,j,k),(i+1,j,k+1),(i+2,j+1,k+1)\}.$$

Thus, the claim is true.

As C_1 is the collection of the simplices $T_{1c}(i,j,k),(i,j,k)\in\mathbb{Z}^3,\ c=1,...,6$ and f is convex on any two neighboring simplices, it is convex on the entire space. If (a) is satisfied, then p is strongly unimodal. The sufficiency of (b), (c), (d), (e), (f) can be proved similarly. $\square$

Consider a lattice point $(x,y,z)\in\mathbb{Z}^3$. For the sake of simplicity assume that $x,y,z\in\{0,1\}$. We have the following six (3!) possibilities in connection with the components of the lattice point (x, y, z):

$$\begin{aligned} &x \le y \le z \ , \ x \le z \le y \ , \\ &y \le x \le z \ , \ y \le z \le x \ , \\ &z \le x \le y \ , \ z \le y \le x \ . \end{aligned} \tag{4.2.3}$$

The binary vectors satisfying (4.2.3) are the vertices of the simplices in Subdivision 1. For example, the relation $x \le y \le z$ provides us with the lattice points:

$$(0,0,0), (0,0,1), (0,1,1), (1,1,1)$$

that are the vertices of the simplex of type T_{15}. In the following section we use this idea to find a subdivision of a hypercube and we present the conditions that ensure the strong unimodality of a multivariate discrete distribution.

4.3 A Sufficient Condition for a Multivariate Discrete Distribution to be Strongly Unimodal

We present a sufficient condition that ensures the strong unimodality of a discrete probability function defined on $\mathbb{Z}^n$. The function f defined on $\mathbb{R}^n$ that we fit to the values of $-\log p(.)$ is piecewise linear. In view of this we need a subdivision of $\mathbb{R}^n$ into non-overlapping convex polyhedra such that $f(\mathbf{x}) = -\log p(\mathbf{x}),\ \mathbf{x} \in \mathbb{Z}^n$, is linear on each of them.

We consider a lattice point $(i_1, i_2, ..., i_n)$ and the $n!$ possibilities in connection with the values of its components. Each of those $n!$ relations provides us with the vertices of a subdividing simplex as we have discussed in Section 4.2. Let $\mathcal{S}_1, ..., \mathcal{S}_{n!}$ designate the resulting subdividing simplices of a hypercube:

$$\begin{aligned} \mathcal{S}_1 = \ & conv\{(i_1, ..., i_n),\ (i_1+1, i_2, i_3, i_4, ..., i_n),\ (i_1+1, i_2+1, i_3, i_4, ..., i_n), \\ & (i_1+1, i_2+1, i_3+1, i_4, ..., i_n),\ \ \ldots\ ,\ (i_1+1, ..., i_n+1)\}, \\ \mathcal{S}_2 = \ & conv\{(i_1, ..., i_n),\ (i_1+1, i_2, i_3, i_4, ..., i_n),\ (i_1+1, i_2, i_3+1, i_4, ..., i_n), \\ & (i_1+1, i_2+1, i_3+1, i_4, ..., i_n),\ \ \ldots\ ,\ (i_1+1, ..., i_n+1)\}, \\ \vdots \\ \mathcal{S}_n = \ & conv\{(i_1, ..., i_n),\ (i_1+1, i_2, i_3, ..., i_{n-1}, i_n),\ (i_1+1, i_2, i_3, ..., i_{n-1}, i_n+1), \\ & (i_1+1, i_2+1, i_3, ..., i_{n-1}, i_n+1),\ \ \ldots\ ,\ (i_1+1, ..., i_n+1)\}, \end{aligned}$$

$$\begin{aligned}\mathcal{S}_{n+1} = \quad & conv\{(i_1,...,i_n),\ (i_1,i_2+1,i_3,i_4,...,i_n),(i_1+1,i_2+1,i_3,i_4,...,i_n),\\ & (i_1+1,i_2+1,i_3+1,i_4,...,i_{n-1},i_n),\ \ \ldots\ ,\ (i_1+1,...,i_n+1)\},\\ \mathcal{S}_{n+2} = \quad & conv\{(i_1,...,i_n),\ (i_1,i_2+1,i_3,...,i_n),(i_1,i_2+1,i_3+1,i_4,...,i_n),\\ & (i_1+1,i_2+1,i_3+1,i_4,...,i_{n-1},i_n),\ \ \ldots\ ,\ (i_1+1,...,i_n+1)\},\\ \vdots & \\ \mathcal{S}_{2n} = \quad & conv\{(i_1,...,i_n),\ (i_1,i_2+1,i_3,...,i_n),\ (i_1,i_2+1,i_3,...,i_{n-1},i_n+1),\\ & (i_1+1,i_2+1,i_3,...,i_{n-1},i_n+1),\ \ \ldots\ ,\ (i_1+1,...,i_n+1)\},\\ \vdots & \\ \mathcal{S}_{n!} = \quad & conv\{(i_1,...,i_n),\ (i_1,...,i_{n-1},i_n+1),(i_1,...,i_{n-2},i_{n-1}+1,i_n+1),\\ & (i_1+1,i_2,...,i_{n-2},i_{n-1}+1,i_n+1),\ \ \ldots\ ,\ (i_1+1,...,i_n+1)\}.\end{aligned}$$

Note that $|\mathcal{S}_1| = ... = |\mathcal{S}_{n!}| = n+1$ and simplices $\mathcal{S}_i$, $\mathcal{S}_j$, $i \neq j$ have a common facet if they have n common vertices. The sufficiency condition for a multivariate discrete probability function to be strongly unimodal is given by the use of any two neighboring simplices with one common facet.

Let p be the probability function of a discrete distribution on $\mathbb{Z}^n$ and $p(i_1,...,i_n)$ the value of p at $(i_1,...,i_n) \in \mathbb{Z}^n$. Let C denote the collection of simplices $\mathcal{S}_1,...,\mathcal{S}_{n!}$.

Theorem 9. *Suppose that p satisfies the following conditions:*

Condition I.

$$p(\mathbf{x})p(\mathbf{y}) \ \leq \ p(\mathbf{x}\vee\mathbf{y})p(\mathbf{x}\wedge\mathbf{y}) \tag{4.3.1}$$

where $\mathbf{x} = (i_1+\varepsilon_1,...,i_n+\varepsilon_n)$, $\mathbf{y} = (i_1+\delta_1,...,i_n+\delta_n)$ *and*

$$\mathbf{x}\vee\mathbf{y} = (\max(i_1+\varepsilon_1,i_1+\delta_1),...,\max(i_n+\varepsilon_n,i_n+\delta_n))$$

$$\mathbf{x}\wedge\mathbf{y} = (\min(i_1+\varepsilon_1,i_1+\delta_1),...,\min(i_n+\varepsilon_n,i_n+\delta_n))$$

and $\varepsilon_j,\delta_j \in \{0,1\}$, $j=1,...,n$ *defined such that for* $k=2,...,n$

$$\sum_{j=1}^{n}\varepsilon_j = \sum_{j=1}^{n}\delta_j = k-1\ , \quad \sum_{j=1}^{n}\varepsilon_j\delta_j = k-2\ .$$

<u>*Condition II.*</u>

$$p(i_1+\gamma_1+1,...,i_n+\gamma_n+1)p(i_1,...,i_n) \ \leq \ p(i_1+1,...,i_n+1)p(i_1+\gamma_1,...,i_n+\gamma_n) \ , \quad (4.3.2)$$

where $\gamma_j \in \{0,1\}, \ j=1,...,n$ *and* $\sum_{j=1}^{n} \gamma_j = 1.$

<u>*Condition III.*</u>

$$p(i_1-\alpha_1,...,i_n-\alpha_n)p(i_1+1,...,i_n+1) \ \leq \ p(i_1,...,i_n)p(i_1-\alpha_1+1,...,i_n-\alpha_n+1), \quad (4.3.3)$$

where $\alpha_j \in \{0,1\}, \ j=1,...,n$ *and* $\sum_{j=1}^{n} \alpha_j = 1.$

Then p *is strongly unimodal.*

Proof. Let $L(c,i_1,i_2,...,i_n),(i_1,i_2,...,i_n) \in \mathbb{Z}^n, \ c=1,2,...,n!$ denote the linear function on $\mathbb{R}^n$ which coincides on the vertices of C with $-\log p(.)$ and define

$$f(\mathbf{y}) = \begin{cases} L(c,i_1,i_2,...,i_n) & \text{if} \quad \mathbf{y} \in \mathcal{S}_c(i_1,i_2,...,i_n),(i_1,i_2,...,i_n) \in \mathbb{Z}^n \ , \\ \infty & \text{if} \quad \mathbf{y} \notin C \ . \end{cases}$$

It is easy to see that f coincides on $\mathbb{Z}^n$ with $-\log p(.)$.

