# PROPERTIES AND SOLUTIONS OF A CLASS OF STOCHASTIC PROGRAMMING PROBLEMS WITH PROBABILISTIC CONSTRAINTS

**BY KUNIKAZU YODA**

**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**

**András Prékopa**

**and approved by**

**New Brunswick, New Jersey**

**October, 2013**

## ABSTRACT OF THE DISSERTATION

# Properties and solutions of a class of stochastic programming problems with probabilistic constraints

**by Kunikazu Yoda**

**Dissertation Director: András Prékopa**

We consider two types of probabilistic constrained stochastic linear programming problems and one probability bounding problem.

The first type involves a random left-hand side matrix whose rows are independent and normally distributed. The quasi-concavity of the constraining function needed for the convexity of the problem is ensured if the factors of the function are uniformly quasi-concave. A necessary and sufficient condition is given for that property to hold. We show practical application in optimal portfolio construction.

The second type is the stochastic multidimensional knapsack problem which involves a random left-hand side matrix with independent components and 0-1 decision variables. We show that the problem is convex, under some condition on the parameters, for special continuous and discrete distributions: gamma, normal, Poisson, and binomial. Numerical experiments suggest that the problem can be solved as efficiently as its deterministic version for moderate sized instances.

In the last problem, we formulate the linear programming problems that give improved lower and upper bounds on the probability of the union of events when the probabilities of some individual or intersections of events in a first few terms of the inclusion-exclusion principle are 0 or very small.

## Acknowledgements

First and foremost, I would like to express my sincere gratitude to my advisor Dr. András Prékopa for his guidance, support, patience, and encouragement throughout my study. I am very grateful to him for I have learned a great deal about how we tackle difficult problems and proceed our research. His bright mind, particularly in the way of thinking about mathematical problems, has stimulated my intellectual development. I would like to thank other committee members, Dr. Endre Boros, Dr. Adi Ben-Israel, Dr. Vladimir Gurvich, Dr. Melike Baykal-Gursoy, and Dr. Munevver Mine Subasi, for their valuable assistance. I would like to thank Terry Hart, Clare Smietana, Lynn Agre, Katie D'Agosta, and my colleagues at RUTCOR for their help and friendly environment they have provided. I would like to acknowledge that part of this research was supported by the National Science Foundation under the grant number CMMI-085663.

I would like to thank the parents of me and my wife and my other relatives for their support. Above all, I would like to thank my wife Fumie for her endless support and warm encouragement and my two sons Ryoto and Shuya, who bring joy to our lives.

# Dedication

To my sons Ryoto and Shuya and my wife Fumie.

# Table of Contents

**Abstract** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . ii

**Acknowledgements** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . iii

**Dedication** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . iv

**List of Tables** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1

**List of Figures** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 2

**1. Introduction** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 3

**2. Uniform quasi-concavity in probabilistic constrained stochastic programming** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 11

2.1. Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 11

2.2. Preliminary results . . . . . . . . . . . . . . . . . . . . . . . . . . . 13

2.3. The main result . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 16

2.4. Application in portfolio optimization . . . . . . . . . . . . . . . . . . 19

**3. Convexity and solutions of the stochastic multidimensional knapsack problem with probabilistic constraints** . . . . . . . . . . . . . . . . . . 24

3.1. Introduction. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 24

3.2. Formulation of the problem. . . . . . . . . . . . . . . . . . . . . . . . 26

3.3. Convexity of the stochastic multidimensional knapsack problem. . . . . 30

3.3.1. Convexity result for the gamma distribution. . . . . . . . . . . . 30

3.3.2. Convexity result for the normal distribution. . . . . . . . . . . . 32

3.3.3. Convexity result for the Poisson distribution. . . . . . . . . . . . 37

3.3.4. Convexity result for the binomial distribution. . . . . . . . . . . 39

3.4. Solutions and computational examples. . . . . . . . . . . . . . . . . . . . . . 42

3.4.1. Solutions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 42

3.4.2. Execution times and quality of solutions. . . . . . . . . . . . . . 42

3.4.3. Project selection problem. . . . . . . . . . . . . . . . . . . . . . 48

**4. Improved bounds on the probability of the union of events, some intersections of which are empty** . . . . . . . . . . . . . . . . . . . . . . . . . 50

4.1. Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 50

4.2. Improved bounds by the maximum independent set problem and its extension . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 53

4.3. Numerical examples . . . . . . . . . . . . . . . . . . . . . . . . . . . 57

**References** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 60

**Vita** . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 66

# List of Tables

2.1. Expected losses in May 2009. . . . . . . . . . . . . . . . . . . . . . . . . 22

2.2. Covariance Matrix in May 2009. . . . . . . . . . . . . . . . . . . . . . . . 22

2.3. Values of nine assets, May 2009. . . . . . . . . . . . . . . . . . . . . . . . 23

3.1. CPU times (in seconds) for computing optimal solutions. . . . . . . . . . 45

3.2. Objective values and probabilities (in %) satisfying the joint knapsack constraints (part 1). . . . . . . . . . . . . . . . . . . . . . . . . . . . . 46

3.3. Objective values and probabilities (in %) satisfying the joint knapsack constraints (part 2). . . . . . . . . . . . . . . . . . . . . . . . . . . . . 47

3.4. Parameters for the projects. . . . . . . . . . . . . . . . . . . . . . . . . . . 49

4.1. Lower and upper bounds on the union of events. . . . . . . . . . . . . . 59

# List of Figures

1.1. Illustration of a reservoir system to protect a downstream area from floods. 5

# Chapter 1

# Introduction

We study three topics relating to probabilistic constraints in this paper. The first two topics concern stochastic linear programming problems with probabilistic constraints. The first topic, presented in Chapter 2, is based on Prékopa, Yoda, and Subasi [65] that characterizes a class of quasi-concavity when the random left-hand side matrix follows normal distributions. The second topic, presented in Chapter 3, is based on Yoda and Prékopa [79] that involves a stochastic version of the multidimensional knapsack problem which is considered a probabilistic constrained stochastic programming problem with 0-1 decision variables and a random left-hand side matrix. The last topic, presented in Chapter 4, is based on Yoda and Prékopa [80] that concerns the probability bounding problem for the union of events.

In a variety of industrial and engineering problems, such as production planning and scheduling, logistics, financial modeling, and telecommunications network design, there is a need to make an optimal decision under uncertainty. There are several ways to handle uncertainty in optimization problems. The classical stochastic programming with recourse, the stochastic programming with probabilistic constraints, and the robust optimization can be mentioned. Our choice is the stochastic programming with probabilistic constraints, which was modeled by Prékopa [51, 53] (see also [58, 60] for summary) in the following way:

$$\text{minimize } f(\boldsymbol{x}) \tag{1.0.1a}$$

$$\text{subject to } \; h_0(\boldsymbol{x}) := \mathbb{P}\left(g_i(\boldsymbol{x},\boldsymbol{\eta}) \geq 0,\ i=1,\dots,r\right) \geq p, \tag{1.0.1b}$$

$$h_i(\boldsymbol{x}) \geq 0,\ i=1,\dots,s, \tag{1.0.1c}$$

where $\boldsymbol{x} \in \mathbb{R}^n$ is the decision vector, $\boldsymbol{\eta} \in \mathbb{R}^\ell$ is a random vector, $f(\boldsymbol{x}),\ g_i(\boldsymbol{x},\boldsymbol{y}),\ h_i(\boldsymbol{x})$

are some functions, and $p$ is a prescribed probability level, e.g., $p = 0.9, 0.95,$ or 0.99, chosen by the decision maker in order to model the reliability of the system. The probability level $p$ ensures that the state of the system remains within the allowable subset with a probability at least as high as $p$ regardless of outcomes of the random parameters.

Stochastic programming with probabilistic constraints was first introduced by Charnes, Cooper, and Symonds [11]. Their models are based on *individual probabilistic constraints*, where instead of using the constraint (1.0.1b), the following constraints are used.

$$\mathbb{P}\left(g_i(\boldsymbol{x}, \boldsymbol{\eta}) \geq 0\right) \geq p_i, \quad i = 1, \ldots, r, \tag{1.0.2}$$

where $p_i$'s are probability levels chosen by the decision maker. If random variables $g_i(\boldsymbol{x}, \boldsymbol{\eta}),\ i = 1, \ldots, r$ are independent of each other, then the use of (1.0.2) is appropriate. Note that in this case the constraint (1.0.1b) has a simpler form:

$$\mathbb{P}\left(g_i(\boldsymbol{x}, \boldsymbol{\eta}) \geq 0,\ i = 1, \ldots, r\right) = \prod_{i=1}^{r} \mathbb{P}\left(g_i(\boldsymbol{x}, \boldsymbol{\eta}) \geq 0\right) \geq p, \tag{1.0.3}$$

which is not the same as (1.0.2). In general, however, when random variables $g_i(\boldsymbol{x}, \boldsymbol{\eta}),\ i = 1, \ldots, r$ are dependent, the *joint probabilistic constraint* (1.0.1b) must be used. Let us look at an example of the reservoir system taken from Prékopa [58], shown in Figure 1.1, to be built to protect a downstream area from floods caused by random inflows of water. Assume for simplicity that a flood can occur once in a year when the random amounts of water to be retained by reservoirs 1 and 2 are $\xi_1$ and $\xi_2$, respectively. We want to find optimal capacities $x_1$ and $x_2$ of the two reservoirs so that a flood occurs no more frequently than once in a hundred years. The amount of water that overflows from reservoir 1 is $[\xi_1 - x_1]_+$. Thus a flood can be prevented if and only if $[\xi_1 - x_1]_+ + \xi_2 \leq x_2$, which is equivalent to $\xi_1 + \xi_2 \leq x_1 + x_2,\ \xi_2 \leq x_2$. So we must use the joint probabilistic constraint:

$$\mathbb{P}\begin{pmatrix} \xi_1 + \xi_2 & \leq x_1 + x_2 \\ \xi_2 & \leq x_2 \end{pmatrix} \geq 0.99,$$

while the individual constraints in this case are meaningless.

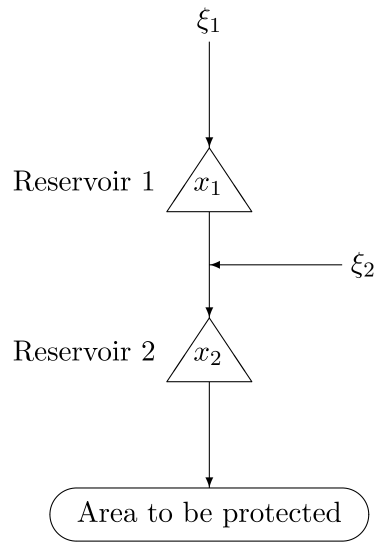


Figure 1.1: Illustration of a reservoir system to protect a downstream area from floods.

Logconcavity is an important concept in stochastic programming. The notion of an $r$-times positive sequence was first introduced by Fekete (see the collection of letters between Fekete and Pólya [21]). For the case of $r = 2$ the definition provides us with the same notion that we call today a logconcave sequence. The generalization to the continuous case is straightforward. Important theoretical foundations of logconcave measures and logconcave functions in the multivariate case are found in Prékopa [52, 53, 54, 55] (see also [58, 60] for summary). First we review logconcavity and quasi-concavity.

**Definition 1.0.1** (Logconcave function)**.** *A nonnegative function $f : S \mapsto \mathbb{R}_+$ defined on a convex subset $S$ of $\mathbb{R}^n$ is said to be logarithmically concave (or logconcave for short) if for all $\boldsymbol{x}, \boldsymbol{y} \in S$ and $\lambda \in (0, 1)$ we have*

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

If $f(\boldsymbol{x}) > 0$ for $\boldsymbol{x} \in S$, then this means that $\log f(\boldsymbol{x})$ is a concave function on $S$. Every concave function that is nonnegative on its domain is logconcave. The product of logconcave functions is also logconcave.

**Definition 1.0.2** (Logconcave measure: Prékopa [52])**.** *A probability measure $P$ defined on the Borel sets of $\mathbb{R}^n$ is said to be logarithmically concave (or logconcave for short)* *if for every pair of convex subsets $A, B$ of $\mathbb{R}^n$ and $\lambda \in (0,1)$ we have*

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

*where $\lambda A + (1-\lambda)B = \{\, \lambda x + (1-\lambda)y \mid x \in A, y \in B \,\}$.*

Quasi-concavity is a generalization of concavity and it has application in microeconomics and finance as utility functions and measures of risk.

**Definition 1.0.3** (Quasi-concave function)**.** *A function $f : S \mapsto \mathbb{R}$ defined on a convex subset $S$ of $\mathbb{R}^n$ is said to be quasi-concave if for all $\boldsymbol{x}, \boldsymbol{y} \in S$ and $\lambda \in (0,1)$ we have*

$$f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \geq \min(f(\boldsymbol{x}), f(\boldsymbol{y})).$$

An alternative definition of a quasi-concave function $f(\boldsymbol{x})$ is that the upper level set $\{\, \boldsymbol{x} \mid f(\boldsymbol{x}) \geq \alpha \,\}$ for any $\alpha$ is convex. A logconcave function is also quasi-concave. Quasi-concavity is preserved in non-decreasing transformations (i.e. if $g : \mathbb{R}^n \mapsto \mathbb{R}$ is quasi-concave and $h : \mathbb{R} \mapsto \mathbb{R}$ non-decreasing, then $f = h \circ g$ is quasi-concave). The sum or product of quasi-concave functions on the same domain is not quasi-concave, in general. In Chapter 2 we introduce a new class of quasi-concave functions, called uniformly quasi-concave functions, where the sum and product (for positive functions) of them is also a quasi-concave function.

Next we review the basic theorems for logconcavity.

**Theorem 1.0.1** (Prékopa [52, 53])**.** *If the probability measure $\mathbb{P}$ is absolutely continuous with respect to the Lebesgue measure and is generated by a logconcave probability density function then the measure $\mathbb{P}$ is logconcave.*

**Theorem 1.0.2** (Prékopa [52, 53])**.** *If $\boldsymbol{\xi} \in \mathbb{R}^n$ is a random vector, the probability distribution of which is logconcave, then the probability distribution function $F(\boldsymbol{x}) = \mathbb{P}(\boldsymbol{\xi} \leq \boldsymbol{x})$ is a logconcave function in $\mathbb{R}^n$.*

**Theorem 1.0.3** (Prékopa [53])**.** *If $g_1(\boldsymbol{x}, \boldsymbol{y}), \ldots, g_r(\boldsymbol{x}, \boldsymbol{y})$ are quasi-concave functions of the variables $\boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{y} \in \mathbb{R}^m$ and $\boldsymbol{\xi} \in \mathbb{R}^m$ is a random variable that has logconcave probability distribution, then the function $G(\boldsymbol{x}) = \mathbb{P}(g_1(\boldsymbol{x}, \boldsymbol{\xi}) \geq 0, \ldots, g_r(\boldsymbol{x}, \boldsymbol{\xi}) \geq 0)$, $\boldsymbol{x} \in \mathbb{R}^n$ is logconcave.*

**Theorem 1.0.4** (Prékopa [53]). *If $f(\boldsymbol{x}, \boldsymbol{y}), \boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{y} \in \mathbb{R}^m$ is a logconcave function, then*

$$\int_{\mathbb{R}^m} f(\boldsymbol{x}, \boldsymbol{y}) d\boldsymbol{y}, \quad \boldsymbol{x} \in \mathbb{R}^n$$

*is also a logconcave function.*

In many applications the functions $f(\boldsymbol{x}), h_i(\boldsymbol{x})$ in (1.0.1) are linear and the stochastic constraints $g_i(\boldsymbol{x}, \boldsymbol{\eta}) \geq 0,\ i = 1, \ldots, r$ have the linear form $\boldsymbol{\xi} - \Xi \boldsymbol{x} \geq 0$. The probabilistic constraint then specializes as

$$h_0(\boldsymbol{x}) := \mathbb{P}(\Xi \boldsymbol{x} \leq \boldsymbol{\xi}) \geq p. \tag{1.0.4}$$

One of $\Xi \in \mathbb{R}^{m \times n}$ and $\boldsymbol{\xi} \in \mathbb{R}^m$ is random and the other is constant. Difficulty in this problem is that the set $\mathcal{S} := \{ \boldsymbol{x} \mid h_0(\boldsymbol{x}) \geq p \}$ is nonconvex and the function $h_0(\boldsymbol{x})$ on $\mathcal{S}$ is nonsmooth or even discontinuous, in general.

For the case of a constant matrix denoted by $T$ (instead of $\Xi$) and a random vector $\boldsymbol{\xi}$, in which (1.0.4) is expressed as $\mathbb{P}(T\boldsymbol{x} \leq \boldsymbol{\xi}) \geq p$, several results are known. For the case of continuously distributed $\boldsymbol{\xi}$ general theorems are available to ensure the convexity of the set $\{ \boldsymbol{x} \mid \mathbb{P}(T\boldsymbol{x} \leq \boldsymbol{\xi}) \geq p \}$ (see Prékopa [58]). The solution of problems where $\boldsymbol{\xi}$ in (2.1.2) is a discrete random vector is more recent. The key concept here is that of a $p$-efficient point, introduced in Prékopa [57] and further developed and used in Prékopa, Vizvári, and Badics [64], Dentcheva, Prékopa, and Ruszczyński [18], and Boros et al. [7]. See also Beraldi and Ruszczyński [4], Vizvári [74], and Luedtke, Ahmed, and Nemhauser [48].

For the case of a random matrix $\Xi = [\boldsymbol{\xi}^1, \ldots, \boldsymbol{\xi}^m]^T$ and a constant vector denoted by $\boldsymbol{d} = (d_1, \ldots, d_m)^T$ (instead of $\boldsymbol{\xi}$), in which (1.0.4) is expressed as $\mathbb{P}(\Xi\boldsymbol{x} \leq \boldsymbol{d}) \geq p$, few results are known. The earliest papers dealing with this case for the normally distributed rows are by Kataoka [40] and van de Panne and Popp [72], in which, however, $\Xi$ has only one ($m = 1$) row and to establish the convexity of the set $\left\{ \boldsymbol{x} \;\middle|\; \mathbb{P}((\boldsymbol{\xi}^1)^T\boldsymbol{x} \leq d_1) \geq p \right\}$ is relatively easy. If $\Xi$ has more than one row, then even if they are independent, it is not easy to ensure the convexity of the set $\{ \boldsymbol{x} \mid \mathbb{P}(\Xi\boldsymbol{x} \leq \boldsymbol{d}) \geq p \}$. The first paper where convexity theorems are presented for this case is by Prékopa [54]. The paper by Henrion and Strugarek [36] can be mentioned, where significant progress in this direction has been made. An important result in the class of log-concave symmetric distribution has been found by Lagoa, Li, and Sznaier [46].

In Chapter 2, we study the problem with a random left-hand side matrix $\Xi$ whose rows are independent and normally distributed, in which the probabilistic constraint is expressed as $h_0(\boldsymbol{x}) := \mathbb{P}(\Xi\boldsymbol{x} \leq \boldsymbol{d}) \geq p$. For the convexity of the problem the quasi-concavity of $h_0(\boldsymbol{x})$ is needed, which is ensured if the factors of $h_0(\boldsymbol{x})$ are uniformly quasi-concave. We give a necessary and sufficient condition on the parameters of the normal distributions for that property to hold.

In Chapter 3, our work concerns the knapsack problem, which is one of the most fundamental combinatorial optimization problems with a wealth of applications in industries. The most basic form, the 0-1 one-dimensional single knapsack problem, can be stated as follows: Given a set of items, each with a value and a size, determine a subset maximizing the total value while keeping the total size within a given capacity. We consider the 0-1 knapsack problem in this paper as opposed to the bounded or unbounded knapsack problem. One generalization of the problem includes the multidimensional knapsack problem (or multiply constrained knapsack problem), where each item has multiple attributes (sizes), such as length, volume, weight, etc., and a knapsack has a capacity for each attribute.