Claim: Condition (4.3.1) ensures the convexity of f within a hypercube and conditions (4.3.2), (4.3.3) in different hypercubes.

Proof of the Claim: On each simplex we define a linear function as the equation of the hyperplane determined by the vertices of the simplex and the corresponding values of $-\log p(.)$. The collection of these linear pieces form the function f. We also ensure convexity of the function f on any two neighboring simplices with a common facet, i.e., the following is satisfied: for any $\mathbf{z}_i = (z_{i1},...,z_{in}) \ , i=0,1,...,n$ and $\mathbf{y}=(y_1,...,y_n)$

$$\frac{\begin{vmatrix} f(\mathbf{y}) & f(\mathbf{z}_0) & f(\mathbf{z}_1) & f(\mathbf{z}_2) & \dots & f(\mathbf{z}_n) \\ 1 & 1 & 1 & 1 & \dots & 1 \\ y_1 & z_{01} & z_{11} & z_{21} & \dots & z_{n1} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ y_n & z_{0n} & z_{1n} & z_{2n} & \dots & z_{nn} \end{vmatrix}}{\begin{vmatrix} 1 & 1 & 1 & \dots & 1 \\ z_{01} & z_{11} & z_{21} & \dots & z_{n1} \\ \vdots & \vdots & \vdots & & \vdots \\ z_{0n} & z_{1n} & z_{2n} & \dots & z_{nn} \end{vmatrix}} \geq 0 \quad (4.3.4)$$

where $f = -\log p(.)$.

First note that both $(i_1, ..., i_n)$ and $(i_1+1, ..., i_n+1)$ are the vertices of simplex $\mathcal{S}_j$ for any j. Therefore any $\mathcal{S}_j$ has $n+1$ neighbors: $n-1$ of them are in the same hypercube and two in different hypercubes. For the sake of simplicity let $(i_1, ..., i_n) = (0, ..., 0)$ and consider the simplex $\mathcal{S}_1$:

$$\boxed{conv\{(0,...,0),(1,0,0,...,0),(1,1,0,0,...,0),(1,1,1,0,0,...,0),\ldots,(1,...,1,1,0),(1,...,1)\}}$$

whose neighbors in the same hypercube are:

$$\begin{array}{|l|}\hline
conv\{(0,...,0),(0,1,0,...,0),(1,1,0,0,...,0),(1,1,1,0,0,...,0),\ldots,(1,...,1,1,0),(1,...,1)\},\\
conv\{(0,...,0),(1,0,0,...,0),(1,0,1,0,...,0),(1,1,1,0,0,...,0),\ldots,(1,...,1,1,0),(1,...,1)\},\\
conv\{(0,...,0),(1,0,0,...,0),(1,1,0,0,...,0),(1,1,0,1,0,...,0),\ldots,(1,...,1,1,0),(1,...,1)\},\\
\quad\vdots\\
conv\{(0,...,0),(1,0,0,...,0),(1,1,0,0,...,0),(1,1,1,0,0,...,0),\ldots,(1,...,1,0,1),(1,...,1)\},\\
\hline\end{array}$$

and those in two different hypercubes are:

$$\begin{array}{|l|}\hline
conv\{(2,1,...,1),(1,0,...,0),(1,1,0,0,...,0),(1,1,1,0,...,0),\ldots,(1,...,1,1,0),(1,...,1)\},\\
conv\{(0,...,0),(1,0,...,0),(1,1,0,0,...,0),(1,1,1,0,...,0),\ldots,(1,...,1,1,0),(0,...,0,-1)\}.\\
\hline\end{array}$$

Here the first $n-1$ neighbors of $\mathcal{S}_1$ are listed in such a way that in the first neighboring simplex of $\mathcal{S}_1$ only the second vertex, in the second neighboring simplex only the third vertex, ..., and finally in the $(n-1)$st neighboring simplex only the nth vertex is different than the one in $\mathcal{S}_1$. In the nth neighboring simplex the first and in the $(n+1)$st neighboring simplex the last vertex is different than the one in $\mathcal{S}_1$.

Let $\mathbf{z}_i, i = 0, ..., n$ be the vertices of $\mathcal{S}_1$ in the same order given above and $\mathbf{y}$ the vertex of the neighboring simplex, that does not belong to $\mathcal{S}_1$. Note that exactly k components of vertex $\mathbf{z}_k$ are 1. We will alternatively use the notation f_k instead of $f(\mathbf{z}_k)$.

Since the denominator of (4.3.4) is 1, the convexity of f is ensured if its numerator is nonnegative. We consider the following cases:

Case 1. Let $\mathbf{y} = (0, 1, 0, ..., 0)$. The numerator of (4.3.4) is equivalent to:

$$\begin{vmatrix} f(\mathbf{y}) & f_0 & f_1 & f_2 & \dots & f_n \\ 1 & 1 & 1 & 1 & \dots & 1 \\ 0 & 0 & 1 & 1 & \dots & 1 \\ 1 & 0 & 0 & 1 & \dots & 1 \\ 0 & 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & 0 & \dots & 1 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \end{vmatrix} . \tag{4.3.5}$$

If we subtract $(n+1)$st column from the $(n+2)$nd column, nth column from the $(n+1)$st column, ..., the second column from the third and finally from the first one and develop the determinant according to the second row, we obtain

$$\begin{vmatrix} f(\mathbf{y}) - f_0 & f_1 - f_0 & f_2 - f_1 & f_3 - f_2 & \dots & f_n - f_{n-1} \\ 0 & 1 & 0 & 0 & \dots & 0 \\ 1 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \end{vmatrix} . \tag{4.3.6}$$

It follows that

$$f(\mathbf{y}) + f_1 \geq f_0 + f_2 \ . \tag{4.3.7}$$

Since $f = -logp(.)$, we have

$$p(0, 1, 0, ..., 0)p(1, 0, 0, ..., 0) \leq p(0, 0, ..., 0)p(1, 1, 0, ..., 0) \ .$$

Case 2. Let $\mathbf{y} = (1, 0, 1, ..., 0)$. The numerator of (4.3.4) is:

$$\begin{vmatrix} f(\mathbf{y}) & f_0 & f_1 & f_2 & \dots & f_n \\ 1 & 1 & 1 & 1 & \dots & 1 \\ 1 & 0 & 1 & 1 & \dots & 1 \\ 0 & 0 & 0 & 1 & \dots & 1 \\ 1 & 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & 0 & \dots & 1 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \end{vmatrix} . \tag{4.3.8}$$

We apply a column subtraction method similar to the one used in Case 1. We subtract $(n+1)$st column from the $(n+2)$nd column, nth column from the $(n+1)$st column, ..., the second column from the third and finally from the first one. Next we develop the determinant according to the second row and subtract the second column from the first column of the resulting determinant. In this case (4.3.8) is equivalent to the following determinant:

$$\begin{vmatrix} f(\mathbf{y}) - f_1 & f_1 - f_0 & f_2 - f_1 & f_3 - f_2 & \dots & f_n - f_{n-1} \\ 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ 1 & 0 & 0 & 1 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \end{vmatrix} \tag{4.3.9}$$

which provides us with the result

$$f(\mathbf{y}) + f_2 \geq f_1 + f_3 \ . \tag{4.3.10}$$

This is the same as

$$p(1, 0, 1, 0, ..., 0)p(1, 1, 0, ..., 0) \leq p(1, 0, ..., 0)p(1, 1, 1, 0, ..., 0) \ .$$

If exactly k components of $\mathbf{y}$ are 1, then, by the use of the column subtraction method described above, the numerator of (4.3.4) can be obtained as

$$f(\mathbf{y}) + f_k \geq f_{k-1} + f_{k+1} \ . \tag{4.3.11}$$

Hence, we obtain all conditions of type (4.3.1) for the case of simplex $\mathcal{S}_1$.