A couple of works are known for the probabilistic constrained stochastic one-dimensional knapsack problem. Goyal and Ravi [32] showed a polynomial time approximation scheme via a parametric LP reformulation when the random item attributes are independent and normally distributed. Fortz and Poss [23] showed that the problem can be linearized when the random item attributes are independent and follow normal or gamma distributions under some regularity condition. Our work can be considered as an extension of the latter work to the multidimensional knapsack problem. While we study the stochastic multidimensional knapsack problem, our convexity result holds for a broader class of stochastic combinatorial optimization problems whose underlying deterministic models are formulated by linear inequalities and 0-1 decision variables.

In Chapter 4, we consider the probability bounds on the union of random events, which have applications in stochastic programming with probabilistic constraints, reliability of networks, and other sciences involving in stochastic systems. While computing exact probabilities in high-dimensional spaces is often intractable, the information on low dimensional probabilities helps to obtain good approximations.

The inclusion-exclusion principle (see de Moivre [15], da Silva [13], and Sylvester [68]) gives the exact probability of the union of events but the formula is impractical if the number of events is large. The Bonferroni inequalities (see Bonferroni [5]) give upper and lower bounds using only a first few terms of the inclusion-exclusion principle. These bounds are usually very weak. The best possible (sharp) bounds using few terms (the number of the terms is called the order of the bounds) have been found in closed forms. The second order sharp lower bound was obtained by Dawson and Sankoff [14] and its upper bound by Kwerel [44, 45] and Sathe et al. [67]. The third order sharp bounds were obtained by Kwerel [44, 45] and Boros and Prékopa [8]. The fourth order sharp upper bound was obtained by Boros and Prékopa [8]. While the fifth or higher order sharp bounds have not been known in closed forms, Prékopa [56] observed that all these bounds are the optimal objective values of binomial moment problems obtained from the formulation by Hailperin [35]. By simply using the first few terms which are aggregated information, we lose the information in individual events. Hailperin [35] provided the Boolean probability bounding scheme, which was initiated by Boole [6], utilizing the probabilities of individual and intersections of events appeared in the first few terms. Although these bounds are much better than those from the binomial moment problems, the formulation is impractical if the number of events is large due to the exponential number of decision variables.

Probability bounds that utilize structures of events have been studied. Hunter's upper bound (see Hunter [37] and Worsley [77]) uses graph structures. It was generalized by Tomescu [71] and improved on by Bukszár and Prékopa [9]. Prékopa and Gao [62] defined the linear programming problems balancing the size of the formulation and the quality of bounds. Prékopa, M. Subasi, and E. Subasi [63] gave the sharp bounds assuming unimodality of the probability distribution. Our contribution is that we formulate the linear programming problems that give improved bounds when the probabilities of some individual or intersections of events are 0 or very small.

# Chapter 2

# Uniform quasi-concavity in probabilistic constrained stochastic programming

## 2.1 Introduction

The stochastic programming problem, termed programming under probabilistic constraints can be formulated in the following way:

$$\begin{aligned} &\text{minimize} \quad f(\boldsymbol{x}) \qquad\qquad (2.1.1)\\ &\text{subject to } h_0(\boldsymbol{x}) = \mathbb{P}\left(g_i(\boldsymbol{x},\boldsymbol{\xi}) \geq 0,\ i = 1,\dots,r\right) \geq p \\ &\qquad\qquad h_i(\boldsymbol{x}) \geq 0,\ i = 1,\dots,m, \end{aligned}$$

where $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{\xi} \in \mathbb{R}^q$, $f(\boldsymbol{x}), g_i(\boldsymbol{x},\boldsymbol{y}), i = 1,\dots,r,\ h_i(\boldsymbol{x}), i = 1,\dots,m$ are some functions and $p$ is a fixed large probability, e.g., $p = 0.9, 0.95, 0.99$. In many applications the stochastic constraints have the form $\boldsymbol{\xi} - T\boldsymbol{x} \geq 0$ and the probabilistic constraint specializes as

$$h_0(\boldsymbol{x}) = \mathbb{P}(T\boldsymbol{x} \leq \boldsymbol{\xi}) \geq p. \qquad (2.1.2)$$

For the case of continuously distributed random vector $\boldsymbol{\xi}$ general theorems are available to ensure the convexity of the set determined by the probabilistic constraint in (2.1.1). For example, if $g_i,\ i = 1,\dots,r$ are concave or at least quasi-concave in all variables and $\boldsymbol{\xi}$ has a logconcave p.d.f., then the function $h_0(\boldsymbol{x})$ is logconcave and the set $\{\, \boldsymbol{x} \mid h_0(\boldsymbol{x}) \geq p \,\}$ is convex (see, e.g., Prékopa [58, 60]). This implies that if $\boldsymbol{\xi}$ has the above-mentioned property, then the set determined by the constraint (2.1.2) is convex. Many applications of the model with probabilistic constraint (2.1.2) have been carried out, for the cases of some special continuous multivariate distributions such as normal, gamma, and Dirichlet, and problem solving packages have been developed (see, e.g., Prékopa [60], Deák [17], Kall and Mayer [39], and Szántai [69]).

The solution of problems where $\boldsymbol{\xi}$ in (2.1.2) is a discrete random vector is more recent. The key concept here is that of a $p$-efficient point, introduced in Prékopa [57] and further developed and used in Prékopa, Vizvári, and Badics [64], Dentcheva, Prékopa, and Ruszczyński [18], and Boros et al. [7]. See also other methods in Beraldi and Ruszczyński [4], Vizvári [74], and Luedtke, Ahmed, and Nemhauser [48].

For the case of a random $T$ in the constraint (2.1.2), few results are known. The earliest papers dealing with a random matrix $T$ in the probabilistic constraint are Kataoka [40] and van de Panne and Popp [72]. In these papers, however, there is only one stochastic constraint and to establish the concavity of the set $\{\, \boldsymbol{x} \mid \mathbb{P}(T\boldsymbol{x} \leq \boldsymbol{\xi}) \geq p \,\}$ is relatively easy (see, the proof of Lemma 2.2.2).

The first paper where convexity theorems are presented for the set of feasible solutions and random matrix $T$ has more than one row, is by Prékopa [54]. If $T$ has more than one row, then even if they are independent, it is not easy to ensure the convexity of the set $\{\, \boldsymbol{x} \mid \mathbb{P}(T\boldsymbol{x} \leq \boldsymbol{\xi}) \geq p \,\}$. The paper by Henrion and Strugarek [36] can be mentioned, where significant progress in this direction has been made. The problem is that the product or sum of quasi-concave functions is not quasi-concave, in general. We briefly recall the results of the paper by Prékopa [54] (see also Prékopa [58] pp. 312–314).

**Theorem 2.1.1** (Prékopa [54])**.** *Let $\boldsymbol{\xi}$ be constant and $T$ a random matrix with independent, normally distributed rows (or columns) such that their covariance matrices are constant multiples of each other. Then $h(\boldsymbol{x}) = \mathbb{P}(T\boldsymbol{x} \leq \boldsymbol{\xi})$ is a quasi-concave function on the set $\{\, \boldsymbol{x} \mid h(\boldsymbol{x}) \geq 1/2 \,\}$.*

We introduce a special class of quasi-concave functions.

**Definition 2.1.1** (Uniformly quasi-concave functions)**.** *Let $h_1(\boldsymbol{x}), \ldots, h_r(\boldsymbol{x})$ be quasi-concave functions on a convex set $E \in \mathbb{R}^n$. We say that they are uniformly quasi-concave functions if for any $\boldsymbol{x}, \boldsymbol{y} \in E$ either*

$$\min(h_i(\boldsymbol{x}), h_i(\boldsymbol{y})) = h_i(\boldsymbol{x}), \;\; i = 1, \ldots, r$$

*or*

$$\min(h_i(\boldsymbol{x}), h_i(\boldsymbol{y})) = h_i(\boldsymbol{y}),\ \ i = 1, \ldots, r.$$

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.

In this chapter we look at probabilistic constraints of the type

$$\mathbb{P}(T\boldsymbol{x} \leq \boldsymbol{b}) \geq p, \tag{2.1.3}$$

where $T$ is a random matrix that has independent, normally distributed rows and $\boldsymbol{b}$ is a constant vector. The constraining function in (2.1.3) is the product of special quasi-concave functions and we show that the uniform quasi-concavity of the factors implies that the covariance matrices of the rows are constant multiples of each other. Section 2 and 3 are devoted to this. In section 4 we show that this very special type of probabilistic constraint is still applicable to solve portfolio optimization problems. We present some numerical results in this respect.

## 2.2 Preliminary results

First we provide a necessary condition for continuously differentiable and uniformly quasi-concave functions $h_1(\boldsymbol{x}), \ldots, h_r(\boldsymbol{x})$ on an open convex set.

**Lemma 2.2.1.** *If $h_1(\boldsymbol{x}), \ldots, h_r(\boldsymbol{x})$ are continuously differentiable and uniformly quasi-concave on an open convex set $E$, then any nonzero gradients $\nabla h_i(\boldsymbol{x}), \nabla h_j(\boldsymbol{x})$ are positive multiples of each other, i.e., for any $i, j \in \{1, \ldots, r\}$, there exists a positive-valued function $\alpha_{ij}(\boldsymbol{x}) = 1/\alpha_{ji}(\boldsymbol{x}) > 0$ defined on $E_{ij} = \{\, \boldsymbol{x} \in E \mid \nabla h_i(\boldsymbol{x}) \neq 0, \nabla h_j(\boldsymbol{x}) \neq 0 \,\} = E_{ji}$ such that for all $\boldsymbol{x} \in E_{ij}$ we have*

$$\nabla h_i(\boldsymbol{x}) = \alpha_{ij}(\boldsymbol{x}) \nabla h_j(\boldsymbol{x}) \tag{2.2.1}$$

*Proof.* We show that (2.2.1) holds for all $\boldsymbol{x} \in E_{ij}$ by contradiction. Suppose that for some $\boldsymbol{x} \in E_{ij}$ we cannot find an $\alpha_{ij}(\boldsymbol{x}) > 0$ satisfying (2.2.1). Without loss of generality we assume that $i = 1, j = 2$.

By the Farkas Lemma, either one of the following two systems has a solution

(i) $\nabla h_2(\boldsymbol{x})^T \boldsymbol{d} \leq 0,\ \nabla h_1(\boldsymbol{x})^T \boldsymbol{d} > 0$

(ii) $\nabla h_1(\boldsymbol{x}) = \lambda \nabla h_2(\boldsymbol{x}),\ \lambda \geq 0$

First, note that since $\nabla h_1(\boldsymbol{x}) \neq 0$ and $\nabla h_2(\boldsymbol{x}) \neq 0$, $\lambda = 0$ cannot be a solution of (ii). Also, $\lambda > 0$ cannot be a solution of (ii), otherwise we can define $\alpha_{12}(\boldsymbol{x}) = \lambda > 0$. Hence, (i) has a solution $\boldsymbol{d}_1$. Similarly, since $\nabla h_2(\boldsymbol{x}) = \alpha_{21}(\boldsymbol{x}) \nabla h_1(\boldsymbol{x})$ does not hold for any defined value of $\alpha_{21}(\boldsymbol{x}) = 1/\alpha_{12}(\boldsymbol{x}) > 0$ by the assumption, (i) with 1 and 2 interchanged has a solution $\boldsymbol{d}_2$. So we have

$$\nabla h_2(\boldsymbol{x})^T \boldsymbol{d}_1 \leq 0,\ \nabla h_1(\boldsymbol{x})^T \boldsymbol{d}_1 > 0,$$
$$\nabla h_1(\boldsymbol{x})^T \boldsymbol{d}_2 \leq 0,\ \nabla h_2(\boldsymbol{x})^T \boldsymbol{d}_2 > 0.$$

Let $\boldsymbol{d} := \boldsymbol{d}_1 - \boldsymbol{d}_2$. Then it follows that

$$\nabla h_1(\boldsymbol{x})^T \boldsymbol{d} > 0,\ \nabla h_2(\boldsymbol{x})^T \boldsymbol{d} < 0. \tag{2.2.2}$$

Note that $\boldsymbol{d} \neq 0$. By the use of finite Taylor series expansions we can write:

$$h_1(\boldsymbol{x} + \varepsilon \boldsymbol{d}) = h_1(\boldsymbol{x}) + (\nabla h_1(\boldsymbol{x})^T \boldsymbol{d})\varepsilon + o(\varepsilon), \tag{2.2.3}$$
$$h_2(\boldsymbol{x} + \varepsilon \boldsymbol{d}) = h_2(\boldsymbol{x}) + (\nabla h_2(\boldsymbol{x})^T \boldsymbol{d})\varepsilon + o(\varepsilon). \tag{2.2.4}$$

Since $E_{12}$ is an open set, we can select $\varepsilon > 0$ small enough so that

$$\exists \boldsymbol{y} := \boldsymbol{x} + \varepsilon \boldsymbol{d} \in E_{12},\ \boldsymbol{y} \neq \boldsymbol{x}, \quad h_1(\boldsymbol{y}) > h_1(\boldsymbol{x}), \quad h_2(\boldsymbol{y}) < h_2(\boldsymbol{x})$$

Hence $h_1(\boldsymbol{x}), \ldots, h_r(\boldsymbol{x})$ are not uniformly quasi-concave, which is a contradiction. $\square$

For $r = 1$, let us consider the function

$$h(\boldsymbol{x}) = \mathbb{P}(T\boldsymbol{x} \leq b), \tag{2.2.5}$$

where $T$ is a random row vector and $b$ is a constant. The following lemma was first proved by Kataoka [40] and van de Panne and Popp [72]. See also Prékopa [58].

**Lemma 2.2.2** (Kataoka [40] and van de Panne and Popp [72]). *If $T$ has normal distribution, then the function $h(\boldsymbol{x})$ is quasi-concave on the set*

$$\left\{ \boldsymbol{x} \;\middle|\; \mathbb{P}(T\boldsymbol{x} \leq b) \geq \frac{1}{2} \right\}.$$

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

$$h_i(\boldsymbol{x}) = \mathbb{P}(T_i\boldsymbol{x} \leq b_i), \; i = 1, \ldots, r, \tag{2.2.6}$$

where each row vector $T_i, \; i = 1, \ldots, r$ has normal distribution with mean vector $\boldsymbol{\mu}_i = \mathbb{E}(T_i^T)$ and covariance matrix $C_i = \mathbb{E}((T_i^T - \boldsymbol{\mu}_i)(T_i^T - \boldsymbol{\mu}_i)^T)$, and $b = (b_1, \ldots, b_r)^T$ is constant.

Suppose $b_i > 0, \; i = 1, \ldots, r$. Let us define set $E$ as follows:

$$\begin{aligned} &E \text{ is convex.} \\ &E \supset B \supset \{0\} \text{ for some open set } B. \\ &\text{Each } h_i(\boldsymbol{x}), \; i = 1, \ldots, r \text{ is quasi-concave on } E. \end{aligned} \tag{2.2.7}$$

One example of such $E$ is

$$E = \bigcap_{i=1}^{r} \left\{ \boldsymbol{x} \;\middle|\; h_i(\boldsymbol{x}) \geq \frac{1}{2} \right\}. \tag{2.2.8}$$

Note that by lemma 2.2.2, $h_i(\boldsymbol{x})$ is quasi-concave on the convex set $E_i = \{ \boldsymbol{x} \mid h_i(\boldsymbol{x}) \geq 1/2 \}$ and that for a sufficiently small open ball $B_\epsilon(0) = \{ \boldsymbol{x} \mid \|\boldsymbol{x}\| < \epsilon \}$ around the origin, $h_i(\boldsymbol{x}) \geq 1/2, \; \forall \boldsymbol{x} \in B_\epsilon(0)$, thus $E_i \supset B_\epsilon(0)$. Also note that the intersection of convex sets is a convex set. If rows $T_1, \ldots, T_r$ of $T$ are independent and $h_1(\boldsymbol{x}), \ldots, h_r(\boldsymbol{x})$ are uniformly quasi-concave, then $h(\boldsymbol{x}) = \mathbb{P}(T_i\boldsymbol{x} \leq b_i, \; i = 1, \ldots, r) = \prod_{i=1}^{r} \mathbb{P}(T_i\boldsymbol{x} \leq b_i) = h_1(\boldsymbol{x}) \cdots h_r(\boldsymbol{x})$ is quasi-concave on $E$.

Suppose $b_i > 0$ and $C_i$ is positive definite for $i = 1, \ldots, r$.

$$h_i(\boldsymbol{x}) = \begin{cases} \Phi\left(\dfrac{b_i - \boldsymbol{\mu}_i^T\boldsymbol{x}}{\sqrt{\boldsymbol{x}^T C_i \boldsymbol{x}}}\right) & \text{for } \boldsymbol{x} \neq 0, \\ \mathbb{P}(0 \leq b_i) = 1 & \text{for } \boldsymbol{x} = 0. \end{cases} \tag{2.2.9}$$

Since

$$\lim_{\boldsymbol{x} \to 0} h_i(\boldsymbol{x}) = \lim_{t \to \infty} \Phi(t) = 1 = h_i(0),$$

$h_i(\boldsymbol{x})$ is continuous at $\boldsymbol{x}=0$. Let us calculate the gradient of $h_i(\boldsymbol{x})$ for $\boldsymbol{x} \in \operatorname{int}(E)\setminus\{0\}$.

$$\begin{aligned}
\nabla h_i(\boldsymbol{x}) &= \nabla\Phi\left(\frac{b_i-\boldsymbol{\mu}_i^T\boldsymbol{x}}{\sqrt{\boldsymbol{x}^TC_i\boldsymbol{x}}}\right) \\
&= \varphi\left(\frac{b_i-\boldsymbol{\mu}_i^T\boldsymbol{x}}{\sqrt{\boldsymbol{x}^TC_i\boldsymbol{x}}}\right)\nabla\frac{b_i-\boldsymbol{\mu}_i^T\boldsymbol{x}}{\sqrt{\boldsymbol{x}^TC_i\boldsymbol{x}}} \\
&= \varphi\left(\frac{b_i-\boldsymbol{\mu}_i^T\boldsymbol{x}}{\sqrt{\boldsymbol{x}^TC_i\boldsymbol{x}}}\right)\frac{-\sqrt{\boldsymbol{x}^TC_i\boldsymbol{x}}\boldsymbol{\mu}_i-(b_i-\boldsymbol{\mu}_i^T\boldsymbol{x})C_i\boldsymbol{x}/\sqrt{\boldsymbol{x}^TC_i\boldsymbol{x}}}{\boldsymbol{x}^TC_i\boldsymbol{x}} \\
&= -\varphi\left(\frac{b_i-\boldsymbol{\mu}_i^T\boldsymbol{x}}{\sqrt{\boldsymbol{x}^TC_i\boldsymbol{x}}}\right)\frac{(\boldsymbol{x}^TC_i\boldsymbol{x})\boldsymbol{\mu}_i+(b_i-\boldsymbol{\mu}_i^T\boldsymbol{x})C_i\boldsymbol{x}}{(\boldsymbol{x}^TC_i\boldsymbol{x})^{3/2}},
\end{aligned} \tag{2.2.10}$$

where $\varphi(t)$ is the p.d.f. of the one-dimensional standard normal distribution.

$$\varphi(t) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}t^2\right).$$

For any fixed $\boldsymbol{x} \neq 0$, we have

$$\begin{aligned}
&\lim_{\varepsilon\downarrow 0}\nabla h_i(\varepsilon\boldsymbol{x}) \\
&= -\lim_{\varepsilon\downarrow 0}\varphi\left(\frac{b_i}{\varepsilon\sqrt{\boldsymbol{x}^TC_i\boldsymbol{x}}}-\frac{\boldsymbol{\mu}_i^T\boldsymbol{x}}{\sqrt{\boldsymbol{x}^TC_i\boldsymbol{x}}}\right)\left\{\frac{(\boldsymbol{x}^TC_i\boldsymbol{x})\boldsymbol{\mu}_i-(\boldsymbol{\mu}_i^T\boldsymbol{x})C_i\boldsymbol{x}}{\varepsilon\sqrt{\boldsymbol{x}^TC_i\boldsymbol{x}}}+\frac{b_iC_i\boldsymbol{x}}{\varepsilon^2(\boldsymbol{x}^TC_i\boldsymbol{x})^{3/2}}\right\} \\
&= 0.
\end{aligned}$$

Hence $\lim_{\boldsymbol{x}\to 0}\nabla h_i(\boldsymbol{x}) = 0$ and $\nabla h_i(\boldsymbol{x})$ is continuous at $\boldsymbol{x}=0$. Therefore $h_1(\boldsymbol{x}),\ldots,h_r(\boldsymbol{x})$ are continuously differentiable on the open convex set $\operatorname{int}(E)$.