Case 3. Now, let $\mathbf{y} = (2, 1, ..., 1)$, which is the vertex of the nth neighbor of the simplex $\mathcal{S}_1$ in different hypercube. In this case we use the following column subtraction method: subtract the $(n+2)$nd column from the first one, the $(n+1)$st column from the $(n+2)$nd one, the nth column from the $(n+1)$st one, ..., and finally the second column from the third one. The resulting determinant is equivalent to:

$$\begin{vmatrix} f(\mathbf{y}) - f_n & f_1 - f_0 \\ 1 & 1 \end{vmatrix} \ . \tag{4.3.12}$$

In this case we obtain

$$f(\mathbf{y}) + f_0 \geq f_1 + f_n \ , \tag{4.3.13}$$

which is the same as

$$p(2, 1, ..., 1)p(0, 0, ..., 0) \leq p(1, 0, ..., 0)p(1, ..., 1) \ .$$

Case 4. Let $\mathbf{y} = (0, ..., 0, -1)$. If we use the column subtraction method in Case 1, the numerator of (4.3.4) is obtained as follows:

$$\begin{vmatrix} f(\mathbf{y}) - f_0 & f_n - f_{n-1} \\ -1 & 1 \end{vmatrix} \ . \tag{4.3.14}$$

Hence, we have

$$f(\mathbf{y}) + f_n \geq f_0 + f_{n-1} \ , \tag{4.3.15}$$

that is equivalent to

$$p(0, ..., 0, -1)p(1, ..., 1, 1) \leq p(0, ..., 0, 0)p(1, ..., 1, 0)$$

Thus, we obtain all conditions of type (4.3.2) and (4.3.3) for the case of simplex $\mathcal{S}_1$.

Let $\mathbf{z}_k$, $k = 0, ..., n$ designate the vertices of a simplex $\mathcal{S}_j$ for any j, whose k components are equal to 1 and $\mathbf{y}$ the vertex of the neighboring simplex that does not belong to $\mathcal{S}_j$. In case of a simplex $\mathcal{S}_j$, the determinant in denominator of (4.3.4) is either 1 or -1. Moreover, by the use of elementary row operations, the numerator of (4.3.4) can be transformed into one of the determinants in Cases 1-4. Hence, the convexity conditions for f are of the forms (4.3.11), (4.3.13), (4.3.15) which provide us with conditions (4.3.1), (4.3.2) and (4.3.3). Thus, p is strongly unimodal. $\square$

4.4 Maximization of a Strongly Unimodal Multivariate Discrete Distribution

We want to find the maximum of a strongly unimodal probability function $p(\mathbf{x}),\ \mathbf{x} \in \mathbb{Z}^n$. The problem is the same as the minimum value of the function

$$-\log p(\mathbf{x}) \ \text{ if } \ \mathbf{x} \in \mathbb{Z}^n$$

that we will be looking at.

In addition to the strongly unimodality of $p(\mathbf{x}),\ \mathbf{x} \in \mathbb{Z}^n$ we assume that there exists a subdivision of $\mathbb{R}^n$ into simplices with pairwise disjoint interiors such that all vertices of all simplices are elements of $\mathbb{Z}^n$ and a function $f(\mathbf{x}),\ \mathbf{x} \in \mathbb{R}^n$ that is linear on each of them, otherwise convex in $\mathbb{R}^n$ and $f(\mathbf{x}) = -\log p(\mathbf{x})$ if $\mathbf{x} \in \mathbb{Z}^n$.

Probability functions frequently take zero values on some points of $\mathbb{Z}^n$. If $p(\mathbf{x}) = 0$, then by definition $f(\mathbf{x}) = \infty$ and therefore any $\mathbf{x}$ with this property can be excluded from the optimization. We can also restrict the optimization to bounded sets. In fact, since $\sum_{\mathbf{x} \in \mathbb{Z}^n} p(\mathbf{x}) = 1$, it follows that there exists a vector $\mathbf{b}$ such that the minimum of f is taken in the set $\{\mathbf{x} \mid |\mathbf{x}| \leq \mathbf{b}\}$. Such a $\mathbf{b}$ can easily be found without the knowledge of the minimum of f, we simply take a $\mathbf{b}$ with large enough components. For simplicity we assume that the minimum of f is taken at some point of the set $\{x \mid 0 \leq \mathbf{x} \leq \mathbf{b}\}$ and such a $\mathbf{b}$ is known.

Probability functions sometimes are of the type where the nonzero probabilities fill up the lattice points of a simplex. An example is the multinomial distribution where

$p(x_1, ..., x_n) > 0$ for the elements of the set

$$\{\mathbf{x} \in \mathbb{Z}^n \mid \mathbf{x} \geq 0,\ x_1 + ... + x_n \leq n\}.$$

Taking this into account, we will be looking at the problems

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ \text{subject to} \quad & \\ & 0 \leq \mathbf{x} \leq \mathbf{b}, \quad \mathbf{x} \ \text{ integer} \end{aligned} \tag{4.4.1}$$

where $\mathbf{b}$ has positive integer components and

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ \text{subject to} \quad & \\ & x_1 + ... + x_n \leq b \\ & \mathbf{x} \geq 0, \quad \mathbf{x} \ \text{ integer} \end{aligned} \tag{4.4.2}$$

where b is a positive integer.

Our assumption regarding the subdivision of $\mathbb{R}^n$ into simplices and the function $f(\mathbf{x}),\ \ \mathbf{x} \in \mathbb{R}^n$ carry over to the feasible sets in problems (4.4.1) and (4.4.2) in a natural way.

The integrality restriction of $\mathbf{x}$ can be removed from both problems (4.4.1) and (4.4.2). In fact, our algorithm not only produces an optimal $\mathbf{x}$ but also a subdividing simplex of which $\mathbf{x}$ is an element and at least one of the vertices of the resulting simplex also is an optimal solution.

First we consider problem (4.4.1). Let N be the number of lattice points of the set $\{\mathbf{x} \mid 0 \leq \mathbf{x} \leq \mathbf{b}\}$. Any function value $f(\mathbf{x}),\ 0 \leq \mathbf{x} \leq \mathbf{b}$ can be obtained by λ-representation, as the optimum value of a linear programming problem:

$$
\begin{aligned}
f(\mathbf{x}) \quad &= \min_{\lambda} \ \sum_{k=1}^{N} f(\mathbf{z}_k)\lambda_k \\
\text{subject to} \quad & \\
&\sum_{k=1}^{N} \mathbf{z}_k \lambda_k = \mathbf{x} \\
&\sum_{k=1}^{N} \lambda_k = 1 \\
&0 \le \mathbf{x} \le \mathbf{b} \\
&\lambda \ge 0 \ .
\end{aligned}
\tag{4.4.3}
$$

By the use of problem (4.4.3), problem (4.4.1) can be written in the following way:

$$
\begin{aligned}
\min_{\mathbf{x}} f(\mathbf{x}) \quad &= \min_{\lambda, \mathbf{x}} \ \sum_{k=1}^{N} f(\mathbf{z}_k)\lambda_k \\
\text{subject to} \quad & \\
&\sum_{k=1}^{N} \mathbf{z}_k \lambda_k = \mathbf{x} \\
&\sum_{k=1}^{N} \lambda_k = 1 \\
&\mathbf{x} \le \mathbf{b} \\
&\lambda \ge 0 \ , \quad \mathbf{x} \ge 0.
\end{aligned}
\tag{4.4.4}
$$

Introducing slack variables we rewrite the problem as

$$
\begin{aligned}
\min_{\mathbf{x}} f(\mathbf{x}) \quad &= \min_{\lambda, \mathbf{x}} \ \sum_{k=1}^{N} f(\mathbf{z}_k)\lambda_k \\
\text{subject to} \quad & \\
&\sum_{k=1}^{N} \mathbf{z}_k \lambda_k - \mathbf{x} = 0 \\
&\sum_{k=1}^{N} \lambda_k = 1 \\
&\mathbf{x} + \mathbf{u} = \mathbf{b} \\
&\lambda \ge 0 \ , \quad \mathbf{x} \ge 0 \ , \quad \mathbf{u} \ge 0.
\end{aligned}
\tag{4.4.5}
$$

In order to construct an initial dual feasible basis to problem (4.4.5) we use the following theorem by Prékopa and Li [75] (Theorem 2.1).

Theorem 10. *Suppose that* $\mathbf{z}_1, ..., \mathbf{z}_k$ *are elements of a convex polyhedron* K *that is subdivided into* r*-dimensional simplices* $S_1, ..., S_h$ *with pairwise disjoint interiors and the set of all of their vertices is equal to* $\{\mathbf{z}_1, ..., \mathbf{z}_k\}$. *Suppose further that there exists a convex function* $f(\mathbf{z}), \mathbf{z} \in K$, *continuous on* K *and linear on any of the simplices* $S_1, ..., S_h$ *with different normal vectors on different simplieces such that* $f_i = f(\mathbf{z}_i), i = 1, ..., k$. *Let* $B_1, ..., B_h$ *be those* $(n+1) \times (n+1)$ *parts of the matrix of equality constraints of problem (4.4.3), the upper* $n \times (n+1)$ *parts of which are the sets of vertices of the simplices* $S_1, ..., S_h$, *respectively. Then* $B_1, ..., B_h$ *are the dual feasible bases of problem (4.4.3) and each of them is dual nondegenerate*

If the above-mentioned normal vectors are not all different, the assertion that the vertices of any simplex form a dual feasible basis, remains true but these bases are no longer all dual nondegenerate, as it turns out from the proof of the theorem.

Let S_1, ..., $S_{n!}$ designate the subdividing simplices. Let us rewrite problem (4.4.5) into more detailed form:

$$\min \sum_{k=1}^{N} f(\mathbf{z}_k)\lambda_k$$

subject to

$$x_1 + u_1 = b_1$$
$$\vdots$$
$$x_n + u_n = b_n$$

$$\sum_{k=1}^{N} \begin{pmatrix} z_{k1} \\ \vdots \\ z_{kn} \end{pmatrix} \lambda_k - \begin{pmatrix} x_1 \\ \vdots \\ 0 \end{pmatrix} - ... - \begin{pmatrix} 0 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \tag{4.4.6}$$

$$\sum_{k=1}^{N} \lambda_k = 1$$

$$\lambda \geq 0\,, \quad \mathbf{x} \geq 0\,, \quad \mathbf{u} \geq 0.$$

0	$\cdots$	0	0	$\cdots$	0	$f(\mathbf{z}_1)$	$\cdots$	$f(\mathbf{z}_n)$	
1	$\cdots$	0	1	$\cdots$	0	0	$\cdots$	0	b_1
$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
0	$\cdots$	1	0	$\cdots$	1	0	$\cdots$	0	b_n
-1	$\cdots$	0	0	$\cdots$	0	z_{11}	$\cdots$	z_{N1}	0
$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
0	$\cdots$	-1	0	$\cdots$	0	z_{1n}	$\cdots$	z_{Nn}	0
0	$\cdots$	0	0	$\cdots$	0	1	$\cdots$	1	1
Block 0						Block 1			

Table 4.1: Coefficient matrix of problem (4.4.6), together with the objective function coefficients and the right-hand side vector

It is easy to see that the rank of the matrix of equality constraints in (4.4.6) is $2n+1$. Let $v_1,\ v_2,\ \ldots,\ v_n,\ y_1,\ y_2,\ \ldots,\ y_n,\ w$ designate the dual variables. The coefficient matrix of problem (4.4.6) has a special structure illustrated in Table 4.1.

First, let us introduce the notations:

$$A = \begin{pmatrix} 1 & \ldots & 0 & 1 & \ldots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 1 & 0 & \ldots & 1 \end{pmatrix}, \quad T = \begin{pmatrix} -1 & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & -1 & 0 & \ldots & 0 \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

To initiate the dual algorithm we compose a dual feasible basis for problem (4.4.6). To accomplish this job we pick an arbitrary simplex S_i whose vertices are $\mathbf{z}_{i_1},\ \ldots,\ \mathbf{z}_{i_{n+1}}$ and form the $2n+1$-component vectors

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \\ \mathbf{z}_{i_1} \\ 1 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \mathbf{z}_{i_{n+1}} \\ 1 \end{pmatrix}. \tag{4.4.7}$$

Then we compute the dual variables $\mathbf{y}$ and w by using the equation:

$$\mathbf{z}_{i_k}^T \mathbf{y} + w = f(\mathbf{z}_{i_k}), \; k = 1, ..., n+1 \; .$$

Next we solve the linear programming problem

$$\begin{aligned} &\min \quad -\mathbf{y}^T T \begin{pmatrix} \mathbf{x} \\ \mathbf{u} \end{pmatrix} \\ &\text{subject to} \\ &\qquad x_1 + u_1 = b_1 \\ &\qquad\quad \vdots \\ &\qquad x_n + u_n = b_n \\ &\qquad \mathbf{x} \geq 0 \; , \quad \mathbf{u} \geq 0, \end{aligned} \tag{4.4.8}$$

by a method that produces a primal-dual feasible basis. Let B be this optimal basis and d a dual vector corresponding to B, i.e., any solution of the equation $d^T B = -\mathbf{y}^T T_B$, where T_B is the part of T which correspond to the basis subscripts. Since A has full rank B and d is uniquely determined.

Problem (4.4.9), however, is equivalent to

$$\begin{aligned} &\min \; y_1 x_1 + ... + y_n x_n \\ &\text{subject to} \\ &\qquad x_1 + u_1 = b_1 \\ &\qquad\quad \vdots \\ &\qquad x_n + u_n = b_n \\ &\qquad \mathbf{x} \geq 0 \; , \quad \mathbf{u} \geq 0 \end{aligned} \tag{4.4.9}$$

which can be solved easily: if $y_i \leq 0$, we take the column of $\mathbf{x}_i$ otherwise we take the column of $\mathbf{u}_i$ into the basis. We have obtained a dual feasible basis for problem (4.4.6). It consists of those vectors that trace out B from A and T_B from T in Block 0, furthermore the previously selected vectors (4.4.7) in Block 1. The dual feasibility is guaranteed by the theorem of Prékopa and Li in [75] (Theorem 2.2).

The next step is to check the primal feasibility of basis. The first n constraint in problem (4.4.6) ensure that in case of any basis the basic $\mathbf{x}_i, \; \mathbf{u}_j, \; (i \neq j)$ variables are positive, since we have the inequality $\mathbf{b} > 0$. Thus, if the basis is not primal feasible, then only the λ variables can take negative value.

If $\lambda_j < 0$, then the column of $\mathbf{z}_{i_j}$ can be chosen to be the outgoing vector. The incoming vector is either a nonbasic column from Block 0 or a nonbasic column from Block 1. The algorithm can be described in the following way.