## 2.3 The main result

In what follows we make use of the following theorem from linear algebra:

**Theorem 2.3.1** (Simultaneous diagonalization of two matrices: Bellman [2])**.** *Given two real symmetric matrices, $A$ and $B$, with $A$ positive definite, there exists a nonsingular matrix $U$ such that*

$$U^TAU = I, \quad U^TBU = \operatorname{diag}(\lambda_1,\lambda_2,\ldots,\lambda_n) = \begin{pmatrix} \lambda_1 & & & O \\ & \lambda_2 & & \\ & & \ddots & \\ O & & & \lambda_n \end{pmatrix} \tag{2.3.1}$$

In the next theorem we present our main result.

**Theorem 2.3.2.** *Suppose $b_i > 0$ and $C_i$ is positive definite for $i = 1, \dots, r$. The functions $h_1(\boldsymbol{x}), \dots, h_r(\boldsymbol{x})$ defined by (2.2.6) (in this case, (2.2.9)) are uniformly quasi-concave on a convex set $E$ satisfying (2.2.7) if and only if each $C_i$ is a constant multiple of a covariance matrix $C$, and*

$$\frac{\boldsymbol{\mu}_1}{b_1} = \cdots = \frac{\boldsymbol{\mu}_r}{b_r}.$$

*Proof.* Sufficiency ($\Leftarrow$) is obvious, so we only show necessity ($\Rightarrow$). It is enough to show that $C_1, C_2$ are constant multiples of each other and that $\boldsymbol{\mu}_1/b_1 = \boldsymbol{\mu}_2/b_2$ for $r \geq 2$. $h_i(\boldsymbol{x})$ is continuously differentiable on the open convex set $\operatorname{int}(E)$. From (2.2.10) we have for $\boldsymbol{x} \neq 0$

$$\boldsymbol{x}^T \nabla h_i(\boldsymbol{x}) = -\varphi\left(\frac{b_i - \boldsymbol{\mu}_i^T \boldsymbol{x}}{\sqrt{\boldsymbol{x}^T C_i \boldsymbol{x}}}\right) \frac{b_i}{\sqrt{\boldsymbol{x}^T C_i \boldsymbol{x}}} < 0.$$

Thus $\nabla h_i(\boldsymbol{x}) \neq 0$ for $\boldsymbol{x} \neq 0$. We know $\lim_{\boldsymbol{x}\to 0} \nabla h_i(\boldsymbol{x}) = 0$. Let $E' := \operatorname{int}(E) \setminus \{0\}$. Then $E' = \{\, \boldsymbol{x} \in \operatorname{int}(E) \mid \nabla h_i(\boldsymbol{x}) \neq 0,\ i \in \{1, \dots, r\}\,\}$. From Lemma 2.2.1 and (2.2.10), there is a positive function $\alpha_{12}(\boldsymbol{x}) > 0$ such that for all $\boldsymbol{x} \in E'$ we have

$$(\boldsymbol{x}^T C_1 \boldsymbol{x})\boldsymbol{\mu}_1 + (b_1 - \boldsymbol{\mu}_1^T \boldsymbol{x}) C_1 \boldsymbol{x} = \alpha_{12}(\boldsymbol{x}) \left\{ (\boldsymbol{x}^T C_2 \boldsymbol{x})\boldsymbol{\mu}_2 + (b_2 - \boldsymbol{\mu}_2^T \boldsymbol{x}) C_2 \boldsymbol{x} \right\} \tag{2.3.2}$$

For small $\varepsilon > 0$ and $\boldsymbol{x} \in E'$, let us replace $\boldsymbol{x}$ with $\varepsilon\boldsymbol{x} \in E'$ in (2.3.2) and divide by $\varepsilon$ for both sides of the equation.

$$\varepsilon(\boldsymbol{x}^T C_1 \boldsymbol{x})\boldsymbol{\mu}_1 + (b_1 - \varepsilon\boldsymbol{\mu}_1^T \boldsymbol{x}) C_1 \boldsymbol{x} = \alpha_{12}(\varepsilon\boldsymbol{x}) \left\{ \varepsilon(\boldsymbol{x}^T C_2 \boldsymbol{x})\boldsymbol{\mu}_2 + (b_2 - \varepsilon\boldsymbol{\mu}_2^T \boldsymbol{x}) C_2 \boldsymbol{x} \right\} \tag{2.3.3}$$

Taking the limit of the both sides of (2.3.3) as $\varepsilon \to 0$ we obtain

$$b_1 C_1 \boldsymbol{x} = (\lim_{\varepsilon\to 0} \alpha_{12}(\varepsilon\boldsymbol{x})) b_2 C_2 \boldsymbol{x}. \tag{2.3.4}$$

Since $0 < \boldsymbol{x}^T C_1 \boldsymbol{x} < \infty,\ 0 < \boldsymbol{x}^T C_2 \boldsymbol{x} < \infty$ for $\boldsymbol{x} \in E'$, the limit

$$\lim_{\varepsilon\to 0} \alpha_{12}(\varepsilon\boldsymbol{x}) = \frac{b_1 \boldsymbol{x}^T C_1 \boldsymbol{x}}{b_2 \boldsymbol{x}^T C_2 \boldsymbol{x}} =: \alpha'_{12}(\boldsymbol{x}) \tag{2.3.5}$$

exists and $0 < \alpha'_{12}(\boldsymbol{x}) < \infty$. Thus we have

$$b_1 C_1 \boldsymbol{x} = \alpha'_{12}(\boldsymbol{x}) b_2 C_2 \boldsymbol{x} \quad \text{for } \boldsymbol{x} \in E'. \tag{2.3.6}$$

Since $C_1$ and $C_2$ are symmetric and $C_2$ is positive definite, from Theorem 2.3.1 there is a nonsingular matrix $U$ such that

$$U^T C_1 U = D, \quad U^T C_2 U = I,$$

where $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_r)$ is a diagonal matrix. Let $\boldsymbol{y} := U^{-1}\boldsymbol{x}$ and $F := \left\{ U^{-1}\boldsymbol{x} \;\middle|\; \boldsymbol{x} \in E' \right\}$. Since $U$ is nonsingular, $F$ is a neighborhood of the origin 0, and $0 \notin F$.

For all $\boldsymbol{y} \in F$ we have by multiplying $U^T$ from left to (2.3.6)

$$\begin{aligned} b_1 D\boldsymbol{y} &= \alpha'_{12}(U\boldsymbol{y}) b_2 \boldsymbol{y} \\ \Rightarrow b_1 \begin{bmatrix} \lambda_1 y_1 \\ \vdots \\ \lambda_r y_r \end{bmatrix} &= \alpha'_{12}(Uy) b_2 \begin{bmatrix} y_1 \\ \vdots \\ y_r \end{bmatrix} \end{aligned}$$

which implies that

$$0 < \alpha'_{12}(\boldsymbol{x}) = \alpha'_{12}(U\boldsymbol{y}) = \frac{b_1\lambda_1}{b_2} = \cdots = \frac{b_1\lambda_r}{b_2} =: \alpha'_{12}$$

is constant. Therefore we have from (2.3.6)

$$C_1 = \alpha'_{12}\frac{b_2}{b_1} C_2. \tag{2.3.7}$$

Let us plug (2.3.7) into (2.3.2).

$$\begin{aligned} &\boldsymbol{x}^T C_2 \boldsymbol{x} \left( \alpha'_{12} b_2 \boldsymbol{\mu}_1 - \alpha_{12}(\boldsymbol{x}) b_1 \boldsymbol{\mu}_2 \right) \\ &+ \left\{ (\alpha'_{12} - \alpha_{12}(\boldsymbol{x})) b_1 b_2 - \left( \alpha'_{12} b_2 \boldsymbol{\mu}_1 - \alpha_{12}(\boldsymbol{x}) b_1 \boldsymbol{\mu}_2 \right)^T \boldsymbol{x} \right\} C_2 \boldsymbol{x} = 0. \end{aligned} \tag{2.3.8}$$

Multiplying (2.3.8) by $\boldsymbol{x}^T$ from left we obtain

$$\left\{ \alpha'_{12} - \alpha_{12}(\boldsymbol{x}) \right\} b_1 b_2 \boldsymbol{x}^T C_2 \boldsymbol{x} = 0 \;\Rightarrow\; \alpha_{12}(\boldsymbol{x}) = \alpha'_{12}. \tag{2.3.9}$$

If we substitute (2.3.9) into (2.3.8), we get

$$\boldsymbol{x}^T C_2 \boldsymbol{x} (b_2 \boldsymbol{\mu}_1 - b_1 \boldsymbol{\mu}_2) = \boldsymbol{x}^T (b_2 \boldsymbol{\mu}_1 - b_1 \boldsymbol{\mu}_2) C_2 \boldsymbol{x}. \tag{2.3.10}$$

Let us introduce $\boldsymbol{w} := U^T(b_2\boldsymbol{\mu}_1 - b_1\boldsymbol{\mu}_2)$. Since $\boldsymbol{x} = U\boldsymbol{y}$ we have

$$\begin{aligned} (\boldsymbol{y}^T \boldsymbol{y}) w &= (\boldsymbol{y}^T \boldsymbol{w}) \boldsymbol{y} \\ \Rightarrow \begin{bmatrix} w_1 \\ \vdots \\ w_r \end{bmatrix} &= \frac{y_1 w_1 + \cdots + y_r w_r}{y_1^2 + \cdots + y_r^2} \begin{bmatrix} y_1 \\ \vdots \\ y_r \end{bmatrix} \end{aligned} \tag{2.3.11}$$

Since (2.3.11) holds for $\boldsymbol{y} = \varepsilon[0, 1, \ldots, 1]^T, \ldots, \boldsymbol{y} = \varepsilon[1, \ldots, 1, 0]^T \in F$ for some small $\varepsilon > 0$, it follows that

$$w_1 = 0, \ldots, w_r = 0 \;\Rightarrow \boldsymbol{w} = 0 \;\Rightarrow\; \frac{\boldsymbol{\mu}_1}{b_1} = \frac{\boldsymbol{\mu}_2}{b_2}. \tag{2.3.12}$$

□

## 2.4 Application in portfolio optimization

In this section we look at a probabilistic constrained stochastic programming problem, where the probabilistic constraint is of type (2.1.2). We assume that $T$ has independent, normally distributed rows and the factors in the product $\prod_{k=1}^{K} \mathbb{P}(T_k x \leq b_k)$ are uniformly quasi concave. The problem is special, but still can be applied, e.g., in portfolio optimization.

Consider $n$ assets and $K$ consecutive periods. Let us introduce the following notations: for $k = 1, \ldots, K$

$$\begin{aligned} T_k &: \text{random loss during the } k\text{-th period} \\ \boldsymbol{\mu}_k = \mathbb{E}[T_k^T] &: \text{expected loss} \\ C_k = \mathbb{E}[(T_k^T - \boldsymbol{\mu}_k)(T_k^T - \boldsymbol{\mu}_k)^T] &: \text{covariance matrix of } T_k\ . \end{aligned}$$

We assume that $T_k, k = 1, \ldots, K$ are independent and normally distributed random vectors and $\boldsymbol{\mu}_k \leq 0,\ k = 1, \ldots, K$. We also assume that the time window of the $K$ periods is relatively short and a linear trend for the expectations prevails. Formally, our assumptions are:

$$\boldsymbol{\mu}_1 = \boldsymbol{\mu} \ \text{ and } \ \boldsymbol{\mu}_{k+1} = \alpha \boldsymbol{\mu}_k, \quad k = 1, \ldots, K-1 \tag{2.4.1}$$

$$C_1 = C \ \text{ and } \ C_{k+1} = \alpha^2 C_k, \quad k = 1, \ldots, K-1. \tag{2.4.2}$$

For the first period, we consider the portfolio optimization problem formulated by

Kataoka [40]:

$$
\begin{aligned}
\text{(Problem 1):}\quad & \text{minimize}\quad b \\
& \text{subject to}\ \Phi\left(\frac{b-\boldsymbol{\mu}^T\boldsymbol{x}}{\sqrt{\boldsymbol{x}^T C\boldsymbol{x}}}\right) \geq p \\
& \qquad\qquad\quad \sum_{j=1}^{n} x_j = 1 \\
& \qquad\qquad\quad x_j \geq 0 \quad \text{for } j = 1,\ldots,n.
\end{aligned}
$$

For the $k$-th period ($k \in \{2,\ldots,K\}$), we consider the following problem.

$$
\begin{aligned}
\text{(Problem } k\text{):}\quad & \text{minimize}\quad b_1 \\
& \text{subject to}\ \prod_{i=1}^{k}\Phi\left(\frac{b_i-\boldsymbol{\mu}_i^T\boldsymbol{x}}{\sqrt{\boldsymbol{x}^T C_i\boldsymbol{x}}}\right) \geq p \\
& \qquad\qquad\quad \sum_{j=1}^{n} x_j = 1 \\
& \qquad\qquad\quad b_{i+1} = \alpha\, b_i \quad \text{for } i = 1,\ldots,k-1 \\
& \qquad\qquad\quad x_j \geq 0 \quad \text{for } j = 1,\ldots,n \\
& \qquad\qquad\quad b_1 \geq 0\ .
\end{aligned}
$$

A related model is presented in Yoda and Prékopa [78], where individual probabilistic constraints are taken for more than one part of the distribution.

By Theorem 2.3.2 the functions $h_1(\boldsymbol{x}),\ldots,h_K(\boldsymbol{x})$ defined by (2.2.9) are uniformly quasi-concave on the convex set

$$E := \cap_{k=1}^{K}\{\,\boldsymbol{x}\mid h_k(\boldsymbol{x}) \geq 1/2\,\} = \cap_{k=1}^{K}\left\{\,\boldsymbol{x}\;\middle|\; b_k \geq \boldsymbol{\mu}_k^T\boldsymbol{x}\,\right\} = \left\{\,\boldsymbol{x}\;\middle|\; b_K \geq \boldsymbol{\mu}_K^T\boldsymbol{x}\,\right\},$$

and hence $h(\boldsymbol{x}) = \prod_{k=1}^{K} h_k(\boldsymbol{x})$ is quasi-concave on $E$. Since the set

$$\{\,\boldsymbol{x}\mid h(\boldsymbol{x}) \geq p,\ \boldsymbol{x}\in E\,\}$$

is convex, the set of feasible solutions of (Problem $k$) is convex.

Below we present a numerical example for the application of the above model. We take the initial expectations and covariance matrix from past history data but then proceed to obtain those values in accordance with the assumption formulated in the model.

**Numerical Example.**

Assets “Dow, S&P500, Nasdaq, NYSECI, 10YrBond” are obtained from Yahoo! Finance (`http://finance.yahoo.com`) and assets “Oil, Gold, Silver, EUR/USD” are obtained from Dukascopy (`http://www.dukascopy.com`). We consider the expected values and the covariance matrix of the daily losses of the nine assets in May 2009. The data is shown in Table 2.1 and Table 2.2.

We assume that in the consecutive periods the expected returns are increased by 1% ($\alpha = 1.01$) and the covariances are increased by $\alpha^2 = (1.01)^2$. The values of the nine assets obtained by the use of (Problem $k$), $k = 1, \ldots, 5$ are given in Table 2.3.

Table 2.1: Expected losses in May 2009.

| Gold | Silver | Nasdaq | S&P500 | Oil | EUR/USD | 10YrBond | Dow | NYSECI |
|---|---|---|---|---|---|---|---|---|
| -1.253 | -3.008 | -0.149 | -0.711 | -1.379 | -0.82 | -1.052 | -0.546 | -1.069 |

Table 2.2: Covariance Matrix in May 2009.

| | Gold | Silver | Nasdaq | S&P500 | Oil | EUR/USD | 10YrBond | Dow | NYSECI |
|---|---|---|---|---|---|---|---|---|---|
| Gold | 5.159 | 7.228 | -1.437 | 1.492 | 3.989 | 2.764 | -5.25 | 1.198 | 2.231 |
| Silver | 7.228 | 19.441 | -0.785 | 6.454 | 10.143 | 4.94 | -5.198 | 5.343 | 9.061 |
| Nasdaq | -1.437 | -0.785 | 15.084 | 11.202 | 1.562 | 0.974 | -0.767 | 9.754 | 12.424 |
| S&P500 | 1.492 | 6.454 | 11.202 | 16.238 | 10.709 | 4.223 | -4.735 | 14.794 | 20.058 |
| Oil | 3.989 | 10.143 | 1.562 | 10.709 | 21.249 | 4.087 | -5.719 | 10.043 | 15.451 |
| EUR/USD | 2.764 | 4.94 | 0.974 | 4.223 | 4.087 | 4.375 | -3.255 | 3.764 | 5.996 |
| 10YrBond | -5.25 | -5.198 | -0.767 | -4.735 | -5.719 | -3.255 | 38.003 | -4.564 | -4.928 |
| Dow | 1.198 | 5.343 | 9.754 | 14.794 | 10.043 | 3.764 | -4.564 | 13.981 | 18.446 |
| NYSECI | 2.231 | 9.061 | 12.424 | 20.058 | 15.451 | 5.996 | -4.928 | 18.446 | 25.706 |

Table 2.3: Values of nine assets, May 2009.

| | Gold | Silver | Nasdaq | S&P500 | Oil | EUR/USD | 10YrBond | Dow | NYSECI |
|---|---|---|---|---|---|---|---|---|---|
| (Problem 1) | 0.5246 | 0.0422 | 0.1267 | 0 | 0.0102 | 0.1433 | 0.1531 | 0 | 0 |
| (Problem 2) | 0.5350 | 0 | 0.1342 | 0 | 0.0103 | 0.1718 | 0.1487 | 0 | 0 |
| (Problem 3) | 0.5266 | 0 | 0.1379 | 0 | 0.0076 | 0.1804 | 0.1475 | 0 | 0 |
| (Problem 4) | 0.5218 | 0 | 0.1399 | 0 | 0.0063 | 0.1849 | 0.1471 | 0 | 0 |
| (Problem 5) | 0.5188 | 0 | 0.1414 | 0 | 0.0053 | 0.1878 | 0.1467 | 0 | 0 |

# Chapter 3

# Convexity and solutions of the stochastic multidimensional knapsack problem with probabilistic constraints

## 3.1 Introduction.

The knapsack problem is one of the most fundamental combinatorial optimization problems with a wealth of applications in industries. The most basic form, the 0-1 one-dimensional single knapsack problem, can be stated as follows: Given a set of items, each with a value and a size, determine a subset maximizing the total value while keeping the total size within a given capacity. The problem is $\mathcal{NP}$-complete to solve exactly, although there is a pseudo-polynomial time algorithm and a fully polynomial-time approximation scheme (FPTAS) (see, e.g., Korte and Vygen [43]). One generalization of the problem includes the multidimensional knapsack problem (or multiply constrained knapsack problem), where each item has multiple attributes (sizes), such as length, volume, weight, etc., and a knapsack has a capacity for each attribute. This variant with a fixed dimension ($\geq 2$) was shown to be $\mathcal{NP}$-complete, and more strongly, has no FPTAS unless $\mathcal{P} = \mathcal{NP}$ by Gens and Levner [29] and Korte and Schrader [42] (see also Kellerer et al. [41]). Another generalization includes the multiple knapsack problem, which is also $\mathcal{NP}$-complete.

In real-life problems we often have to deal with uncertainty. Parameters cannot be predicted exactly but rather estimated probabilistically. Therefore, it is sometimes more desirable to model these parameters with random variables. In this chapter, we study the multidimensional (single) knapsack problem where the item attributes are independent random variables. Under the assumption we need a new principle to formulate the problem and our choice is the probabilistic constrained formulation.

The knapsack problem has been studied for more than a century. A broad overview of the theoretical and the practical results can be found in Kellerer et al. [41]. The deterministic model for the multidimensional knapsack problem has been studied extensively since the 1950s (see Fréville [24] and Fréville and Hanafi [25] for a comprehensive survey). A few works are known for the use of the probabilistic constrained stochastic programming model for the knapsack problem with random item attributes, which can be considered as a special case for the general stochastic programming problem with probabilistic linear constraints. If the randomness is in the technology matrix, then the problem is typically nonconvex. There are some exceptions. If the random variables follow normal distributions, the probabilistic constraints can be rewritten as quadratic constraints for the random matrix with one row (see Kataoka [40], van de Panne and Popp [72], and Prékopa [58]). For the random matrix with more than one row, the first paper where convexity theorems are presented is by Prékopa [54] and an important progress was made by Henrion and Strugarek [36]. Recently, Zymler et al. [81] developed tractable semidefinite programming based approximations by using moment information of the distributions. The use of moment information is also found in Prékopa [59] and Mádi-Nagy and Prékopa [49]. A couple of works are known for the probabilistic constrained stochastic one-dimensional knapsack problem. Goyal and Ravi [32] showed a polynomial time approximation scheme via a parametric LP reformulation when the random item attributes are independent and normally distributed. Fortz and Poss [23] showed that the problem can be linearized when the random item attributes are independent and follow normal or gamma distributions under some regulatory condition. Our work can be considered as an extension of the latter work to the multidimensional knapsack problem. While we study the stochastic multidimensional knapsack problem, our convexity result holds for a broader class of stochastic combinatorial optimization problems whose underlying deterministic models are formulated by linear inequalities and 0-1 decision variables.

Applications of the multidimensional knapsack problem include, but are not limited to, cargo loading (see Bellman and Dreyfus [3]), cutting stock (see Gilmore and Gomory [31]), capital budgeting (see Lorie and Savage [47] and Weingartner [75]), project selection (see Petersen [50]), resource allocation in distributed data processing (see Gavish and Pirkul [28]), computer systems design (see Ferreira et al. [22]), daily management of a satellite (see Vasquez and Hao [73]), and combinatorial auctions (see de Vries and Vohra [16] and Rothkopf et al. [66]). See also a survey paper by Wilbaut et al. [76].

This chapter is organized as follows. In section 3.2, we formulate the probabilistic constrained model for the multidimensional knapsack problem and present our main theorem. In section 3.3, we prove the theorem showing convexity of a relaxed feasible set of the problem for four distributions. Section 3.4 illustrates computational experiments.

## 3.2 Formulation of the problem.

First let us consider the deterministic problem. We are given a set of $n$ items with values $v_1, v_2, \dots, v_n$. Each item has $m$ attributes such as weight, time, budget, etc. and item $j$ consumes $w_{ij} > 0$ units of resource for attribute $i$. We have a single knapsack with $m$ capacities $W_i > 0$ , $i = 1, \dots, m$ for the $m$ attributes, respectively. The goal is to select a subset of items (to be placed into the knapsack) maximizing the total value while keeping the capacities, which can be formulated as follows:

$$\text{maximize} \sum_{j=1}^{n} v_j x_j \tag{3.2.1a}$$

$$\text{subject to} \sum_{j=1}^{n} w_{ij} x_j \le W_i \quad \text{for } i = 1, \dots, m \tag{3.2.1b}$$

$$x_j \in \{0, 1\} \quad \text{for } j = 1, \dots, n. \tag{3.2.1c}$$

Now suppose the attributes are random variables denoted by $\xi_{ij}$'s (in place of $w_{ij}$'s), and formulate the problem as a probabilistic constrained stochastic programming, where constraints (3.2.1b) are replaced by the following joint probabilistic constraint:

$$\mathbb{P}\left( \sum_{j=1}^{n} \xi_{ij} x_j \le W_i \quad \text{for } i = 1, \dots, m \right) \ge q. \tag{3.2.2}$$

Here $q \in (0, 1)$ is a fixed probability level, e.g., $0.9, 0.95, 0.99$.

Assuming the random variables $\xi_{ij}$'s are independent, the joint probabilistic constraint (3.2.2) can be written as follows:

$$\prod_{i=1}^{m} \mathbb{P}\left(\sum_{j=1}^{n} \xi_{ij} x_j \leq W_i\right) \geq q. \tag{3.2.3}$$

Here we only need independence of the random vectors $(\xi_{i1}, \xi_{i2}, \ldots, \xi_{in})$, $i = 1, \ldots, m$, that is, random variables of different attributes in the same item or in different items are independent but those of the same attribute in different items may be dependent.

Here let us mention a property of random variables as follows:

**Definition 3.2.1** (Associated random variables: Esary et al. [20])**.** *We say random variables $\xi_1, \ldots, \xi_n$ are* associated *if*

$$\mathrm{Cov}[f(\boldsymbol{\xi}), g(\boldsymbol{\xi})] \geq 0$$

*for all nondecreasing functions $f$ and $g$ for which $\mathbb{E}[f(\boldsymbol{\xi})]$, $\mathbb{E}[g(\boldsymbol{\xi})]$, and $\mathbb{E}[f(\boldsymbol{\xi})g(\boldsymbol{\xi})]$ exist, where $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_n)$.*

To give an example, consider $\xi_i := \eta_0 + \eta_i,\ i = 1, \ldots, n$ where $\eta_0, \eta_1, \ldots, \eta_n$ are independent random variables. The random variables $\xi_1, \ldots, \xi_n$ are associated by Theorem 2.1 (Independent random variables are associated) and Property $\mathrm{P}_4$ (Nondecreasing functions of associated random variables are associated) in Esary et al. [20].

Note that even if we can't assume the random variables are independent, which is often the case in real-life applications, if we can assume they are associated, we have the following inequality (see Esary et al. [20]):

$$\mathbb{P}\left(\sum_{j=1}^{n} \xi_{ij} x_j \leq W_i \ \text{ for } i = 1, \ldots, m\right) \geq \prod_{i=1}^{m} \mathbb{P}\left(\sum_{j=1}^{n} \xi_{ij} x_j \leq W_i\right).$$

Then the probabilistic constraint (3.2.3) ensures the joint probabilistic constraint (3.2.2). The probabilistic constrained stochastic programming model for the multidimensional knapsack problem with independent random item attributes is formulated as follows:

$$\text{maximize} \sum_{j=1}^{n} v_j x_j \tag{3.2.4a}$$

$$\text{subject to} \prod_{i=1}^{m} \mathbb{P}\left(\sum_{j=1}^{n} \xi_{ij} x_j \leq W_i\right) \geq q \tag{3.2.4b}$$

$$x_j \in \{0,1\} \quad \text{for } j = 1, \ldots, n. \tag{3.2.4c}$$

Let us denote $\boldsymbol{x} = (x_1, \ldots, x_n)$ and

$$F_i(\boldsymbol{x}) := \mathbb{P}\left(\sum_{j=1}^{n} \xi_{ij} x_j \leq W_i\right). \tag{3.2.5}$$

The feasible set of the problem is as follows:

$$\left\{ \boldsymbol{x} \in \mathbb{Z}^n \;\middle|\; \prod_{i=1}^{m} F_i(\boldsymbol{x}) \geq q,\ x_j \in \{0,1\} \quad \text{for } j = 1, \ldots, n \right\}.$$

By relaxing the integrality of $x_j$'s, we consider the following relaxed feasible set:

$$\mathcal{S} := \left\{ \boldsymbol{x} \in \mathbb{R}^n \;\middle|\; \prod_{i=1}^{m} F_i(\boldsymbol{x}) \geq q,\ x_j \in [0,1] \quad \text{for } j = 1, \ldots, n \right\}. \tag{3.2.6}$$

The set $\mathcal{S}$ is convex if $F_i(\boldsymbol{x})$ is logconcave for every $i \in \{1, \ldots, m\}$. The following remark, which is easy to prove, will be used to show the logconcavity of $F_i(\boldsymbol{x})$ for some special distributions in section 3.3.

**Remark 3.2.1.** *Let $h$ be a linear function. A composite function $f(\boldsymbol{x}) = g(h(\boldsymbol{x}))$ is a logconcave function of $\boldsymbol{x} \in \mathbb{R}^n$ if $g(t)$ is a logconcave function of $t \in \mathbb{R}$.*

In all of our cases studied in this chapter, $F_i(\boldsymbol{x})$ has the form

$$F_i(\boldsymbol{x}) = g_i(h_i(\boldsymbol{x}))$$

with $g_i(t)$ defined on $\mathbb{R}$ and $h_i(\boldsymbol{x}) = \boldsymbol{c}_i^T \boldsymbol{x}$ defined on $\mathbb{R}^n$, where $\boldsymbol{c}_i \in \mathbb{R}^n$ is a constant vector. So we only have to show that $g_i(t)$ is logconcave to ensure the logconcavity of $F_i(x)$ and hence the convexity of $\mathcal{S}$.

Here we need to define some notations to present our main theorem. We denote by $\varphi(t)$ and $\Phi(t)$, the p.d.f. and the c.d.f., respectively, of the standard normal distribution:

$$\varphi(t) := \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right), \quad \Phi(t) := \int_{-\infty}^{t} \varphi(u) du. \tag{3.2.7}$$

**Theorem 3.2.1.** *The relaxed feasible set $\mathcal{S}$ in (3.2.6) is convex if for each $i \in \{1, \ldots, m\}$ the independent random variables $\xi_{ij},\ j = 1, \ldots, n$ follow one of the following distributions:*

- *The gamma distribution $\xi_{ij} \sim \Gamma(p_{ij}, \theta_i)$ with shape $p_{ij} > 0$ and scale $\theta_i > 0$.*
- *The normal distribution $\xi_{ij} \sim \mathcal{N}(\mu_{ij}, \lambda_i \mu_{ij})$ with mean $\mu_{ij} > 0$ and variance $\lambda_i \mu_{ij} > 0$ satisfying*

$$\frac{4W_i}{\lambda_i} \geq -\gamma_i^2 + \left[ \frac{1 + \sqrt{1 + \gamma_i(\gamma_i + \varphi(\gamma_i)/\Phi(\gamma_i))}}{\gamma_i + \varphi(\gamma_i)/\Phi(\gamma_i)} \right]^2 \tag{3.2.8a}$$

*where*

$$\gamma_i := \frac{W_i - \nu_i}{\sqrt{\lambda_i \nu_i}}, \quad \nu_i := \sum_{k=1}^{n} \mu_{ik}. \tag{3.2.8b}$$

- *The Poisson distribution $\xi_{ij} \sim Pois(\lambda_{ij})$ with parameter $\lambda_{ij} > 0$.*
- *The binomial distribution $\xi_{ij} \sim B(n_{ij}, p_i)$ with the number of trials $n_{ij} \in \mathbb{N}$ and the success probability in each trial $p_i \in (0, 1)$.*

No special form of the parameters is necessary for the Poisson distribution. For the other three distributions, those special forms are encountered when, for instance, each random variable $\xi_{ij},\ j \in \{1, \ldots, n\}$ is the sum of varying number of independent and identically distributed (i.i.d.) random variables. (We need one more condition (3.2.8) for the normal distribution.) Recall the sum of i.i.d. gamma/normal/Bernoulli random variables is a gamma/normal/binomial random variable. For the case of the binomial distribution, we may consider a practical example where each item has varying number of identical components with the same failure rate and a random item attribute refers to the total number of failed components. We may consider similar examples for the gamma and the normal distributions.

## 3.3 Convexity of the stochastic multidimensional knapsack problem.

### 3.3.1 Convexity result for the gamma distribution.

In this section we consider a case where the independent random variable $\xi_{ij}$ in (3.2.4) has the gamma distribution

$$\xi_{ij} \sim \Gamma(p_{ij}, \theta_i)$$

with shape $p_{ij} > 0$ and scale $\theta_i > 0$ for some $i \in \{1, \ldots, m\}$ and all $j = 1, \ldots, n$. The probability density function (pdf) of $\xi_{ij}$ is as follows.

$$f(y) = \frac{1}{\Gamma(p_{ij})\theta_i} y^{p_{ij}-1} \exp\left(-\frac{y}{\theta_i}\right) \quad \text{for } y > 0,$$

where we defined the gamma function:

$$\Gamma(p) := \int_0^\infty t^{p-1} e^{-t} dt \;\text{ for } p > 0. \tag{3.3.1}$$

Note that $\mathrm{Var}(\xi_{ij}) = \theta_i \, \mathbb{E}(\xi_{ij}) = p_{ij}\theta_i^2$. Since $x_j \in \{0, 1\}$, it follows that for $\boldsymbol{x} \neq 0$, $\sum_{j=1}^n \xi_{ij} x_j$ is a sum of independent gamma random variables with the common scale $\theta_i$ and thus has the gamma distribution

$$\sum_{j=1}^n \xi_{ij} x_j \sim \Gamma\left(p_i(\boldsymbol{x}),\, \theta_i\right),$$

where we defined $p_i(\boldsymbol{x}) := \sum_{j=1}^n p_{ij} x_j$. Let $P(p, \lambda)$ denote the lower regularized gamma function:

$$P(p, \lambda) := \int_0^\lambda \frac{t^{p-1} e^{-t}}{\Gamma(p)} dt \;\text{ for } p \geq 0,\; \lambda \geq 0. \tag{3.3.2}$$

For any fixed $\lambda > 0$, we define $P(0, \lambda) = \lim_{p \to 0+} P(p, \lambda) = 1$. The function in the joint probabilistic constraint:

$$F_i(\boldsymbol{x}) = \mathbb{P}\left(\sum_{j=1}^n \xi_{ij} x_j \leq W_i\right) = P\left(p_i(\boldsymbol{x}), \frac{W_i}{\theta_i}\right) \tag{3.3.3}$$

is defined for $\boldsymbol{x} \geq 0$. The logconcavity of $P(p, \lambda)$ in the following lemma together with Remark 3.2.1 ensures the logconcavity of $F_i(\boldsymbol{x})$ in (3.3.3), which proves Theorem 3.2.1 for the case of the gamma distribution.

**Lemma 3.3.1.** *For any fixed $\lambda > 0$, the function $P(p, \lambda)$ defined by* (3.3.2) *is strictly decreasing and strictly logconcave for $p \geq 0$.*

*Proof.* First we prove the decreasing property. It follows by theorem 4.8.2 in Prékopa [58] that

$$1 - P(p, \lambda) = \int_{\lambda}^{\infty} \frac{t^{p-1} e^{-t}}{\Gamma(p)} dt$$

is strictly increasing for $p \geq 0$. Hence $P(p, \lambda)$ is strictly decreasing for $p \geq 0$.

Next we prove the logconcavity by showing the second derivative of $\ln F(k, \lambda)$ is negative. Simple calculation shows that for $p > 0$

$$\begin{aligned} \frac{d^2}{dp^2} \ln P(p, \lambda) = & \frac{\displaystyle\int_0^{\lambda} t^{p-1} (\ln t)^2 e^{-t} dt}{\displaystyle\int_0^{\lambda} t^{p-1} e^{-t} dt} - \left( \frac{\displaystyle\int_0^{\lambda} t^{p-1} (\ln t) e^{-t} dt}{\displaystyle\int_0^{\lambda} t^{p-1} e^{-t} dt} \right)^2 \\ & - \left[ \frac{\displaystyle\int_0^{\infty} t^{p-1} (\ln t)^2 e^{-t} dt}{\displaystyle\int_0^{\infty} t^{p-1} e^{-t} dt} - \left( \frac{\displaystyle\int_0^{\infty} t^{p-1} (\ln t) e^{-t} dt}{\displaystyle\int_0^{\infty} t^{p-1} e^{-t} dt} \right)^2 \right]. \end{aligned} \tag{3.3.4}$$

Let us introduce a random variable $X$ that has the following continuous and strictly logconcave p.d.f.:

$$g(x) := \frac{e^{px} e^{-e^x}}{\Gamma(p)} \quad \text{for } x \in \mathbb{R},$$

where $p > 0$ is now a constant. Note that

$$g(x) > 0, \quad \int_{-\infty}^{\infty} g(x) dx = \int_0^{\infty} \frac{t^{p-1} e^{-t}}{\Gamma(p)} dt = 1, \quad \frac{d^2 \ln g(x)}{dx^2} = -e^x < 0.$$

The second derivative (3.3.4) can be written as

$$\begin{aligned} & \frac{\displaystyle\int_{-\infty}^{\ln \lambda} x^2 g(x) dx}{\displaystyle\int_{-\infty}^{\ln \lambda} g(x) dx} - \left( \frac{\displaystyle\int_{-\infty}^{\ln \lambda} x g(x) dx}{\displaystyle\int_{-\infty}^{\ln \lambda} g(x) dx} \right)^2 - \left[ \frac{\displaystyle\int_{-\infty}^{\infty} x^2 g(x) dx}{\displaystyle\int_{-\infty}^{\infty} g(x) dx} - \left( \frac{\displaystyle\int_{-\infty}^{\infty} x g(x) dx}{\displaystyle\int_{-\infty}^{\infty} g(x) dx} \right)^2 \right] \\ & = \mathbb{E}(X^2 | X \leq \ln \lambda) - \mathbb{E}^2(X | X \leq \ln \lambda) - \left[ (\mathbb{E}(X^2) - \mathbb{E}^2(X) \right] \\ & = \mathbb{E}(X^2 | X \leq v) - \mathbb{E}^2(X | X \leq v) - \left[ \mathbb{E}(X^2) - \mathbb{E}^2(X) \right]. \quad \text{(We denote } v := \ln \lambda) \end{aligned}$$

Since the random variable $X$ has a continuous and strictly logconcave p.d.f. $g(x)$, it follows by theorem 2.2 and 2.3 in Prékopa [61] that

$$\mathbb{E}(X^2|X \leq v) - \mathbb{E}^2(X|X \leq v)$$

is strictly increasing in $v$. We have proved that for any $v \in \mathbb{R}$, the following inequality holds:

$$\mathbb{E}(X^2|X \leq v) - \mathbb{E}^2(X|X \leq v) < \mathbb{E}(X^2) - \mathbb{E}^2(X),$$

which implies $d^2 \ln P(p, \lambda)/dp^2 < 0$. Hence $P(p, \lambda)$ is strictly logconcave for $p \geq 0$. $\square$

### 3.3.2 Convexity result for the normal distribution.

In this section we consider a case where the independent random variable $\xi_{ij}$ in (3.2.4) has the normal distribution

$$\xi_{ij} \sim \mathcal{N}(\mu_{ij},\, \lambda_i\, \mu_{ij})$$

with mean $\mu_{ij} > 0$ and variance $\lambda_i \mu_{ij} > 0$ satisfying (3.2.8) for some $i \in \{1, \ldots, m\}$ and all $j = 1, \ldots, n$. The pdf of $\xi_{ij}$ is as follows.

$$f(y) = \frac{1}{\sqrt{2\pi\lambda_i\, \mu_{ij}}} \exp\left(-\frac{(y - \mu_{ij})^2}{2\lambda_i\, \mu_{ij}}\right) \quad \text{for } y \in \mathbb{R}$$

Note that $\mathrm{Var}(\xi_{ij}) = \lambda_i\, \mathbb{E}(\xi_{ij}) = \lambda_i \mu_{ij}$. The condition (3.2.8) is satisfied if $\nu_i$ is smaller than the threshold determined by $W_i$ and $\lambda_i$ (see (3.3.17)). Since $x_j \in \{0, 1\}$, it follows that for $\boldsymbol{x} \neq 0$, $\sum_{j=1}^n \xi_{ij} x_j$ is a sum of independent normal random variables and thus has the normal distribution

$$\begin{aligned}\sum_{j=1}^n \xi_{ij} x_j \sim \mathcal{N}&\left(\sum_{j=1}^n \mu_{ij} x_j,\, \sum_{j=1}^n \lambda_i \mu_{ij} x_j^2\right)\\ &= \mathcal{N}\left(\mu_i(\boldsymbol{x}),\, \lambda_i \mu_i(\boldsymbol{x})\right), \quad (\because x_j^2 = x_j)\end{aligned}$$

where we defined $\mu_i(\boldsymbol{x}) := \sum_{j=1}^n \mu_{ij} x_j$. Let us introduce a function:

$$G_i(\mu) := \Phi\left(\frac{W_i - \mu}{\sqrt{\lambda_i \mu}}\right) \quad \text{for } \mu \geq 0. \tag{3.3.5}$$

We define $G_i(0) = \lim_{\mu\to 0+} G_i(\mu) = 1$. The function in the joint probabilistic constraint:

$$F_i(\boldsymbol{x}) = \mathbb{P}\left(\sum_{j=1}^{n} \xi_{ij}x_j \le W_i\right) = G_i(\mu_i(\boldsymbol{x})) \tag{3.3.6}$$

is defined for $\boldsymbol{x} \ge 0$.