Dual algorithm to maximize strongly unimodal functions

Let us introduce the notations:

$$H = \{1, 2, ..., h\}, \quad K = \{1, 2, ..., k\},$$

$$Z_j = \{\mathbf{z}_k \mid \mathbf{z}_k \text{ is a vertex of the simplex } \mathcal{S}_j \ , \ k \in K\}, \quad j \in H,$$

$$L_j = \{k \mid \mathbf{z}_k \in Z_j\}, \quad j \in H.$$

Step 0. Pick arbitrarily a simplex $\mathcal{S}_j$ and let $\mathbf{z}_i, \ i \in \{i_1, ..., i_{n+1}\}$ be the collection of its vertices and

$$I^{(0)} \leftarrow L_j = \{i_1, ..., i_{n+1}\}$$

Go to Step 1.

Step 1. Solve for $\mathbf{y}, \ w$ the system of linear equations:

$$\mathbf{z}_{i_k}^T \mathbf{y} + w = f(\mathbf{z}_{i_k}), \quad k \in I^{(0)},$$

where $\mathbf{y} \in \mathbb{R}^n, \ w \in \mathbb{R}$ Go to Step 2.

Step 2. Compose a dual feasible basis B by including the vectors (4.4.7), any column of $\mathbf{x}_i$ if $y_i \leq 0$ and any column of $\mathbf{u}_i$ if $y_i > 0$. Go to Step 3.

Step 3. Check the primal feasibility of basis B. If $\lambda \geq 0$, stop, the basis is optimal. If $\lambda \ngeq 0$, then pick $\lambda_q < 0$ arbitrarily and remove $\mathbf{z}_q$ from the basis. Go to Step 4.

Step 4. *Determination of incoming vector.* The following columns may enter:

(1) A nonbasic column from Block 0.

(2) A nonbasic column $\mathbf{z}_j$ from Block 1.

Updating formulas

(1) Given $I^{(k)}$, in order to update a column from Block 0 which traces out the nonbasic column a_p from A, solve the following system of linear equations:

$$\begin{aligned} A_B d_p &= a_p \\ T_B d_p + \sum_{i \in I^{(k)}} \mathbf{z}_i d_i &= t_p \\ \sum_{i \in I^{(k)}} d_i &= 1 \ , \end{aligned} \tag{4.4.10}$$

where A_B, T_B are the parts of A and T, respectively, corresponding to basis B; d_p is a vector with suitable size and t_p is the p column of the matrix T.

Compute the reduced costs:

$$\bar{c}_p = \sum_{i \in I^{(k)}} f(\mathbf{z}_i) d_i \ .$$

(2) Assume that a nonbasic column $\mathbf{z}_j$ from Block 1, where $|I^{(k)}| < n+1$, $j \neq q$ and $\{j\} \cup I^{(k)} \backslash \{q\}$ is a subset of L_l for some $l \in H$, enters the basis. Let $\hat{I}^{(k)}$ designate the set of all possible j from Block 1 satisfying above requirements. To update the column containing $\mathbf{z}_j$, $j \in \hat{I}^{(k)}$, we solve the system of linear equations:

$$\begin{aligned} A_B r_j &= 0 \\ T_B r_j + \sum_{i \in \hat{I}^{(k)}} \mathbf{z}_i d_i &= \mathbf{z}_j \\ \sum_{i \in \hat{I}^{(k)}} d_i &= 1 \ , \end{aligned} \tag{4.4.11}$$

where r_j is a vector with suitable size.

Compute the reduced costs:

$$\bar{f}_j = -f(\mathbf{z}_j) \ .$$

Determination of the vector that enters the basis

Let $\tilde{d}^{\,T} = (d_p^T, d_{i_1}, ..., d_{i_{n+1}})$ and $\tilde{d}(q)$ be the qth component of $\tilde{d}$ in (4.4.10). Let $\tilde{r}^{\,T} = (r_j^T, d_{i_1}, ..., d_{i_{n+1}})$ and $\tilde{r}(q)$ be the q component of $\tilde{r}$ in (4.4.11). Then the incoming vector is determined by taking the minimum of the following minima:

$$min_{\tilde{d}(q)<0}\left\{\frac{\sum_{i\in I^{(k)}} f(\mathbf{z}_i)d_i}{\tilde{d}(q)}\right\}\ , \tag{4.4.12}$$

$$min_{\tilde{r}(q)<0}\left\{\frac{-f(\mathbf{z}_j)}{\tilde{r}(q)}\right\}\ . \tag{4.4.13}$$

If the minimum is attained in (4.4.12), let

$$I^{(k+1)} = I^{(k)}\backslash\{q\}\ .$$

Update the basis B by replacing the outgoing vector by the column of a_p in Block 0.

If the minimum is attained in (4.4.13), then the column of $\mathbf{z}_j$ is the incoming vector. Let

$$I^{(k+1)} = I^{(k)} \cup \{j\}\backslash\{q\}\ .$$

Update the basis B by replacing the outgoing vector by the column of $\mathbf{z}_j$ in Block 1. Go to Step 3.

If no two linear pieces of the function $f(\mathbf{x})$ are on the same hyperplane, then cycling cannot occur, i.e., no simplex that has been used before returns. Otherwise an anti-cycling procedure has to be used: lexicographic dual algorithm (see, e.g., [76]) or Bland's rule [10].

We can also find bounds for the optimum value of problem (4.4.6) by the use of the fast bounding technique by Prékopa and Li [75]. First we construct a dual feasible basis as described before. If $\mathbf{v},\ \mathbf{y},\ w$ are the corresponding dual vectors, then $b^T\mathbf{v}+w$ is a lower bound for the optimum value of problem (4.4.6). In order to find an upper bound we use any method that produces a pair of primal and dual optimal solutions (not necessarily an optimal basis). Having the optimal $(\hat{\mathbf{x}}^T, \hat{\mathbf{u}}^T)$, we arbitrarily pick a simplex $\mathcal{S}_k$ and represent $(\hat{\mathbf{x}}^T, \hat{\mathbf{u}}^T)$ as the convex combination of the vertices of S_k. If all coefficients are nonnegative, then we stop. Otherwise we delete the corresponding vertex from the simplex and update the basis by including the vertex of the neighboring simplex into the basis which is not a vertex of the current simplex. If the representation of the vector $-T(\hat{\mathbf{x}}^T, \hat{\mathbf{u}}^T)$ is

$$-T(\hat{\mathbf{x}}^T, \hat{\mathbf{u}}^T) = \sum_{\mathbf{z}_j\in\mathcal{S}} \mathbf{z}_j\lambda_j,$$

where $\sum_j \lambda_j = 1$, $\lambda_j \geq 0$, then the upper bound is given by

$$-\mathbf{y}^T T(\hat{\mathbf{x}}^T, \hat{\mathbf{u}}^T) + \sum_j f(\mathbf{z}_j)\lambda_j \ .$$

The solution of problem (4.4.2) can be accomplished in the same way, only trivial modifications are needed. If the minimum of f is taken in the set $\{\mathbf{x} \in \mathbb{Z}^n \mid \mathbf{x} > 0,\ x_1 + ... + x_n \leq b\}$, then we assume that $f(\mathbf{x}) = M$ $(M > 0)$ for every $\mathbf{x}$ that does not belong to this set, where M is large enough (or ∞). In this case, problem (4.4.2) can be solved by the use of above-mentioned methods.

In continuous optimization one of the important properties of convex functions is the coincidence between their local and global minima. A function is $g : \mathbb{Z}^n \to \mathbb{R}$ called integrally convex if and only if its extension $\tilde{g} : \mathbb{R}^n \to \mathbb{R}$ is convex. In this case a global minimum for (continuous) function $\tilde{g}$ is a global minimum for (discrete) function g, and vice versa.

4.5 Examples

We present four multivariate discrete distributions that are strongly unimodal. The properties of the distributions in the examples of this section can be found in [39].

Example 1. A function $f : X = X_1 \times X_2 \times ... \times X_n \to [0, \infty)$ is said to be *multivariate totally positive of order* 2, *MTP*$_2$, if for all $\mathbf{x},\ \mathbf{y} \in X$ (see [41])

$$f(\mathbf{x} \vee \mathbf{y}) f(\mathbf{x} \wedge \mathbf{y}) \geq f(\mathbf{x}) f(\mathbf{y}),$$

where

$$\mathbf{x} \vee \mathbf{y} = (\max(x_1, y_1), \max(x_2, y_2), ... \max(x_n, y_n)),$$

$$\mathbf{x} \wedge \mathbf{y} = (\min(x_1, y_1), \min(x_2, y_2), ... \min(x_n, y_n)) \ .$$

It is easy to see that if p is *MTP*$_2$, then condition (4.3.1) of Theorem 2 is satisfied.

Consider the negative multinomial distribution with the probability function:

$$p(x_1, ..., x_n) = \frac{(k - 1 + \sum_{i=1}^n x_i)!}{(k-1)!} p_0^k \prod_{i=1}^n \frac{p_i^{x_i}}{x_i!}$$

$$x_i = 0, 1, 2, ... \quad i = 1, 2, ..., n$$

$$\sum_{i=0}^{n} p_i = 1, \ 0 < p_i < 1, \ i = 1, 2, .., n.$$

Karlin and Rinott [41] proved that the negative multinomial distribution is MTP_2. Therefore p satisfies condition (4.3.1) of Theorem 2.

Conditions (4.3.2) and (4.3.3) are satisfied if

$$\begin{gathered} n - 1 - 2x_1 + x_2 + ... + x_n \geq 0, \\ n - 1 + x_1 - 2x_2 + x_3 + ... + x_n \geq 0, \\ \vdots \\ n - 1 + x_1 + x_2 + ... + x_{n-1} - 2x_n \geq 0 \ . \end{gathered} \tag{4.5.1}$$

Thus, the negative multinomial distribution is strongly unimodal if (4.5.1) holds.

Example 2. The multivariate hypergeometric distribution has the following probability function:

$$\frac{p(x_1, ..., x_{n-1}) = \prod_{i=1}^{n-1} \begin{pmatrix} m_i \\ x_i \end{pmatrix} \begin{pmatrix} m - m_1 - ... - m_{n-1} \\ k - x_1 - ... - x_{n-1} \end{pmatrix}}{\begin{pmatrix} m \\ k \end{pmatrix}}$$

$$0 \leq x_i \leq m_i, \quad i = 1, 2, ..., n-1$$

$$\sum_{i=1}^{n-1} x_i \leq k, \quad \sum_{i=1}^{n-1} m_i \leq m \ .$$

One can show that p satisfies the conditions of Theorem 2. Thus, the multivariate hypergeometric distribution is strongly unimodal.

Example 3. The multivariate negative hypergeometric distribution has probability function:

$$p(\mathbf{x}) = \frac{k!\Gamma(m)\Gamma(m - m_1 - ... - m_{n-1} + k - x_1 - ... - x_{n-1})}{\Gamma(k+m)\Gamma(m - m_1 - ... - m_{n-1})(k - x_1 - ... - x_{n-1})!} \prod_{i=1}^{n-1} \frac{\Gamma(m_i + x_i)}{\Gamma(m_i) x_i!}$$

$$0 \leq x_i \leq m_i, \quad i = 1, 2, ..., n-1$$

$$\sum_{i=1}^{n-1} x_i \leq k, \quad \sum_{i=1}^{n-1} m_i \leq m \ .$$

Since p satisfies the conditions of Theorem 2 it is strongly unimodal.

Example 4. Consider the Dirichlet (or Beta)-compound multinomial distribution

$$\text{Multinomial}(k; p_1, ..., p_{n-1}) \bigwedge_{p_1,...,p_{n-1}} \text{Dirichlet}(\alpha_1, ..., \alpha_{n-1})$$

The probability mass function of this compound distribution is:

$$p(x_1, ..., x_{n-1}) = \frac{k!\Gamma(\alpha)\Gamma(\alpha_n + k - x_1 - ... - x_{n-1})}{\Gamma(k+\alpha)\Gamma(\alpha_n)(k - x_1 - ... - x_{n-1})!} \prod_{i=1}^{n-1} \frac{\Gamma(\alpha_i + x_i)}{\Gamma(\alpha_i) x_i!} \ .$$

$$\alpha_n = \alpha - \sum_{i=1}^{n-1} \alpha_i, \quad \sum_{i=1}^{n-1} x_i \leq k, \quad x_i \geq 0 \ .$$

The function p satisfies conditions (4.3.1), (4.3.2) and (4.3.3) of Theorem 2. Thus, it is strongly unimodal.

References

[1] F. Alajaji, H. Kuai, G. Takahara, A Lower Bound for the Probability of a Finite Union of Events Discrete Applied Mathematics 215 (2000), 147–158.

[2] K. J. Arrow, S. Karlin, H. Scarf, A min-max solution of an inventory problem, Studies in the Mathematical Theory of Inventory and Production, Standford University Press, 1958.

[3] J. L. Balintfy, R. D. Armstong, A chance constrained multiple choice programming algorithm with applications in: Stochastic Programming (Proc. Internat. Conf., Univ. Oxford, Oxford, England, 1974, M.A.H. Dempster), Academic Press, London, New York, pp. 301–325, 1980.

[4] R. Barlow, F. Proschan, Statistical theory of reliability and life testing probability models, Holt, Rinehartand Winston Inc., 1975.

[5] O. Barndorff-Nielsen, Unimodality and exponential families, Comm. Statist. 1 (1973), 189–216.

[6] J. Battersby, Network Analysis for Planning and Scheduling (3rd edn.), Macmillan, 1970.

[7] C. M. Bender, D.C. Brody, B.K. Meister, Inverse of a Vandermonde Matrix, 2001.

[8] P. Beraldi, A. Ruszczyński, A branch and bound method for stochastic integer problems under probabilistic constraints, Optimization Methods and Software 17 (2002), 359–382.

[9] A. Bjorck, V. Pereyra, Solutions of Vandermonde systems of equations, Mathematics of Computations 24 (1970), 893–903.

[10] R. G. Bland, New pivoting rules for the simplex method, Math. of Oper. Res. 2 (1977), 103–107.

[11] C. E. Bonferroni, Teoria Statistica delle classi e calcoro delle probabilitá, Volume in onore di Riccardo Dalla Volta, Universitá di Frenze 2 (1937), 1–62.

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

[13] G. Boole, Of Propositions Numericallu Definite, In Transactions of the Cambridge Philosophical Society, Part II, XI, reprint as Study IV in the next reference, 1868.

[14] G. Boole, Collected Logical Works, Vol I, Studies in Logic and Probability, R. Rhees, (Ed.), Open Court Publ. Co., LaSalle, I11, 1952.

[15] E. Boros, A. Prékopa, Closed form two-sided bounds for probabilities that exactly r and at least r out of n events occur, Math. Oper. Res. 14 (1989), 317–347.

[16] E. Boros, A. Prékopa, Probabilistic Bounds and Algorithms for the Maximum Satisfiability Problem, Annals of Operations Research 21 (1989), 109–126.

[17] J. Bukszár, A. Prékopa, Probability bounds with cherry trees, Math. Oper. Res. 26 (2001), 174 - 192.

[18] J. Bukszár, T. Szántai, Probability bounds given by hypercherry trees, Optimization Methods and Software 17 (3) (2002), 409–422.

[19] J. Bukszár, Hypermultitrees and Bonferroni Inequalities, Math. Ineq. and Applications 6 (2003), 727–745.