The following lemma about $\varphi(x)/\Phi(x)$ is required to prove the logconcavity of $G_i(\mu)$. Most of (i) and (ii) are well-known, but we present them for completeness.

**Lemma 3.3.2.** *Let us denote $\rho(x) := \varphi(x)/\Phi(x)$. For $x \in \mathbb{R}$, we have the following:*

*(i)* $\rho(x)$ *is positive, strictly decreasing, strictly logconcave, and strictly convex.*

$\lim_{x\to-\infty} \rho(x) = \infty$. $\lim_{x\to\infty} \rho(x) = 0$.

*(ii)* $x + \rho(x)$ *is positive and strictly increasing.*

$\lim_{x\to-\infty} (x + \rho(x)) = 0$. $\lim_{x\to\infty} (x + \rho(x)) = \infty$.

*(iii)* $\rho(x)(x + \rho(x)) \in (0, 1)$ *is strictly decreasing.*

$\lim_{x\to-\infty} \rho(x)(x + \rho(x)) = 1$. $\lim_{x\to\infty} \rho(x)(x + \rho(x)) = 0$.

*(iv)* $1/(x + \rho(x))$ *is positive, strictly decreasing, and strictly convex.*

*Proof.* (i) Clearly $\rho(x) = \varphi(x)/\Phi(x) > 0$. We have $\rho'(x) = (\ln \Phi(x))'' < 0$ since $\Phi(x)$ is strictly logconcave. We prove in (ii) the strict logconcavity of $\rho(x)$, which is equivalent to the strict increasing property of $x+\rho(x)$ because $(\ln \rho(x))' = -(x+\rho(x))$. We prove in (iii) the strict convexity of $\rho(x)$, which is equivalent to the strict decreasing property of $\rho(x)(x + \rho(x))$ because $\rho'(x) = -\rho(x)(x + \rho(x))$. The proof of the limits are easily derived by l'Hôpital's rule.

(ii) We can express $x + \rho(x) = -(\ln \Phi(x))''/\rho(x) > 0$ since $\Phi(x)$ is strictly logconcave. Alternatively, we can express

$$x + \rho(x) = \frac{x\Phi(x) + \varphi(x)}{\Phi(x)} = \frac{\int_{-\infty}^{x}(t\Phi(t) + \varphi(t))'dt}{\Phi(x)} = \frac{\int_{-\infty}^{x}\Phi(t)dt}{\Phi(x)} = \frac{1}{\left(\ln \int_{-\infty}^{x}\Phi(t)dt\right)'}. \tag{3.3.7}$$

Since $\int_{-\infty}^{x} \Phi(t)dt$ is an integral of a strictly logconcave distribution function, it is strictly logconcave. Hence the denominator of (3.3.7) is a strictly decreasing function. The proof of the limits are easily derived by l'Hôpital's rule.

(iii) First we prove

$$0 < \rho(x)(x + \rho(x)) < 1. \tag{3.3.8}$$

The left inequality is obvious from the positivities in (i) and (ii). Since $x + \rho(x)$ is strictly increasing by (ii), it follows that $0 < (x + \rho(x))' = 1 - \rho(x)(x + \rho(x))$, which proves the right inequality. Then the proof of the limits are easily derived by l'Hôpital's rule. To prove the strict decreasing property we show the following inequality:

$$\begin{aligned} f(x) &:= -\frac{1}{\rho(x)}(\rho(x)(x + \rho(x)))' \\ &= (x + \rho(x))^2 + \rho(x)(x + \rho(x)) - 1 > 0 \quad \text{for } x \in \mathbb{R}. \end{aligned} \tag{3.3.9}$$

Either $f(x) > 0$ for all $x \in \mathbb{R}$ or there exists $x_0 \in \mathbb{R}$ such that $f(x_0) \leq 0$. We show the latter case doesn't hold by contradiction. Assume it holds. By the definition of $f(x)$ in (3.3.9) we have the following implication:

$$\begin{aligned} & f'(x) = 2(x + \rho(x))(1 - \rho(x)(x + \rho(x))) - \rho(x)((x + \rho(x))^2 + \rho(x)(x + \rho(x)) - 1) \leq 0 \\ \Rightarrow\ & f(x) \geq \frac{2}{\rho(x)}(x + \rho(x))(1 - \rho(x)(x + \rho(x))) > 0 \quad (\because \text{positivities in (i) and (ii), (3.3.8)}), \end{aligned} \tag{3.3.10}$$

which implies that $f'(x_0) > 0$. Since $f(x)$ is continuously differentiable, there exists $x_1 < x_0$ such that $f(x_1) < f(x_0) \leq 0$. Since $\lim_{x \to -\infty} f(x) = 0$ and $f(x_1) < 0$, there exists $x_2 \in (-\infty, x_1)$ such that $f(x_1) < f(x_2) < 0$ and $f'(x_2) < 0$, which contradicts (3.3.10).

(iv) The positivity and the decreasing property follows directly from (ii). To prove the strict convexity we show the following inequality:

$$\begin{aligned} g(x) &:= (x + \rho(x))^3 \left( \frac{1}{x + \rho(x)} \right)'' \\ &= 2(1 - \rho(x)(x + \rho(x)))^2 - \rho(x)(x + \rho(x))f(x) \\ &= 2(1 - \rho(x)\eta(x))^2 - \rho(x)\eta(x)f(x) \ > 0 \quad \text{for } x \in \mathbb{R}, \end{aligned} \tag{3.3.11}$$

where we defined $\eta(x) := x + \rho(x)$ and used $f(x) = \eta(x)^2 + \rho(x)\eta(x) - 1$ defined in (3.3.9). Either $g(x) > 0$ for all $x \in \mathbb{R}$ or there exists $x_0 \in \mathbb{R}$ such that $g(x_0) \leq 0$. We show the latter case doesn't hold by contradiction. Assume it holds. By the definition of $g(x)$ in (3.3.11) we have the following implication:

$$\begin{aligned} g' &= \rho\left[4(1-\rho\eta)f + f^2 - 2\eta^2(1-\rho\eta) + \rho\eta f\right] \leq 0 \\ \Rightarrow\ g &\geq 2(1-\rho\eta)^2 + 4(1-\rho\eta)f - 2\eta^2(1-\rho\eta) + f^2 \\ &= 2(1-\rho\eta)f + f^2 > 0 \quad (\because (3.3.8), (3.3.9)), \end{aligned} \tag{3.3.12}$$

which implies that $g'(x_0) > 0$. Since $g(x)$ is continuously differentiable, there exists $x_1 < x_0$ such that $g(x_1) < g(x_0) \leq 0$. Since $\lim_{x\to-\infty} g(x) = 0$ and $g(x_1) < 0$, there exists $x_2 \in (-\infty, x_1)$ such that $g(x_1) < g(x_2) < 0$ and $g'(x_2) < 0$, which contradicts (3.3.12). □

The logconcavity of $G_i(\mu)$ in the following lemma together with Remark 3.2.1 ensures the logconcavity of $F_i(\boldsymbol{x})$ in (3.3.6), which proves Theorem 3.2.1 for the case of the normal distribution.

**Lemma 3.3.3.** *The function $G_i(\mu)$ defined by* (3.3.5) *is strictly decreasing for $\mu \geq 0$. It is logconcave for $\mu \in [0, \nu_i]$ if and only if the condition* (3.2.8) *is satisfied.*

*Proof.* We prove the decreasing property for $x \geq 0$ and the logconcavity for $x \in [0, \nu]$ of the function:

$$F(x) := \Phi\left(\frac{a}{\sqrt{x}} - b\sqrt{x}\right) \quad \text{for } x \geq 0,$$

where $a, b, \nu > 0$ for simplicity of the notation. We define $F(0) = \lim_{x\to 0+} F(x) = 1$. Note that $G_i(\mu) = F(\mu)$ with $a = W_i/\sqrt{\lambda_i}$, $b = 1/\sqrt{\lambda_i}$, and $\nu = \nu_i$. Let us denote

$$g(x; a, b) := ax^{-1/2} - bx^{1/2} \quad \text{and} \quad \rho(z) := \frac{\varphi(z)}{\Phi(z)}. \tag{3.3.13}$$

First we prove the decreasing property of $F(x) = \Phi(g(x; a, b))$. Clearly $\Phi(z)$ is strictly increasing for $z \in \mathbb{R}$. The following shows that $g(x; a, b)$ is strictly decreasing for $x > 0$.

$$g'(x; a, b) = -\frac{1}{2}(ax^{-\frac{3}{2}} + bx^{-\frac{1}{2}}) < 0 \quad \text{for } x > 0.$$

It follows that $\Phi(g(x;a,b))$ is strictly decreasing for $x \geq 0$.

Next we prove the logconcavity. The function $F(x) = \Phi(g(x;a,b))$ is logconcave for $x \in [0,\nu]$ if and only if for all $x \in (0,\nu]$,

$$0 \geq (\ln F(x))'' = \left(\frac{\varphi(g)g'}{\Phi(g)}\right)' = \big(\rho(g)g'\big)' = \rho(g)\left[g'' - (g+\rho(g))(g')^2\right]$$
$$\Leftrightarrow g + \rho(g) \geq \frac{g''}{(g')^2} = \frac{x^{\frac{1}{2}}(3a+bx)}{(a+bx)^2}. \tag{3.3.14}$$

By solving the quadratic equation

$$g = ax^{-1/2} - bx^{1/2} \text{ (in (3.3.13))} \;\Leftrightarrow\; b(x^{1/2})^2 + gx^{1/2} - a = 0$$

with respect to $x^{1/2} > 0$, we have

$$x^{\frac{1}{2}} = \frac{\sqrt{g^2+4ab} - g}{2b}. \tag{3.3.15}$$

By using (3.3.15), the condition (3.3.14) for $a$ and $b$ is equivalent to

$$\frac{2\sqrt{g^2+4ab} + g}{g^2+4ab} \leq g + \rho(g) \quad \text{for all } x \in (0,\nu].$$

This is equivalent to

$$\sqrt{g^2+4ab} \leq \frac{1 - \sqrt{1 + g(g+\rho(g))}}{g+\rho(g)} \tag{3.3.16a}$$

or

$$\sqrt{g^2+4ab} \geq \frac{1 + \sqrt{1 + g(g+\rho(g))}}{g+\rho(g)} \tag{3.3.16b}$$

for all $x \in (0,\nu]$. Note that $1 + g(g+\rho(g)) = (g+\rho(g))^2 + 1 - \rho(g)(g+\rho(g)) > 0$ by Lemma 3.3.2 (iii). No $a,b$ exists under (3.3.16a), for instance when $g(g+\rho(g)) > 0$, which holds for sufficiently small $x > 0$. Since both sides of (3.3.16b) are positive, the condition (3.3.16b) can be expressed as

$$4ab \geq -g^2 + \left[\frac{1}{g+\rho(g)} + \sqrt{h(g)}\right]^2 = -g^2 + \frac{1}{(g+\rho(g))^2} + \frac{2\sqrt{h(g)}}{g+\rho(g)} + h(g). \tag{3.3.17}$$

Here we defined

$$h(z) := \frac{1 + z(z+\rho(z))}{(z+\rho(z))^2} = 1 + \frac{1 - \rho(z)(z+\rho(z))}{(z+\rho(z))^2} \quad (> 0 \text{ by (3.3.8) in Lemma 3.3.2 (iii)})$$
$$= 1 - \left(\frac{1}{z+\rho(z)}\right)',$$

which is decreasing for $z \in \mathbb{R}$ by Lemma 3.3.2 (iv). The sum $-g^2 + 1/(g+\rho(g))^2$ in (3.3.17) is decreasing in $g$ because

$$\begin{aligned} -\frac{(z+\rho(z))^3}{2}\left(-z^2 + \frac{1}{(z+\rho(z))^2}\right)' &= (z+\rho(z))^4 - \rho(z)(z+\rho(z))^3 + \{1-\rho(z)(z+\rho(z))\} \\ &> (z+\rho(z))^4 - \{1-\rho(z)(z+\rho(z))\}^2 \\ &\quad (\because \text{(3.3.11) in Lemma 3.3.2 (iv)}) \\ &> 0. \quad (\because \text{(3.3.8),(3.3.9) in Lemma 3.3.2 (ii),(iii)}) \end{aligned}$$

By Lemma 3.3.2 (ii), $g + \rho(g)$ is increasing in $g$. Thus the right-hand side of (3.3.17) is decreasing for $g \in \mathbb{R}$ and hence it is increasing for $x > 0$, regardless of $a$ and $b$. So the condition (3.3.14) for $a$ and $b$ is equivalent to

$$4ab \geq -g(\nu; a, b)^2 + \left[\frac{1+\sqrt{1+g(\nu;a,b)\{g(\nu;a,b)+\rho(g(\nu;a,b))\}}}{g(\nu;a,b)+\rho(g(\nu;a,b))}\right]^2. \tag{3.3.18}$$

Therefore $F(x)$ is logconcave for $x \in [0, \nu]$ under the condition for $a$ and $b$ in (3.3.18). $\square$

### 3.3.3 Convexity result for the Poisson distribution.

In this section we consider a case where the independent random variable $\xi_{ij}$ in (3.2.4) has the Poisson distribution

$$\xi_{ij} \sim \mathrm{Pois}(\lambda_{ij}),$$

with parameter $\lambda_{ij} > 0$ for some $i \in \{1, \ldots, m\}$ and all $j = 1, \ldots, n$. The probability mass function (pmf) of $\xi_{ij}$ is as follows.

$$f(k) = \frac{\lambda_{ij}^k}{k!} e^{-\lambda_{ij}} \quad \text{for } k = 0, 1, \ldots$$

Note that $\mathrm{Var}(\xi_{ij}) = \mathbb{E}(\xi_{ij}) = \lambda_{ij}$. Since $x_j \in \{0,1\}$, it follows that for $\boldsymbol{x} \neq 0$, $\sum_{j=1}^n \xi_{ij} x_j$ is a sum of independent Poisson random variables and thus has the Poisson distribution

$$\sum_{j=1}^n \xi_{ij} x_j \sim \mathrm{Pois}\left(\lambda_i(\boldsymbol{x})\right),$$

where we defined $\lambda_i(\boldsymbol{x}) := \sum_{j=1}^n \lambda_{ij} x_j$. Let $Q(p, \lambda)$ denote the upper regularized gamma function:

$$Q(p,\lambda) := 1 - P(p,\lambda) = \int_\lambda^\infty \frac{t^{p-1}e^{-t}}{\Gamma(p)} dt \quad \text{for } p \geq 0,\ \lambda \geq 0. \tag{3.3.19}$$

Here $\Gamma(p)$ and $P(p, \lambda)$ are defined by (3.3.1) and (3.3.2), respectively. It is well known that

$$\sum_{k=0}^{N} \frac{\lambda^k}{k!} e^{-\lambda} = Q(N+1, \lambda)$$

for any nonnegative integer $N$. The function in the joint probabilistic constraint:

$$F_i(\boldsymbol{x}) = \mathbb{P}\left( \sum_{j=1}^n \xi_{ij} x_j \leq W_i \right) = Q(\lfloor W_i \rfloor + 1, \lambda_i(\boldsymbol{x})) \tag{3.3.20}$$

is defined for $\boldsymbol{x} \geq 0$. The logconcavity of $Q(p, \lambda)$ in the following lemma together with Remark 3.2.1 ensures the logconcavity of $F_i(\boldsymbol{x})$ in (3.3.20), which proves Theorem 3.2.1 for the case of the Poisson distribution.

**Lemma 3.3.4.** *For any fixed $p \geq 1$, the function $Q(p, \lambda)$ defined by* (3.3.19) *is strictly decreasing and logconcave for $\lambda \geq 0$.*

*Proof.* First we prove the decreasing property. It follows that for $\lambda > 0$,

$$\frac{dQ(p,\lambda)}{d\lambda} = -\frac{\lambda^{p-1}e^{-\lambda}}{\Gamma(p)} < 0.$$

Hence $Q(p, \lambda)$ is strictly decreasing for $\lambda \geq 0$.

Next we prove the logconcavity. Let us introduce the following continuous and logconcave p.d.f.:

$$f(y) := \frac{y^{p-1}e^{-y}}{\Gamma(p)} \quad \text{for } y \geq 0.$$

Note that

$$f(y) \geq 0, \quad \int_0^\infty f(y)dy = 1, \quad \frac{d^2}{dy^2} \ln f(y) = -\frac{p-1}{y^2} \leq 0.$$

It follows from Theorem 4.2.4 in Prékopa [58] that

$$1 - \int_{-\infty}^{\lambda} f(y)dy = \int_\lambda^\infty f(y)dy = \int_\lambda^\infty \frac{y^{p-1}e^{-y}}{\Gamma(p)} dy = Q(p,\lambda)$$

is logconcave for $\lambda \geq 0$. $\square$

### 3.3.4 Convexity result for the binomial distribution.

In this section we consider a case where the independent random variable $\xi_{ij}$ in (3.2.4) has the binomial distribution

$$\xi_{ij} \sim B(n_{ij},\, p_i)$$

with the number of trials $n_{ij} \in \mathbb{N}$ and the success probability in each trial $p_i \in (0,1)$ for some $i \in \{1, \ldots, m\}$ and all $j = 1, \ldots, n$. The pmf of $\xi_{ij}$ is as follows.

$$f(k) = \binom{n_{ij}}{k} p_i^k (1-p_i)^{n_{ij}-k} \quad \text{for } k = 0, 1, \ldots, n_{ij}$$

Note that $\mathrm{Var}(\xi_{ij}) = p_i\, \mathbb{E}(\xi_{ij}) = n_{ij} p_i^2$. Since $x_j \in \{0,1\}$, it follows that for $\boldsymbol{x} \neq 0$, $\sum_{j=1}^n \xi_{ij} x_j$ is a sum of independent binomial random variables with the common success probability $p_i$ and thus has the binomial distribution

$$\sum_{j=1}^n \xi_{ij} x_j \sim B\left(n_i(\boldsymbol{x}),\, p_i\right),$$

where we defined $n_i(\boldsymbol{x}) := \sum_{j=1}^n n_{ij} x_j$. Let $I(p; a, b)$ denote the regularized beta function:

$$I(p; a, b) := \frac{\int_0^p y^{a-1}(1-y)^{b-1} dy}{\int_0^1 y^{a-1}(1-y)^{b-1} dy}.$$

Let us define a continuous function:

$$J(z; c, p) := \begin{cases} I(1-p; z-c, c+1) & \text{for } z > c \\ 1 & \text{for } z \le c \end{cases}, \tag{3.3.21}$$

where $c > -1$ and $p \in (0,1)$. It is well known that

$$0 < J(z; c, p) < 1 \ \text{ for } z > c, \quad \lim_{z \to c+} J(z; c, p) = 1, \quad \lim_{z \to \infty} J(z; c, p) = 0.$$

It is well known (and easy to prove, e.g., by induction) that

$$\sum_{k=0}^{\min(N,n)} \binom{n}{k} p^k (1-p)^{n-k} = J(n; N, p)$$

for any nonnegative integers $n$ and $N$. The function in the joint probabilistic constraint:

$$F_i(\boldsymbol{x}) = \mathbb{P}\left(\sum_{j=1}^n \xi_{ij} x_j \le W_i\right) = J\left(n_i(\boldsymbol{x}); \lfloor W_i \rfloor, p_i\right) \tag{3.3.22}$$

is defined for $\boldsymbol{x} \geq 0$. The logconcavity of $J(z; c, p)$ in the following lemma together with Remark 3.2.1 ensures the logconcavity of $F_i(\boldsymbol{x})$ in (3.3.22), which proves Theorem 3.2.1 for the case of the binomial distribution.

**Lemma 3.3.5.** *For any fixed $c \geq 0$ and $p \in (0, 1)$, the function $J(z; c, p)$ defined by* (3.3.21) *is decreasing and logconcave for $z \in \mathbb{R}$.*

*Proof.* First we prove the decreasing property. Let us designate $f(y, z; c) := y^c(1 - y)^{z-c-1}$. For $z > c$, the derivative of $J(z; c, p) = I(1 - p; z - c, c + 1) = 1 - I(p; c + 1, z - c) = \int_p^1 f(y, z; c)dy / \int_0^1 f(y, z; c)dy$ is calculated as follows:

$$\frac{dJ(z;c,p)}{dz} = \frac{d}{dz}\frac{\int_p^1 f(y,z;c)dy}{\int_0^1 f(y,z;c)dy}$$
$$= J(z;c,p)\left(\frac{\int_p^1 f(y,z;c)\ln(1-y)dy}{\int_p^1 f(y,z;c)dy} - \frac{\int_0^1 f(y,z;c)\ln(1-y)dy}{\int_0^1 f(y,z;c)dy}\right). \tag{3.3.23}$$

The derivative with respect to $p$ of the first term in the parenthesis in (3.3.23) is

$$\frac{d}{dp}\frac{\int_p^1 f(y,z;c)\ln(1-y)dy}{\int_p^1 f(y,z;c)dy}$$
$$= \frac{f(p,z;c)}{\left(\int_p^1 f(y,z;c)dy\right)^2}\left[\int_p^1 f(y,z;c)\left(\ln(1-y) - \ln(1-p)\right)dy\right]$$
$$< 0.$$
$$(\because f(p,z;c) > 0,\ f(y,z;c) > 0 \text{ and } \ln(1-y) - \ln(1-p) < 0 \text{ on } y \in (p,1))$$

Thus the first term in the parenthesis in (3.3.23) is a strictly decreasing function of $p$, which implies $J(z; c, p)$ is strictly decreasing for $z \geq c$ and hence decreasing for $z \in \mathbb{R}$.

Next we prove the logconcavity. For $z > c$, the second derivative of $\ln J(z; c, p)$ is calculated as follows:

$$\frac{d^2}{dz^2}\ln J(z;c,p) =$$
$$\frac{\int_p^1 f(y,z;c)\left\{\ln(1-y)\right\}^2 dy}{\int_p^1 f(y,z;c)dy} - \left(\frac{\int_p^1 f(y,z;c)\ln(1-y)dy}{\int_p^1 f(y,z;c)dy}\right)^2$$
$$- \left[\frac{\int_0^1 f(y,z;c)\left\{\ln(1-y)\right\}^2 dy}{\int_0^1 f(y,z;c)dy} - \left(\frac{\int_0^1 f(y,z;c)\ln(1-y)dy}{\int_0^1 f(y,z;c)dy}\right)^2\right]. \tag{3.3.24}$$

Let us introduce a random variable $X$ that has the following continuous and logconcave p.d.f.:

$$g(x) = \frac{(1-e^{-x})^c e^{-x(z-c)}}{\int_0^\infty (1-e^{-y})^c e^{-y(z-c)} dy} \quad \text{for } x \geq 0, \tag{3.3.25}$$

where $c \geq 0$ and $z > c$ are fixed. Note that

$$g(x) \geq 0, \quad \int_0^\infty g(x)dx = 1, \quad \frac{d^2 \ln g(x)}{dx^2} = -\frac{ce^{-x}}{(1-e^{-x})^2} \leq 0.$$

Then by changing variable of integration by $x = \ln 1/(1-y)$, the second derivative (3.3.24) can be written as

$$\frac{\int_{\ln 1/(1-p)}^\infty x^2 g(x)dx}{\int_{\ln 1/(1-p)}^\infty g(x)dx} - \left(\frac{\int_{\ln 1/(1-p)}^\infty xg(x)dx}{\int_{\ln 1/(1-p)}^\infty g(x)dx}\right)^2 - \left[\frac{\int_0^\infty x^2 g(x)dx}{\int_0^\infty g(x)dx} - \left(\frac{\int_0^\infty xg(x)dx}{\int_0^\infty g(x)dx}\right)^2\right]$$

$$= \mathbb{E}(X^2|X \geq \ln 1/(1-p)) - \mathbb{E}^2(X|X \geq \ln 1/(1-p)) - \left[\mathbb{E}(X^2) - \mathbb{E}^2(X)\right]$$

$$= \mathbb{E}(X^2|X \geq v) - \mathbb{E}^2(X|X \geq v) - \left[\mathbb{E}(X^2) - \mathbb{E}^2(X)\right]. \quad \text{(We denote } v := \ln 1/(1-p))$$

Since the random variable $X$ has a continuous and logconcave p.d.f. $g(x)$, it follows by theorem 2.1 in Prékopa [61] that

$$\mathbb{E}(X^2|X \geq v) - \mathbb{E}^2(X|X \geq v)$$

is decreasing in $v$. We have proved that for any $v > 0$ the following inequality holds:

$$\mathbb{E}(X^2|X \geq v) - \mathbb{E}^2(X|X \geq v) \leq \mathbb{E}(X^2) - \mathbb{E}^2(X),$$

which implies $d^2 \ln J(z;c,p)/dz^2 \leq 0$, and thus $J(z;c,p)$ is logconcave for $z \geq c$. We prove that $J(z;c,p)$ is logconcave for all $z$ by verifying that the following inequality is satisfied.

$$J(\mu z_1 + (1-\mu)z_2; c, p) \geq J(z_1; c, p)^\mu J(z_2; c, p)^{1-\mu} \tag{3.3.26}$$

The fact that (3.3.26) is satisfied for $z_1, z_2 \geq c$ follows from the logconcavity of $J(z;c,p)$ for $z \geq c$ so that we have two cases left. For any $z_1, z_2 < c$, the equality clearly holds (both sides are 1) for (3.3.26). For any $z_1 < c,\ z_2 \geq c$, we have

$$\begin{aligned} J(\mu z_1 + (1-\mu)z_2; c, p) &\geq J(\mu c + (1-\mu)z_2; c, p) \quad (\because J(z;c,p) \text{ is decreasing in } z) \\ &\geq J(c;c,p)^\mu J(z_2;c,p)^{1-\mu} \quad (\because J(z;c,p) \text{ is logconcave for } z \geq c) \\ &= J(z_1;c,p)^\mu J(z_2;c,p)^{1-\mu}. \end{aligned}$$

Therefore, $J(z; c, p)$ is logconcave for $z \in \mathbb{R}$. $\square$

## 3.4 Solutions and computational examples.

### 3.4.1 Solutions

We introduce a known solution to convex mixed-integer nonlinear programming (MINLP) problems to briefly explain how convexity of the relaxed problem helps to find a globally optimal integer solution. DICOPT [34] is a program for solving MINLP problems based on extensions of the outer approximation method. The method generates accumulating tangent hyper-planes of the relaxed problem improving successively linear approximation of nonlinear convex functions that underestimate the objective function and overestimate the feasible region for the case of convex problems. The discrete optimization is performed via a mixed-integer linear programming (MILP), which provides a lower bound (in the case of minimization) on the objective function which increases monotonically as iterations proceed due to the accumulation of linear approximations.

### 3.4.2 Execution times and quality of solutions.

We carried out an experiment to measure the CPU times for computing optimal solutions of the problems and to compare the solutions from the deterministic model and the stochastic models. There are a number of open source and commercial solvers for MINLP available today. We used GAMS [27] (release 24.1.2, 64-bit) as a modeling system with DICOPT as an MINLP solver. We selected CPLEX [38] (version 12.5.1.0) as an MILP solver and SNOPT [30] (version 7.2-12) as a nonlinear programming (NLP) solver both called internally from DICOPT in solving subproblems. GAMS has built-in support for various special math functions including the regularized gamma function, the regularized beta function, and the standard normal c.d.f., which cover all of our cases. It also has support for derivatives of those functions used by NLP solvers. In the case of the Poisson distribution, due to errors in computing the exact derivatives of the regularized gamma function by GAMS, we selected as an NLP solver KNITRO [10] (version 8.1.1), which has an option to compute derivatives by finite-difference approximations. This, however, can seriously degrate the performance and the likelyhood of converging to a solution. The experiment was carried out on the NEOS Server [12, 19, 33], where each job was run on a processor clocked at 2.79 GHz with the maximum memory limit of 400 MB. Both absolute and relative MIP tolerance options in the solver were set to 0 so we obtained the exact optimal 0-1 solutions under the system's accuracy in fractional calculations. A time limit of three hours was imposed in solving a problem.

We performed tests on instances created from the data files for multidimensional knapsack problems in OR-Library [1]. Since the original data was prepared for the deterministic model, we needed to generate test data for the stochastic models. Each instance in the original data consists of:

$$\text{item value } v_j, \quad \text{item attribute } w_{ij}, \quad \text{capacity } W_i \quad \text{for } i = 1, \ldots, m, \quad j = 1, \ldots, n.$$

Test data was prepared in the following way. We used the same

$$\text{item value } v_j, \quad \text{capacity } W_i \quad \text{for } i = 1, \ldots, m, \quad j = 1, \ldots, n$$

as those in the original data and the common probability level $q = 0.9$ for all instances. The parameters for the distributions of random item attribute $\xi_{ij}$ were generated from $w_{ij}$ as follows.

$$\begin{array}{ll}
r_i \text{ randomly generated in } [0.05, 0.10] & \\
\text{Gamma } [\xi_{ij} \sim \Gamma(p_{ij}, \theta_i)] : & \theta_i = r_i, \quad p_{ij} = w_{ij}/\theta_i \\
\text{Normal } [\xi_{ij} \sim \mathcal{N}(\mu_{ij}, \lambda_i \mu_{ij})] : & \lambda_i = r_i, \quad \mu_{ij} = w_{ij} \\
\text{Poisson } [\xi_{ij} \sim \text{Pois}(\lambda_{ij})] : & \lambda_{ij} = w_{ij} \\
\text{Binomial } [\xi_{ij} \sim B(n_{ij}, p_i)] : & p_i = 1 - r_i, \quad n_{ij} = \lfloor w_{ij}/p_i \rceil
\end{array}$$

In the above generation the mean $\mathbb{E}[\xi_{ij}]$ is equal to $w_{ij}$ (approximately in the case of the binomial distribution) and the variance $\text{Var}[\xi_{ij}]$ is 5 to 10% of the mean (except for the Poisson distribution where the variance is equal to the mean). For each original instance we generated four test instances corresponding to gamma, normal, Poisson, and binomial, in each of which all items follow the same type of distribution.

Table 3.1 shows the CPU time for finding an optimal solution. The first column is the instance name, the second and the third columns the size of the instance, and the fourth through the eighth columns the CPU times, each of which is the median of five runs, for the deterministic model (in column "Det") and the four stochastic models (in columns "$\Gamma$", "$\mathcal{N}$", "Pois", "$B$"). Tables 3.2 and 3.3 compare the optimal objective values ("obj" columns) and the probabilities satisfying the joint knapsack constraints ("jkc%" columns) by using the optimal solutions on the deterministic model and the stochastic models.

Table 3.1: CPU times (in seconds) for computing optimal solutions.

| | | | CPU time | | | | |
|---|---|---|---|---|---|---|---|
| Instance | $m$ | $n$ | Det | $\Gamma$ | $\mathcal{N}$ | Pois | $B$ |
| 1-1 | 10 | 6 | 0.012 | 0.036 | 0.026 | 0.107 | 0.070 |
| 1-2 | 10 | 10 | 0.013 | 0.031 | 0.030 | 0.231 | 0.045 |
| 1-3 | 10 | 15 | 0.014 | 0.042 | 0.037 | 0.278 | 0.045 |
| 1-4 | 10 | 20 | 0.015 | 0.035 | 0.051 | 0.347 | 0.059 |
| 1-5 | 10 | 28 | 0.016 | 0.037 | 0.045 | 0.889 | 0.097 |
| 1-6 | 5 | 39 | 0.019 | 0.066 | 0.102 | 1.889 | 0.152 |
| 1-7 | 5 | 50 | 0.019 | 0.116 | 0.155 | 0.672 | 0.398 |
| 2-WE01 | 5 | 30 | 0.016 | 0.028 | 0.028 | 1.060 | 0.046 |
| 2-WE06 | 5 | 40 | 0.015 | 0.032 | 0.044 | 0.369 | 0.049 |
| 2-WE10 | 5 | 50 | 0.023 | 0.041 | 0.053 | 0.235 | 0.068 |
| 2-PB4 | 2 | 29 | 0.016 | 0.040 | 0.038 | 0.058 | 0.040 |
| 2-PB5 | 10 | 20 | 0.047 | 0.058 | 0.081 | 4.238 | 0.095 |
| 2-PB6 | 30 | 40 | 0.032 | 0.128 | 0.119 | 1.397 | 0.204 |
| cb1-10 | 5 | 100 | 0.92 | 11.18 | 11.11 | 3.57 | 10.87 |
| cb1-20 | 5 | 100 | 0.85 | 0.39 | 0.39 | 2.25 | 0.40 |
| cb1-30 | 5 | 100 | 1.76 | 0.68 | 0.84 | 4.45 | 0.55 |
| cb2-10 | 5 | 250 | 3.32 | 43.80 | 43.85 | 110.04 | 45.92 |
| cb2-20 | 5 | 250 | 29.35 | 104.72 | 101.13 | 44.00 | 23.13 |
| cb2-30 | 5 | 250 | 26.93 | 9.81 | 5.39 | 21.24 | 4.54 |
| cb3-10 | 5 | 500 | 1850.82 | 225.02 | 186.61 | 2538.41 | 207.92 |
| cb3-20 | 5 | 500 | 783.11 | 144.88 | 86.04 | 1128.51 | 68.36 |
| cb3-30 | 5 | 500 | 888.81 | 110.85 | 115.05 | 1545.49 | 79.91 |
| cb4-10 | 10 | 100 | 5.88 | 45.49 | 43.03 | 14.14 | 102.90 |
| cb4-20 | 10 | 100 | 12.65 | 21.58 | 27.28 | 11.30 | 22.74 |
| cb4-30 | 10 | 100 | 3.27 | 13.89 | 15.35 | 3.39 | 12.03 |
| cb5-10 | 10 | 250 | 6336.91 | 3852.86 | 3748.23 | 7198.34 | 4211.01 |
| cb5-20 | 10 | 250 | 1479.26 | 1173.38 | 945.91 | 8332.73 | 1153.43 |
| cb5-30 | 10 | 250 | 523.66 | 8136.13 | 8557.09 | 1.54 | 7341.99 |
| cb7-10 | 30 | 100 | 40.12 | 395.83 | 384.33 | 134.55 | 355.68 |
| cb7-20 | 30 | 100 | 57.77 | 118.59 | 132.74 | 146.41 | 129.56 |
| cb7-30 | 30 | 100 | 51.52 | 329.71 | 347.56 | 7.55 | 518.70 |

Table 3.2: Objective values and probabilities (in %) satisfying the joint knapsack constraints (part 1).

| | | | Deterministic solution | | | | | Stochastic solutions | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | jkc% | | | | $\Gamma$ | | $\mathcal{N}$ | | Pois | | $B$ | |
| Instance | $m$ | $n$ | obj | $\Gamma$ | $\mathcal{N}$ | Pois | $B$ | obj | jkc% | obj | jkc% | obj | jkc% | obj | jkc% |
| 1-1 | 10 | 6 | 3800 | 93.8 | 94.0 | 50.3 | 98.5 | 3800 | 93.8 | 3800 | 94.0 | 3300 | 93.6 | 3800 | 98.5 |
| 1-2 | 10 | 10 | 8706.1 | 52.4 | 52.3 | 34.8 | 52.0 | 8577.8 | 99.9 | 8577.8 | 99.9 | 8577.8 | 91.1 | 8577.8 | 99.9 |
| 1-3 | 10 | 15 | 4015 | 78.8 | 79.0 | 22.4 | 75.0 | 3915 | 90.3 | 3915 | 99.4 | 3605 | 90.2 | 3945 | 90.6 |
| 1-4 | 10 | 20 | 6120 | 28.0 | 27.8 | 21.9 | 30.2 | 6030 | 95.7 | 6030 | 95.8 | 5780 | 90.6 | 6040 | 90.8 |
| 1-5 | 10 | 28 | 12400 | 35.9 | 35.8 | 23.4 | 44.5 | 12310 | 93.3 | 12310 | 93.4 | 11930 | 91.1 | 12310 | 94.7 |
| 1-6 | 5 | 39 | 10618 | 23.1 | 23.0 | 10.5 | 24.6 | 10456 | 91.4 | 10456 | 91.6 | 9949 | 90.0 | 10456 | 92.6 |
| 1-7 | 5 | 50 | 16537 | 13.0 | 12.9 | 9.3 | 20.2 | 16330 | 90.7 | 16330 | 90.9 | 15816 | 90.1 | 16358 | 90.4 |
| 2-WE01 | 5 | 30 | 4554 | 96.2 | 96.2 | 57.9 | 97.2 | 4554 | 96.2 | 4554 | 96.2 | 4424 | 97.8 | 4554 | 97.2 |
| 2-WE06 | 5 | 40 | 5557 | 57.5 | 57.3 | 43.8 | 70.4 | 5533 | 96.9 | 5533 | 97.0 | 5442 | 93.3 | 5533 | 98.8 |
| 2-WE10 | 5 | 50 | 6339 | 93.0 | 93.1 | 52.6 | 96.1 | 6339 | 93.0 | 6339 | 93.1 | 6226 | 95.1 | 6339 | 96.1 |
| 2-PB4 | 2 | 29 | 95168 | 67.1 | 67.0 | 41.4 | 80.7 | 93183 | 92.1 | 93183 | 92.4 | 87627 | 91.9 | 94461 | 93.0 |
| 2-PB5 | 10 | 20 | 2139 | 93.0 | 93.1 | 15.1 | 95.7 | 2139 | 93.0 | 2139 | 93.1 | 1948 | 91.1 | 2139 | 95.7 |
| 2-PB6 | 30 | 40 | 776 | 78.7 | 78.7 | 18.2 | 77.0 | 765 | 99.9 | 765 | 99.9 | 732 | 99.7 | 765 | 99.9 |