[20] A. Burkauskas, On the Convexity Problem of Probabilistic Constrained Stochastic Programming Problems (in Hungarian). Alkalmazott Matematikai Lapok (Applied Mathematical Papers) 12 (1986), 77–90.

[21] A. Charnes, W. W. Cooper, G. H. Symonds, Cost horizons and certainty equivalents; an approach to stochastic programming of heating oil, Management Science 4 (1958), 235–263.

[22] D. A. Dawson, A. Sankoff, An inequality for probabilities, Proceedings of the American Mathematical Society 18 (1967), 504–507.

[23] D. Dentcheva, A. Prékopa, A. Ruszczyński, Concavity and efficient points of discrete distribution in probabilistic programming, Math. Program. A, 89 (2000), 55–77.

[24] D. Dentcheva, A. Prékopa, A. Ruszczyński, On Convex Probabilistic Programming with Discrete Distributions. Nonlinear Analysis 47 (2001), 1997–2009.

[25] D. Dentcheva, A. Prékopa, A. Ruszczyński, Bounds for integer stochastic programs with probabilistic constraints, Discrete Applied Mathematics 124 (2002), 55–65.

[26] D. Dentcheva, B. Lai, A. Ruszczyński, Efficient point methods for probabilistic optimization problems, Mathematical Methods of Operations Research 60 (2004), 331–346.

[27] C. I. Fábián, A. Prékopa, O. Ruf-Fiedler, On a dual method for a speacially structured linear programming problem with application to stochastic programming, Optimization Methods and Software 17 (2002), 445–492.

[28] P. Favati, F. Tardella, Convexity in nonlinear integer programming, Ricerca Operativa 53 (1990) 3–44.

[29] M. Fekete, G. Pólya, Überein ein Problem von Laguerre, Rediconti del Circolo Matematico di Palermo 23 (1912), 89–120.

[30] M. Fréchet, Les Probabilités Associées a un Systéme d'Événement Compatibles et Dépendants, Actualités Scientifique et Industrielles, Nos. 859, 942 (1940/43), Paris.

[31] J. Galambos, R. Mucci, Inequalities for linear combinations of binomial moments, Publ. Math. 27 (1980), 263–269.

[32] J. Galambos, I. Simonelli, Bonferroni-type inequalities with applications, Springer, Wahrscheinlichkeits, 1996.

[33] L. Gao, A. Prékopa, Lower and Upper Bounds for the Probability that at least r and Exactly r out of n Events Occur, Mathematical Inequalities and Applications 5 (2002), 315–333.

[34] J. Gessel and G. Viennot, *Binomial determinants, paths, and hook length formulae*, Advences in Mathematics, **58**, (1985), 300–321.

[35] T. Hailperin, Best Possible Inequalities for the probability of a Logical Function of Events, Am. Math. Monthly 72 (1965), 343–359.

[36] R. Henrion, C. Strugarek, Convexity of Chance Constraints with Independent Random Variables, Stochastic Programming E-Print Series (SPEPS) 9 (2006), to appear in: Computational Optimization and Applications.

[37] F. S. Hillier, G. J. Lieberman, Introduction to Operations Research, New York: McGraw Hill, 2001.

[38] D. Hunter, An upper bound for the probability of a union, Journal of Applied Probability 13 (1976), 597–603.

[39] N. L. Johnson, S. Kotz, N. Balakrishnan, Discrete Multivariate Distributions, Wiley, New York, 1997.

[40] C. Jordan, Calculus of finite differences, 3rd ed., Chelsea, New York, 1965.

[41] S. Karlin, Y. Rinott, Classes of orderings of measures and related correlation inequalities: I. Multivariate Totally Positive Distributions, Journal of Multivariate Analysis 10 (1980), 467–498.

[42] S. Kataoka, A Stochastic Programming Model, Econometrica 31 (1963), 181–196.

[43] P. Kelle, On the safety stock problem for random delivery processes, European Journal of Operations Research 17 (1984), 191–200.

[44] P. Kelle, Safety stock planning in a multi-stage production inventory system, Engineering Costs and Production Economics 9 (1985), 231–237.

[45] J. H. B. Kemperman, The general moment problem, a geometric approach, The Annals of Mathematical Statistics 39 (1968), 93–122.

[46] W. Kuo, M.J. Zuo, Optimal Reliability Modeling: Principles and Applications, Wiley, New York, NY, 2003.

[47] S. M. Kwerel, Most Stringent bounds on aggregated probabilities of partially specified dependent probability systems, J. Amer. Stat. Assoc. 70 (1975), 472–479.

[48] C. E. Lemke, The dual method for solve the linear programming problem, Naval Research Logistic Quarterly 22 (1954), 978–981.

[49] J. S. Lin, C. C. Jan, J. Yuan, On reliability evaluation of a capacitated-flow network in terms of minimal pathsets, Networks, 25 (1995), 131–138.

[50] A. Lisnianski, G. Levitin, Multi-state System Reliability: Assessment, Optimization and Applications, World Scientific, Singapore, 2003.

[51] G. Mádi-Nagy, A. Prékopa, On Multivariate Discrete Moment Problems and their Applications to Bounding Expectations and Probabilities. Mathematics of Operations Research 29 (2004) 229–258.

[52] B. B. Mandelbrot, The fractal geometry of nature, Freeman, New York, 1977.

[53] B. L. Miller, H. M. Wagner, Chance Constrained Programming with Joint Constraints, Operations Research 13 (1965), 930–945.

[54] M. Morhác, An iterative error-free algorithm to solve Vandermonde systems, Applied Mathematics and Computation 117 (2001) 45–54.

[55] T. F. Móri, G. J. Székely, A note on the background of several Bonferroni-Galambos type inequalities, J. Applied Probability 22 (1985) 836–843.

[56] K. Murota, Discrete convex analysis, Mathematical Programming 83 (1998), 313–371.

[57] C. van de Panne, W. Popp, Minimum Cost Cattle Feed under Probabilistic Problem Constraint, Management Science 9 (1963), 405–430.

[58] J. G. Pedersen, On strong unimodality of two-dimensional discrete distributions with applications to M-ancillarity, Scand J. Statist. 2 (1975), 99–102.

[59] O. Platz, A sharp upper probability bound for the occurence of at least m out of n events, J. Applied Probability 22 (1985) 978–981.

[60] A. Prékopa, On probabilistic constrained programming, Proceedings of the Princeton Symposium on Mathematical Programming, Princeton University Press, Princeton, NJ, pp. 113–138, 1970.

[61] A. Prékopa, Logarithmic concave functions with applications to stochastic programming, Acta. Sci. Math. 32 (1971), 301–316.

[62] A. Prékopa, Contributions to the theory of stochastic programming, Mathematical Programming 4 (1973), 202–221.

[63] A. Prékopa, On logarithmic concave measures and functions, Acta. Sci. Math. 34 (1973), 335–343.

[64] A. Prékopa, Programming under Probabilistic Constraints with a Random Technology Matrix, Mathematische Operationsforschung und Statistik, Ser. Optimization 5 (1974), 109–116.

[65] A. Prékopa, P. Kelle, Reliability-type inventory models based n stochastic programming, Mathematical Programming Study9 (1978), 43–58.

[66] A. Prékopa, Network planning using two-stage programming under uncertainty, in: Recent Results in Stochastic Programming, Lecture Notes in Economics and Mathematical Systems 179 (P.Kall and A.Prékopa, eds.), Springer-Verlag, Berlin, (1980) 216–237.

[67] A. Prékopa, Boole-Bonferroni inequalities and linear programming, Oper. Res. 36 (1988), 145–162.

[68] A. Prékopa, Totally positive linear programming problems, L.V. Kantorovich Memorial Volume, Oxford Univ. Press, New York, (1989), 197–207.

[69] A. Prékopa, Sharp bounds on probabilities using linear programming, Oper. Res. 38 (1990), 227–239.

[70] A. Prékopa, The discrete moment problem and linear programming, Discrete Applied Mathematics 27 (1990), 235–254.