Table 3.3: Objective values and probabilities (in %) satisfying the joint knapsack constraints (part 2).

| | | | Deterministic solution | | | | | Stochastic solutions | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | jkc% | | | | $\Gamma$ | | $\mathcal{N}$ | | Pois | | $B$ | |
| Instance | $m$ | $n$ | obj | $\Gamma$ | $\mathcal{N}$ | Pois | $B$ | obj | jkc% | obj | jkc% | obj | jkc% | obj | jkc% |
| cb1-10 | 5 | 100 | 24411 | 74.7 | 74.7 | 27.6 | 74.7 | 24315 | 98.3 | 24315 | 98.3 | 24024 | 97.1 | 24315 | 98.1 |
| cb1-20 | 5 | 100 | 44554 | 57.6 | 57.6 | 11.7 | 59.4 | 44511 | 99.0 | 44511 | 99.0 | 44107 | 95.2 | 44511 | 99.1 |
| cb1-30 | 5 | 100 | 59965 | 25.8 | 25.7 | 10.6 | 26.5 | 59915 | 95.5 | 59915 | 95.5 | 59466 | 95.1 | 59915 | 95.7 |
| cb2-10 | 5 | 250 | 58959 | 16.8 | 16.8 | 5.7 | 17.8 | 58785 | 92.6 | 58785 | 92.6 | 58329 | 93.8 | 58785 | 93.4 |
| cb2-20 | 5 | 250 | 107058 | 10.8 | 10.8 | 6.1 | 10.8 | 106891 | 93.5 | 106891 | 93.5 | 106326 | 93.4 | 106900 | 90.6 |
| cb2-30 | 5 | 250 | 154668 | 10.8 | 10.8 | 4.9 | 10.1 | 154517 | 90.4 | 154517 | 90.4 | 153929 | 92.2 | 154517 | 90.1 |
| cb3-10 | 5 | 500 | 120717 | 7.6 | 7.6 | 4.1 | 7.4 | 120490 | 90.1 | 120490 | 90.1 | 119851 | 91.4 | 120490 | 90.2 |
| cb3-20 | 5 | 500 | 219719 | 6.4 | 6.4 | 3.8 | 5.8 | 219476 | 91.7 | 219476 | 91.7 | 218663 | 91.9 | 219476 | 91.0 |
| cb3-30 | 5 | 500 | 299910 | 4.4 | 4.4 | 3.4 | 4.5 | 299661 | 90.6 | 299661 | 90.6 | 298864 | 90.8 | 299663 | 90.1 |
| cb4-10 | 10 | 100 | 22702 | 79.9 | 79.9 | 31.3 | 77.5 | 22567 | 90.2 | 22567 | 90.3 | 22273 | 96.6 | 22551 | 99.6 |
| cb4-20 | 10 | 100 | 41207 | 63.0 | 63.0 | 4.4 | 63.7 | 41096 | 95.2 | 41096 | 95.2 | 40721 | 98.2 | 41096 | 95.6 |
| cb4-30 | 10 | 100 | 60633 | 89.5 | 89.5 | 23.5 | 89.0 | 60515 | 94.7 | 60515 | 94.7 | 60133 | 96.7 | 60515 | 94.5 |
| cb5-10 | 10 | 250 | 59208 | 18.6 | 18.6 | 3.3 | 18.1 | 59064 | 94.7 | 59064 | 94.8 | 58527 | 95.3 | 59064 | 94.3 |
| cb5-20 | 10 | 250 | 106723 | 7.2 | 7.2 | 1.5 | 6.4 | 106558 | 91.6 | 106558 | 91.6 | 105880 | 94.8 | 106558 | 91.3 |
| cb5-30 | 10 | 250 | 149704 | 10.8 | 10.8 | 1.3 | 10.4 | 149485 | 90.9 | 149485 | 91.0 | 148377 | 99.8 | 149485 | 90.6 |
| cb7-10 | 30 | 100 | 20983 | 54.0 | 54.0 | 20.7 | 55.1 | 20862 | 99.4 | 20862 | 99.4 | 20506 | 99.0 | 20862 | 99.5 |
| cb7-20 | 30 | 100 | 41700 | 48.5 | 48.5 | 12.1 | 51.2 | 41620 | 94.9 | 41620 | 95.0 | 41056 | 99.0 | 41620 | 94.7 |
| cb7-30 | 30 | 100 | 60603 | 61.8 | 61.8 | 10.0 | 60.1 | 60471 | 98.9 | 60471 | 98.9 | 53662 | 99.9 | 60471 | 93.6 |

We see that solutions can be computed instantly for smaller sized problems ($n \leq 50$) while it can take minutes to hours to solve larger sized problems ($n \geq 100$). We observe that solving a stochastic version of the problem is not much more difficult, easier for some instances, than solving the deterministic version. In such instances, though, further numerical tests suggest that solving a stochastic version needs more time as a test instance is generated with smaller variances of the random variables and we set a higher probability level $q$, in which case the problem becomes closer to the deterministic version. The objective function value of a stochastic solution is worse by a few percentages than that of the deterministic solution but the probability satisfying the joint knapsack constraint with the stochastic solution is no less than $q = 0.9$ ensured by the model while it is much lower, often less than 0.5 for larger sized problems, with the deterministic solution. We have to treat the results from the Poisson distribution separately from other distributions because we used a different NLP solver with derivatives computed by approximations, which may be the reason that solutions of some instances are not optimal computed in unusually small times such as in 'cb5-30' and 'cb7-30'.

### 3.4.3 Project selection problem.

We illustrate a numerical example of a project selection problem. Suppose we are given a set of $n = 5$ projects. For each project $j \in \{1, 2, \ldots, 5\}$, the following parameters are given. Its estimated profit is $v_j$. It consumes $m = 4$ types of resources. The random amount $\xi_{1j}$ consumed for the resource 1 follows the gamma distribution with shape $p_{1j}$ and scale $\theta_1$. The random amount $\xi_{2j}$ consumed for the resource 2 follows the normal distribution with mean $\mu_{2j}$ and variance $\lambda_2 \mu_{2j}$. The random amount $\xi_{3j}$ consumed for the resource 3 follows the Poisson distribution with parameter $\lambda_{3j}$. The random amount $\xi_{4j}$ consumed for the resource 4 follows the binomial distribution with number of trials $n_{4j}$ and success probability $p_4$. Note that the random amounts $\xi_{ij}$ follow the same type of distribution in the same resource but follow different types of distributions in different resources. All random amounts $\xi_{ij}$ ($i = 1, \ldots, 4$ and $j = 1, \ldots, 5$) are assumed to be independent. The capacities of the total amount of consumption for the four resource types are $W_1, \ldots, W_4$, respectively. The parameters $c_j, p_{1j}, \mu_{2j}, \lambda_{3j}, N_{4j}, W_i$ are shown in Table 3.4. The parameters $\mu_{2j}$, $\lambda_2$, and $W_2$ satisfy the condition (3.2.8). We want

Table 3.4: Parameters for the projects.

| Project | Profit | Resource 1 $\Gamma(p_{1j}, \theta_1)$ | | Resource 2 $\mathcal{N}(\mu_{2j}, \lambda_2\mu_{2j})$ | | Resource 3 $\text{Pois}(\lambda_{3j})$ | Resource 4 $B(n_{4j}, p_4)$ | |
|---|---|---|---|---|---|---|---|---|
| $j$ | $v_j$ | $p_{1j}$ | $\theta_1$ | $\mu_{2j}$ | $\lambda_2$ | $\lambda_{3j}$ | $n_{4j}$ | $p_4$ |
| 1 | 560 | 210.0 | | 86.0 | | 24.0 | 50 | |
| 2 | 500 | 490.0 | | 153.0 | | 41.0 | 100 | |
| 3 | 170 | 350.0 | 0.09 | 112.0 | 0.12 | 37.0 | 210 | 0.02 |
| 4 | 230 | 140.0 | | 91.0 | | 31.0 | 160 | |
| 5 | 140 | 270.0 | | 98.0 | | 53.0 | 152 | |
| | Capacity | $W_1 = 150.0$ | | $W_2 = 203.0$ | | $W_3 = 78$ | $W_4 = 10$ | |

to find a subset of the projects that maximizes the total estimated profit while keeping the capacity constraints with a high probability. With the probability level $q = 0.9$, we can formulate the stochastic multidimensional knapsack problem as follows:

$$
\begin{aligned}
\text{maximize} \quad & \sum_{j=1}^{5} v_j x_j \\
\text{subject to} \quad & \ln P\left(\sum_{j=1}^{5} p_{1j}x_j, \frac{W_1}{\theta_1}\right) + \ln \Phi\left(\frac{W_2 - \sum_{j=1}^{5} \mu_{2j}x_j}{\sqrt{\lambda_2 \sum_{j=1}^{5} \mu_{2j}x_j}}\right) \\
& \quad + \ln Q\left(\lfloor W_3 \rfloor + 1, \sum_{j=1}^{5} \lambda_{3j}x_j\right) + \ln J\left(\sum_{j=1}^{5} n_{4j}x_j; \lfloor W_4 \rfloor, p_4\right) \geq \ln 0.9, \\
& x_j \in \{0,1\} \quad \text{for } j = 1, \ldots, 5,
\end{aligned}
$$

where functions $P$,$\Phi$,$Q$,$J$ are defined by (3.3.2),(3.2.7),(3.3.19),(3.3.21), respectively. The problem is a convex MINLP due to our results and we can use the same software package as before to solve it. The optimal solution is $\boldsymbol{x} = (1, 0, 0, 1, 0)$. So the best choice is to select the projects 1 and 4.

# Chapter 4

# Improved bounds on the probability of the union of events, some intersections of which are empty

## 4.1 Introduction

Computing the probability of the union of events is important in reliability theory, stochastic programming, and other sciences concerned with stochastic systems. In reliability theory, consider a communication network with nodes and arcs, each with a probability of a failure. The node-to-node reliability of a pair of nodes is the probability of the union of events, each of which occurs when a path between the two nodes consists of arcs without failures. The all-terminal reliability is the probability of the union of events, each of which occurs when a spanning tree of the network consists of arcs without failures. In probabilistic constrained stochastic programming, a joint probabilistic constraint for random variables $X_1, \dots, X_n$ specifies a lower bound on $\mathbb{P}(X_1 \le z_1 \cap \dots \cap X_n \le z_n) = 1 - \mathbb{P}(X_1 > z_1 \cup \dots \cup X_n > z_n)$, which involves the probability of the union of events. Although it is hard to compute the exact probability of the union of a large number of events, we can compute an approximation of it by using the information about individual events or intersections of a small number of events.

Let $\{ A_1, A_2, \dots, A_n \}$ be a set of arbitrary events in some probability space and introduce the notation for the probability of the intersections of its subsets.

$$p_I := \mathbb{P}\left( \bigcap_{j \in I} A_j \right) \quad \text{for } I \subset \{1, \dots, n\}\,. \tag{4.1.1}$$

Let $S_0 := 1$ by definition and define

$$S_k := \sum_{\substack{I \subset \{1,\dots,n\} \\ |I| = k}} p_I \quad \text{for } k = 1, 2, \dots, n. \tag{4.1.2}$$

The classical inclusion-exclusion principle (see de Moivre [15] for the concept and the first appearance in a paper in da Silva [13] and later in Sylvester [68]) gives the probability of the union of events as follows:

$$\mathbb{P}(A_1 \cup \cdots \cup A_n) = S_1 - S_2 + \cdots + (-1)^{n-1} S_n. \tag{4.1.3}$$

However, this formula is impractical if the number of events $n$ is large, in which case the calculation of $S_k$ is intractable unless $k$ is close to 1 or $n$ due to $\binom{n}{k}$ sums in (4.1.2). We can still approximate the bounds using a few $S_k$'s. The Bonferroni inequalities (see Bonferroni [5]) states that for $m \leq n$,

$$\mathbb{P}(A_1 \cup \cdots \cup A_n) \begin{Bmatrix} \geq \\ \leq \end{Bmatrix} S_1 - S_2 + \cdots + (-1)^{m-1} S_m \quad \begin{cases} \text{if } m \text{ is even} \\ \text{if } m \text{ is odd} \end{cases}. \tag{4.1.4}$$

These bounds are usually very weak unless $m$ is large.

The best possible (sharp) bounds using only a few $S_k$'s have been found in closed forms. The number of $S_k$'s used is called the order of the bound. The second order sharp lower bound based on $S_1, S_2$ was obtained by Dawson and Sankoff [14] and its upper bound by Kwerel [44, 45] and Sathe et al. [67]. See also Galambos [26] and Prékopa [56]. The third order sharp bounds based on $S_1, S_2, S_3$ were obtained by Kwerel [44, 45] and Boros and Prékopa [8]. The fourth order sharp upper bound based on $S_1, S_2, S_3, S_4$ was given by Boros and Prékopa [8].

While the fifth or higher order sharp bounds have not been known in closed forms, Prékopa [56] observed that all these bounds are the optimal objective values of the binomial moment problems (also regarded as aggregated linear programming problems) obtained from the formulation by Hailperin [35]. See Prékopa [56] for the structures of the dual feasible bases of the problems and Boros and Prékopa [8] for the property of the optimal solution of their dual problems. Let $\nu$ designate the random number of events among $A_1 \ldots, A_n$ that occur. Then we have the following relations (see Takács [70] and Prékopa [58]):

$$\mathbb{E}\left[\binom{\nu}{k}\right] = S_k \quad \text{for } k = 0, 1, \ldots, n, \tag{4.1.5}$$

which can also be written as

$$\sum_{i=0}^{n} \binom{i}{k} v_i = S_k \quad \text{for } k = 0, 1, \ldots, n, \tag{4.1.6}$$

where $v_i := \mathbb{P}(\nu = i)$ for $i = 0, 1, \ldots, n$. The value $S_k$ is called the $k$-th binomial moment of $\nu$. If only $S_1, \ldots, S_m$ are known, the sharp lower and upper bounds on the probability of the union are given by the minimization and maximzation, respectively, of the binomial moment problem as follows (see Hailperin [35] and Prékopa [56]):

$$\begin{aligned} \min(\max) \; & \sum_{i=1}^{n} x_i \\ \text{subject to} \; & \sum_{i=1}^{n} \binom{i}{k} x_i = S_k \quad \text{for } k = 1, \ldots, m \\ & x_i \geq 0 \quad \text{for } i = 1, \ldots, n. \end{aligned} \tag{4.1.7}$$

In practice, however, we are usually not given the values of $S_1, \ldots, S_m$ but we calculate them by (4.1.2) from $p_I$'s, which are in many cases easily calculated for small $|I|$. By simply using the aggregated information $S_1, \ldots, S_m$, we lose the information in individual events. If $p_I,\ I \subset \{1, \ldots, n\},\ 1 \leq |I| \leq m$ are known, the sharp lower and upper bounds on the probability of the union are given by the minimization and maximization, respectively, of the Boolean probability bounding scheme (also regarded as disaggregated linear programming problems), which was initiated by Boole [6] and provided by Hailperin [35], as follows:

$$\begin{aligned} \min(\max) \; & \sum_{\emptyset \neq J \subset \{1,\ldots,n\}} x_J \\ \text{subject to} \; & \sum_{\emptyset \neq J \subset \{1,\ldots,n\}} a_{IJ} x_J = p_I \;\; \text{for } I \subset \{1, \ldots, n\},\ 1 \leq |I| \leq m, \\ & x_J \geq 0 \quad \text{for } \emptyset \neq J \subset \{1, \ldots, n\}, \end{aligned} \tag{4.1.8}$$

where we defined

$$a_{IJ} = \begin{cases} 1 & \text{if } I \subset J \\ 0 & \text{otherwise} \end{cases}.$$

Although these disaggregated problems (4.1.8) give much better bounds than those from the aggregated problems (4.1.7), they are impractical if the number of events $n$ is large due to the exponential number $(2^n - 1)$ of decision variables $x_J$'s.

Probability bounds that utilize structures of individual and intersections of events have been studied improving on those from the aggregated problems but avoiding the exponential size in the disaggregated problems. The first significant result is Hunter's upper bound (see Hunter [37] and Worsley [77]):

$$\mathbb{P}(A_1 \cup \cdots \cup A_n) \leq S_1 - \sum_{(i,j)\in T} p_{ij}, \tag{4.1.9}$$

where $T$ is the heaviest spanning tree of the $n$-node complete graph with each edge $(i,j)$ assigned the weight $p_{ij}$. Hunter's bound was generalized by Tomescu [71] and improved on using special hypergraph structures by Bukszár and Prékopa [9]. Prékopa and Gao [62] defined the linear programming problems which are obtained by partial aggregation and disaggregation, balancing the size of the problem and the quality of bounds. Prékopa, M. Subasi, and E. Subasi [63] gave the sharp bounds where the probability distribution of the occurrences of events is unimodal with known mode.

Section 4.2 presents our main result. We formulate the linear programming problems that give improved lower and upper bounds on the probability of the union of events when some $p_I$'s are known to be 0 or very small. Section 4.3 presents a numerical example comparing the bounds from our work with those obtained from the binomial moment problem.

## 4.2 Improved bounds by the maximum independent set problem and its extension

We provide improved bounds on the probability of the union of events using structures of individual and intersections of events. Suppose we are given all $p_I$ for $1 \leq |I| \leq m$. First let us consider the case for $m = 2$. In practice, some event occurs only when another event occurs and some pair of events never occurs together. If an event $A_i$ occurs only when another event $A_j$ occurs:

$$A_i \subset A_j \Leftrightarrow \mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i) \ \ (p_{ij} = p_i)\,, \tag{4.2.1}$$

then we can eliminate the event $A_i$ to reduce the size $n$ of the problem. The following remark is obvious.