[71] A. Prékopa, Dual method for a one-stage stochastic programming problem with random RHS obeying a discrete probability distribution, Z. Operations Research, 34 (1990), 441–461.

[72] A. Prékopa, E. Boros, On the existence of a feasible flow in a stochastic transportation network, Operations Research, 39 (1991), 119–129.

[73] A. Prékopa, E. Boros, Keh-Wei Lih, The Use of Binomial Moments for Bounding Network Reliability. In: Reliability of Computer and Communication Networks, (F. Roberts, F. Hwang, C. Monma, eds.). DIMACS Series in Discrete Mathematics and Theoretical Computer Science, Vol. 5, AMS/ACM 1991, 197–212.

[74] A. Prékopa, Stochastic Programming, Kluwer Academic Publishers, Dordtecht, Boston, 1995.

[75] A. Prékopa, W. Li, Solution of and bounding in a linearly constrained optimization problem with convex, polyhedral objective function, Mathematical Programming 70 (1995), 1–16.

[76] A. Prékopa, A brief introduction to linear programming, Math.Scientist 21 (1996), 85–111.

[77] A. Prékopa, B. Vizvári, T. Badics, Programming under probabilistic constraint with discrete random variable, in:New Trends in Mathematical Programming (F.Giannessi et al. eds.), Kluwer, 1998.

[78] A. Prékopa, The use of discrete moment bounds in probabilistic constrained stochastic programming models, Annals of Operations Reserach 85 (1999), 21–38.

[79] A. Prékopa, Probabilistic programming, Chapter 5 in: A. Ruszczyński and A. Shapiro (Eds.), Stochastic Programming, Handbooks in Operations Research and Management Science, Vol. 10, Elsevier, Amsterdam, pp. 267–351, 2003.

[80] A. Prékopa, J. Long, T. Szántai, New bounds and approximations for the probability distribution of the length of critical path, in: K. Marti, Yu. Ermoliev, G. Pflug (Eds.), Dynamic Stochastic Optimization, Lecture Notes in Economics and Mathematical Systems 532, Springer, Berlin, Heidelberg, 2004, pp. 293–320.

[81] A. Prékopa, L. Gao, Bounding the Probability of the union of events by the use of aggregation and disaggregation in Linear Programs, Discrete Applied Math. 145 (2005), 444–454.

[82] A. Prékopa, M. Subasi, A convexity theorem in programming under probabilistic constraints, RUTCOR Research Report RRR 32-2007.

[83] A. Prékopa, M. Subasi, E. Subasi, Sharp bounds for the probability of the union of events under unimodality condition, European Journal of Pure and Applied Mathematics 1 (2008) 60–81. Available online at `http://www.ejpam.com`.

[84] F. Roberts, F. Hwang, C.Monma (Eds.), Reliability of Computer and Communication Networks, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, 5, 1991.

[85] N. H. Roberts, V. E. Vesely, D. F. Haasl, F. F. Goldberg, Fault Free Handbook, U.S., Nuclear Regularity Commission, Washington, D.C., Nureg-0492.

[86] S. M. Samuels, W. J. Studden, Bonferroni type probability bounds as an application of the theory of Tchebycsheff system, Probability and Mathematics, Papers in Honor of Samuel Karlin, Academic Press (1989), 271–289.

[87] Y. S. Sathe, M. Pradhan, S. P. Shah, Inequalities of the probability of occurence at least m out of n events, Journal of Applied Probability 17 (1980), 1127–1132.

[88] E. Subasi, M. Subasi, A. Prékopa, Discrete moment problems with distributions known to be unimodal, Mathematical Inequalities and Applications, accepted. Available as RUTCOR Research Report RRR 15-2007.

[89] E. Subasi, Sharp bounds for expectations and probabilities in a distribution with given shape information, Dissertation, RUTCOR, Rutgers Center for Operations Research, January 2007.

[90] E. Subasi, M. Subasi, A. Prékopa, On strong unimodality of multivariate discrete distributions, Discrete Applied Mathematics, accepted.
Available online at `http://www.sciencedirect.com` (doi:10.1016/j.dam.2008.02.010)

[91] L. Takács, On the method of inclusion and exclusion, Journal of the American Math Association 62 (1967), 102–113.

[92] E. Tamm, The Quasi-Convexity of Probability and Quantile Functions (in Russian), Eesti NSV Teaduste Akademia Toimetised, FÄuÄusika-Matemaatika (News of the Estonian Academy of Sciences, Math.-Phys.) 25 (1976), 141–145.

[93] E. Tamm, On g-Concave Functions and Probability Measures (in Russian), Eesti NSV Teaduste Akademia Toimetised, FÄuÄusika- Matemaatika (News of the Estonian Academy of Sciences, Math.-Phys.) 26 (1977), 376–379.

[94] P. Veneziani, Graph-based upper bounds for the probability of the union of events, The Electronic Journal of Combinatorics, 15 (2008), #R28.

[95] B. Vizvári, The integer programming background of a stochastic integer programming algorithm of Dentcheva-Prékopa-Ruszczyński, Optimization Methods and Software, 17 (2002), 543–559.

[96] B. Vizvári, New upper bounds on the probability of events based on graph structures, Mathematical Inequalities and Applications 10 (2007), 217-228.

Vita

Munevver Mine Subasi

2001 - 2008	Teaching/Research Assistant, Rutgers University, NJ, USA
1997 - 2000	Teaching/Research Assistant, Cukurova University, Adana, Turkey
1999	M.S., Mathematics, Cukurova University, Adana, Turkey
1997	B.S., Mathematics, Cukurova University, Adana, Turkey
1997	Minor Degree in Education, Cukurova University, Adana, Turkey
2008	E. Subasi, M. Subasi and A. Prékopa. On strong unimodality of multivariate discrete distributions, *Discrete Applied Mathematics*, accepted. Available online at `http://www.sciencedirect.com` (doi:10.1016/j.dam.2008.02.010)
2008	A. Prékopa, M. Subasi, E. Subasi. Sharp bounds for the probability of the union of events under unimodality condition, *European Journal of Pure and Applied Mathematics* 1 (2008) 60–81 (Invited paper). Available online at `http://www.ejpam.com`
2007	A. Prékopa, M. Subasi. A convexity theorem in programming under probabilistic constraints, RUTCOR Research Report RRR 32-2007.
2007	E. Subasi, M. Subasi, A. Prékopa. Discrete moment problems with distributions known to be unimodal, *Mathematical Inequalities and Applications*, accepted. Available as RUTCOR Research Report RRR 15-2007.
2007	M. Lipkowitz, E. Subasi, V. Anbalagan, W. Zhang, P. L. Hammer, J. Robiz, The AASK Investigators. Logical Analysis of Data (LAD) Combinatorial Biomarkers Predict Rate of Decline of GFR in AASK from Serum SELDI-TOF Mass Spectra, *The American Society of Nephrology*, submitted.
2006	E. Subasi, M. Subasi and A. Prékopa. Sharp bounds for probabilities with given shape information, RUTCOR-Rutgers Center for Operations Research RRR 4-2006.
2005	M. Subasi, E. Subasi, M. Anthony, P.L. Hammer. Using a similarity measure for credible classification, *Discrete Applied Mathematics*, accepted. Available as RUTCOR Research Report RRR 39-2005.

2005 E. Subasi, M. Subasi and A. Prékopa. Discrete moment problem with given shape of the distribution, RUTCOR Research Report RRR 41-2005.

Working Papers

- Sharp bounds for the probabilities that at least r and exactly r out of n events occur (with A.Prékopa, E. Subasi)

- Handling missing data using similarity measure (with M. Anthony, P. L. Hammer, E. Subasi)

- Maximization of strongly unimodal multivariate discrete distributions (with A. Prékopa, E. Subasi)

- Detection of suspicious observations using similarity measure (with E. Subasi)