**Remark 4.2.1.** *For a set of $n$ events, probabilities of individual or intersections of events $p_I$'s defined by* (4.1.1) *and binomial moments $S_k$'s defined by* (4.1.2) *have the following properties.*

*(i) $p_I \geq p_J$ for any $I \subset J \subset \{ 1, \ldots, n \}$.*

*(ii) If $p_I = 0$ for some $I \subset \{ 1, \ldots, n \}$ then $S_n = 0$.*

Suppose some pair of events $(A_i, A_j)$ never occurs together:

$$A_i \cap A_j = \emptyset \Leftrightarrow \mathbb{P}(A_i \cap A_j) = p_{ij} = 0. \tag{4.2.2}$$

If there are one or more such pairs, then the minimum order $\ell$ exists such that $S_\ell = S_{\ell+1} = \cdots = S_n = 0$ because of Remarks 4.2.1 (ii). We can find a good upper bound on $\ell$ by solving the maximum independent set (MIS) problem as follows. Consider an undirected graph $G = (V, E)$ of $n = |V|$ nodes corresponding to the $n$ events $A_1, \ldots, A_n$, respectively. Create an edge between two nodes $A_i$ and $A_j$ if $\mathbb{P}(A_i \cap A_j) = 0$. An independent set (or stable set) is a set of nodes, no two of which are adjacent. The size of an independent set is the number of nodes it contains. The MIS problem is to find a largest independent set for a given graph. Any set of nodes (events) whose size is greater than the maximum contains at least one pair of adjacent nodes (events) whose union is empty. So the probability of the union of the events in the set is 0. Thus all binomial moments whose order are higher than the maximum are 0. The integer programming formulation of the MIS problem is as follows:

$$\begin{aligned} \text{maximize} \ & \sum_{u \in V} x_u \\ \text{subject to} \ & x_u + x_v \leq 1 \quad \text{for } (u,v) \in E, \\ & x_u \in \{0,1\} \quad \text{for } u \in V. \end{aligned} \tag{4.2.3}$$

The problem is $\mathcal{N}P$ hard so it is unlikely we can find the exact solution efficiently. But we only need an upper bound, which can be obtained by solving the following LP relaxation of the problem (4.2.3) expressed in a different way:

$$\begin{aligned} &\text{maximize } \sum_{i=1}^{n} x_i \\ &\text{subject to } \ x_i + x_j \le 1 \quad \text{for } \{i,j\} \subset \{1,\dots,n\} \text{ where } p_{ij} = 0, \\ &\qquad\qquad\quad x_i \ge 0 \quad \text{for } i = 1,\dots,n. \end{aligned} \tag{4.2.4}$$

Next let us consider the general case for $m \le n$. Suppose some tuple of events $(A_{i_1},\dots,A_{i_k})$ never occurs together:

$$A_{i_1} \cap \cdots \cap A_{i_k} = \emptyset \Leftrightarrow \mathbb{P}(A_{i_1} \cap \cdots \cap A_{i_k}) = p_{i_1,\dots,i_k} = 0. \tag{4.2.5}$$

If there are one or more such tuples, then we can obtain a better upper bound on the minimum order $\ell$ by solving the following LP which extends the problem (4.2.4):

$$\begin{aligned} &\text{maximize } \sum_{i=1}^{n} x_i \\ &\text{subject to } \sum_{i \in I} x_i \le k-1 \quad \text{for } I \subset \{1,\dots,n\},\ |I| = k \ \text{ where } p_I = 0 \\ &\qquad\qquad\quad \text{for } k = 2,\dots,m, \\ &\qquad\qquad\quad x_i \ge 0 \quad \text{for } i = 1,\dots,n. \end{aligned} \tag{4.2.6}$$

The IP version ($x_i \in \{0,1\}$ instead of $x_i \ge 0$) of this problem means that we find the maximum number of events that include no tuple of events with size up to $m$ whose intersections are empty. Let $U$ denote the optimal objective value of the problem (4.2.6). We have the following conditions:

$$\begin{aligned} &S_k = 0 \text{ (or more precisely, } p_I = 0 \ \text{ for } I \subset \{1,\dots,n\},\ |I| = k) \\ &\text{for } k = \lfloor U \rfloor + 1,\dots,n. \end{aligned} \tag{4.2.7}$$

Improved lower and upper bounds based on the binomial moment problem (4.1.7) are given by the optimal objective values of the minimization and maximization, respectively, of the following LP:

$$\begin{aligned}
\min(\max)\ & \sum_{i=1}^{n} x_i \\
\text{subject to}\ & \sum_{i=1}^{n} \binom{i}{k} x_i = S_k \ \text{ for } k = 1, \ldots, \min(m, \lfloor U \rfloor), \\
& \sum_{i=1}^{n} \binom{i}{k} x_i = 0 \ \text{ for } k = \lfloor U \rfloor + 1, \ldots, n, \\
& x_i \geq 0 \quad \text{for } i = 1, \ldots, n.
\end{aligned} \tag{4.2.8}$$

Since the solution of this problem gives $x_i = 0$ for $i = \lfloor U \rfloor + 1, \ldots, n$, we can simplify the formulation:

$$\begin{aligned}
\min(\max)\ & \sum_{i=1}^{\lfloor U \rfloor} x_i \\
\text{subject to}\ & \sum_{i=1}^{\lfloor U \rfloor} \binom{i}{k} x_i = S_k \ \text{ for } k = 1, \ldots, \min(m, \lfloor U \rfloor), \\
& x_i \geq 0 \quad \text{for } i = 1, \ldots, \lfloor U \rfloor.
\end{aligned} \tag{4.2.9}$$

Similarly, improved lower and upper bounds based on the Boolean probability bounding scheme (4.1.8) are given by the optimal objective values of the minimization and maximization, respectively, of the following LP:

$$\begin{aligned}
\min(\max)\ & \sum_{\substack{J \subset \{1,\ldots,n\} \\ 1 \leq |J| \leq \lfloor U \rfloor}} x_J \\
\text{subject to}\ & \sum_{\substack{J \subset \{1,\ldots,n\} \\ 1 \leq |J| \leq \lfloor U \rfloor}} a_{IJ} x_J = p_I \ \text{ for } 1 \leq |I| \leq \min(m, \lfloor U \rfloor), \\
& x_J \geq 0 \quad \text{for } \ J \subset \{1, \ldots, n\},\ 1 \leq |J| \leq \lfloor U \rfloor.
\end{aligned} \tag{4.2.10}$$

While the number of decision variables $x_J$'s is still exponential, the reduction from the original problem is exponentially large.

Now let us consider a more general case where the probabilities of some intersections of events are very small instead of 0. Suppose some tuple of events $(A_{i_1}, \ldots, A_{i_k})$ occurs jointly with a very low probability no greater than a fixed $\varepsilon \geq 0$:

$$\mathbb{P}(A_{i_1} \cap \cdots \cap A_{i_k}) = p_{i_1,\ldots,i_k} \leq \varepsilon. \tag{4.2.11}$$

Then consider the following LP similar to (4.2.6):

$$
\begin{aligned}
&\text{maximize} \quad \sum_{i=1}^{n} x_i \\
&\text{subject to} \quad \sum_{i \in I} x_i \leq k-1 \quad \text{for } I \subset \{1,\dots,n\},\ |I| = k \ \text{ where } p_I \leq \varepsilon \\
&\qquad\qquad\quad \text{for } k = 2,\dots,m, \\
&\qquad\qquad\quad x_i \geq 0 \quad \text{for } i = 1,\dots,n.
\end{aligned}
\tag{4.2.12}
$$

Let $U_\varepsilon$ denote the optimal objective value of the problem (4.2.12). Because of Remark 4.2.1 (i), we have the following conditions:

$$
\begin{aligned}
&S_k \leq \varepsilon \binom{n}{k} \ \text{(or more precisely, } p_I \leq \varepsilon \ \text{ for } I \subset \{1,\dots,n\},\ |I| = k) \\
&\text{for } k = \lfloor U_\varepsilon \rfloor + 1,\dots,n.
\end{aligned}
\tag{4.2.13}
$$

Improved lower and upper bounds based on the binomial moment problem (4.1.7) are given by the optimal objective values of the minimization and maximization, respectively, of the following LP:

$$
\begin{aligned}
&\min(\max) \quad \sum_{i=1}^{n} x_i \\
&\text{subject to} \quad \sum_{i=1}^{n} \binom{i}{k} x_i = S_k \ \text{ for } k = 1,\dots,\min(m, \lfloor U_\varepsilon \rfloor), \\
&\qquad\qquad\quad \sum_{i=1}^{n} \binom{i}{k} x_i \leq \varepsilon \binom{n}{k} \ \text{ for } k = \lfloor U_\varepsilon \rfloor + 1,\dots,n, \\
&\qquad\qquad\quad x_i \geq 0 \quad \text{for } i = 1,\dots,n.
\end{aligned}
\tag{4.2.14}
$$

The formulation based on the Boolean probability bounding scheme using $\varepsilon$ and $U_\varepsilon$ is impractical since it contains an exponential number of constraints.

## 4.3 Numerical examples

Consider a unit square $\left\{ (x,y) \in \mathbb{R}^2 \,\middle|\, 0 \leq x \leq 1,\ 0 \leq y \leq 1 \right\}$ on a plain. For each $i \in \{1,\dots,n\}$, select uniformly randomly two distinct x-values $x_{i1}$, $x_{i2}$ where $0 \leq x_{i1} < x_{i2} \leq 1$ and two distinct y-values $y_{i1}$, $y_{i2}$ where $0 \leq y_{i1} < y_{i2} \leq 1$ on the unit square. Let us assign the probability of the event $A_i$ as the area of the rectangle defined by

$\{\ (x,y) \in \mathbb{R}^2 \ \big|\ x_{i1} \le x \le x_{i2},\ y_{i1} \le y \le y_{i2}\ \}$:

$$\mathbb{P}(A_i) = (x_{i2} - x_{i1})(y_{i2} - y_{i1}). \tag{4.3.1}$$

The joint probability of two events $A_i,\ A_j$ is the area of the intersection of the two rectangles associated with them.

$$\mathbb{P}(A_i \cap A_j) = \begin{cases} (\min(x_{i2}, x_{j2}) - \max(x_{i1}, x_{j1}))(\min(y_{i2}, y_{j2}) - \max(y_{i1}, y_{j1})) \\ \quad \text{if } \min(x_{i2}, x_{j2}) > \max(x_{i1}, x_{j1}) \text{ and } \min(y_{i2}, y_{j2}) > \max(y_{i1}, y_{j1}) \\ 0 \quad \text{otherwise} \end{cases} \tag{4.3.2}$$

Similarly, the joint probability of three or more events is the area of the intersection of the rectangles associated with them. A certain number of joint probabilities are expected to be 0 in this example. We compute lower and upper bounds on the probability of the union of the events $A_1, \ldots, A_n$. Table 4.1 compares the bounds obtained from our results (4.2.6) and (4.2.9) and those obtained from the binomial moment problem (4.1.7).

Table 4.1: Lower and upper bounds on the union of events.

| | | Our result | | | | | | Binomial moment problem | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $m = 2$ | | $m = 3$ | | $m = 4$ | | $m = 2$ | | $m = 3$ | | $m = 4$ | |
| $n$ | Exact | LB | UB | LB | UB | LB | UB | LB | UB | LB | UB | LB | UB |
| 10 | 0.7283 | 0.5750 | 1 | 0.6472 | 0.7941 | 0.7117 | 0.7559 | 0.5750 | 1 | 0.6182 | 0.7941 | 0.7117 | 0.7712 |
| 15 | 0.9409 | 0.8173 | 1 | 0.8523 | 1 | 0.9112 | 0.9855 | 0.8173 | 1 | 0.8309 | 1 | 0.9112 | 1 |
| 20 | 0.7950 | 0.6271 | 1 | 0.6746 | 0.8782 | 0.7760 | 0.8478 | 0.6271 | 1 | 0.6485 | 0.8782 | 0.7760 | 0.8630 |
| 30 | - | 0.6250 | 1 | 0.6654 | 1 | 0.7411 | 0.9934 | 0.6250 | 1 | 0.6434 | 1 | 0.7411 | 1 |

# References

[1] J. E. Beasley. OR-Library: Distributing test problems by electronic mail. *Journal of the Operational Research Society*, 41(11):1069–1072, 1990.

[2] R. E. Bellman. *Reduction of General Symmetric Matrices to Dianogal Form*, chapter 4, pages 44–72. SIAM, 2nd edition, 1997.

[3] R. E. Bellman and S. E. Dreyfus. *Applied Dynamic Programming.* Princeton University Press, Princeton, NJ, 1962.

[4] P. Beraldi and A. Ruszczyński. A branch and bound method for stochastic integer problems under probabilistic constraints. *Optimization Methods and Software*, 17(3):359–382, 2002.

[5] C. E. Bonferroni. *Teoria statistica delle classi e calcolo delle probabilità*. Pubblicazioni del R. Istituto superiore di scienze economiche e commerciali di Firenze. Libreria internazionale Seeber, 1936.

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

[7] E. Boros, K. Elbassioni, V. Gurvich, L. Khachiyan, and K. Makino. An intersection inequality for discrete distributions and related generation problems. In J. C. M. Baeten, J. K. Lenstra, J. Parrow, and G. J. Woeginger, editors, *Automata, Languages and Programming*, volume 2719 of *Lecture Notes in Computer Science*, pages 543–555. Springer, 2003.

[8] E. Boros and A. Prékopa. Closed form two-sided bounds for probabilities that at least $r$ and exactly $r$ out of $n$ events occur. *Mathematics of Operations Research*, 14(2):317–342, 1989.

[9] J. Bukszár and A. Prékopa. Probability bounds with cherry trees. *Mathematics of Operations Research*, 26(1):174–192, 2001.

[10] R. H. Byrd, J. Nocedal, and R. A. Waltz. KNITRO: An integrated package for nonlinear optimization. In *Large-scale nonlinear optimization*, pages 35–59. Springer, 2006.

[11] A. Charnes, W. W. Cooper, and G. H. Symonds. Cost horizons and certainty equivalents: an approach to stochastic programming of heating oil. *Management Science*, 4(3):235–263, 1958.

[12] J. Czyzyk, M. Mesnier, and J. Moré. The NEOS server. *IEEE Journal on Computational Science and Engineering*, 5:68–75, 1998.

[13] D. da Silva. Proprietades geraes. *J. de l'Ecole Polytechnique*, cah. 30. I, 1854.

[14] D. A. Dawson and D. Sankoff. An inequality for probabilities. *Proceedings of the American Mathematical Society*, 18(3):504–507, 1967.

[15] A. de Moivre. *The doctrine of chances: or, a method of calculating the probability of events in play*. W. Pearson, 1718.

[16] S. de Vries and R. V. Vohra. Combinatorial auctions: A survey. *INFORMS Journal on Computing*, 15(3):284–309, 2003.

[17] I. Deák. Solving stochastic programming problems by successive regression approximations – numerical results. In K. Marti, Y. Ermoliev, and G. Pflug, editors, *Dynamic Stochastic Optimization*, volume 532 of *Lecture Notes in Economics and Mathematical Systems*, pages 209–224. Springer, 2004.

[18] D. Dentcheva, A. Prékopa, and A. Ruszczyński. Concavity and efficient points of discrete distributions in probabilistic programming. *Mathematical Programming*, 89(1):55–77, 2000.

[19] E. Dolan. The NEOS server 4.0 administrative guide. Technical Memorandum ANL/MCS-TM-250, Mathematics and Computer Science Division, Argonne National Laboratory, 2001.

[20] J. D. Esary, F. Proschan, and D. W. Walkup. Association of random variables, with applications. *The Annals of Mathematical Statistics*, 38(5):1466–1474, 1967.

[21] M. Fekete and G. Pólya. Über ein problem von laguerre. *Rendiconti del Circolo Matematico di Palermo*, 34(1):89–120, 1912.

[22] C. E. Ferreira, M. Grötschel, R. W. A. Martin, S. Kiefl, and L. Krispenz. Some integer programs arising in the design of main frame computers. *Mathematical Methods of Operations Research*, 38(1):77–100, 1993.

[23] B. Fortz and M. Poss. Easy distributions for combinatorial optimization problems with probabilistic constraints. *Operations Research Letters*, 38(6):545–549, 2010.

[24] A. Fréville. The multidimensional 0-1 knapsack problem: An overview. *European Journal of Operational Research*, 155(1):1–21, 2004.

[25] A. Fréville and S. Hanafi. The multidimensional 0-1 knapsack problem–bounds and computational aspects. *Annals of Operations Research*, 139(1):195–227, 2005.

[26] J. Galambos. Bonferroni inequalities. *The Annals of Probability*, 5(4):577–581, 1977.

[27] GAMS Development Corporation, Washington D.C. *GAMS – A User's Guide*. Available in `http://www.gams.com`.

[28] B. Gavish and H. Pirkul. Allocation of databases and processors in a distributed computing system. In J. Akoka, editor, *Management of Distributed Data Processing*, pages 215–231, North Holland, Amsterdam, 1982.

[29] G. V. Gens and E. V. Levner. Complexity and approximation algorithms for combinatorial problems: A survey. Technical report, Central Economic and Mathematical Institute, Academy of Sciences of the USSR, Moscow, 1979.

[30] P. E. Gill, W. Murray, and M. A. Saunders. SNOPT: An SQP algorithm for large-scale constrained optimization. *SIAM Journal on Optimization*, 12:979–1006, 1997.

[31] P. C. Gilmore and R. E. Gomory. The theory and computation of knapsack functions. *Operations Research*, 14(6):1045–1074, 1966.

[32] V. Goyal and R. Ravi. A PTAS for the chance-constrained knapsack problem with random item sizes. *Operations Research Letters*, 38(3):161–164, 2010.

[33] W. Gropp and J. Moré. Optimization environments and the NEOS server. *Approximation Theory and Optimization*, pages 167–182, 1997.

[34] I. Grossmann, J. Viswanathan, A. Vecchietti, R. Raman, and E. Kalvelagen. *DICOPT*. Engineering Research Design Center, Carnegie Mellon University, Pittsburgh, PA and GAMS Development Coporation, Washington D.C. Available in `http://www.gams.com`.

[35] T. Hailperin. Best possible inequalities for the probability of a logical function of events. *The American Mathematical Monthly*, 72(4):343–359, 1965.

[36] R. Henrion and C. Strugarek. Convexity of chance constraints with independent random variables. *Computational Optimization and Applications*, 41(2):263–276, 2008.

[37] D. Hunter. An upper bound for the probability of a union. *Journal of Applied Probability*, pages 597–603, 1976.

[38] International Business Machines Corporation. *IBM ILOG CPLEX V12.1: User's Manual for CPLEX*.

[39] P. Kall and J. Mayer. *Stochastic Linear Programming: Models, Thoery, and Computation*. Springer, 2005.

[40] S. Kataoka. A stochastic programming model. *Econometrica: Journal of the Econometric Society*, 31(1–2):181–196, 1963.

[41] H. Kellerer, U. Pferschy, and D. Pisinger. *Knapsack Problems*. Springer-Verlag, 2004.

[42] B. H. Korte and R. Schrader. On the existence of fast approximation schemes. In O. L. Mangasarian, R. R. Meyer, and S. M. Robinson, editors, *Nonlinear Programming*, volume 4, pages 415–437. Academic press, 1981.

[43] B. H. Korte and J. Vygen. *Combinatorial Optimization: Theory and Algorithms*, chapter 17. Springer-Verlag, 2008.

[44] S. M. Kwerel. Bounds on the probability of the union and intersection of m events. *Advances in Applied Probability*, pages 431–448, 1975.

[45] S. M. Kwerel. Most stringent bounds on aggregated probabilities of partially specified dependent probability systems. *Journal of the American Statistical Association*, 70(350):472–479, 1975.

[46] C. M. Lagoa, X. Li, and M. Sznaier. Probabilistically constrained linear programs and risk-adjusted controller design. *SIAM Journal on Optimization*, 15(3):938–951, 2005.

[47] J. H. Lorie and L. J. Savage. Three problems in rationing capital. *The Journal of Business*, 28(4):229–239, 1955.

[48] J. Luedtke, S. Ahmed, and G. L. Nemhauser. An integer programming approach for linear programs with probabilistic constraints. *Mathematical Programming*, 122(2):247–272, 2010.

[49] G. Mádi-Nagy and A. Prékopa. On multivariate discrete moment problems and their applications to bounding expectations and probabilities. *Mathematics of Operations Research*, 29(2):229–258, 2004.

[50] C. C. Petersen. Computational experience with variants of the Balas algorithm applied to the selection of R&D projects. *Management Science*, 13(9):736–750, 1967.

[51] A. Prékopa. On probabilistic constrained programming. In *Proceedings of the Princeton symposium on mathematical programming*, pages 113–138. Princeton University Press Princeton, NJ, 1970.

[52] A. Prékopa. Logarithmic concave measures with application to stochastic programming. *Acta Scientiarium Mathematicarum (Szeged)*, 32:301–316, 1971.

[53] A. Prékopa. Contributions to the theory of stochastic programming. *Mathematical Programming*, 4(1):202–221, 1973.

[54] A. Prékopa. Programming under probabilistic constraints with a random technology matrix. *Mathematische Operationsforschung und Statistik, Ser. Optimization*, 5(2):109–116, 1974.

[55] A. Prékopa. Logarithmic concave measures and related topics. In M. A. H. Dempster, editor, *Stochastic Programming*, pages 63–82. Academic Press, 1980.

[56] A. Prékopa. Boole-Bonferroni inequalities and linear programming. *Operations Research*, 36(1):145–162, 1988.

[57] A. Prékopa. Dual method for the solution of a one-stage stochastic programming problem with random RHS obeying a discrete probability distribution. *Zeitschrift für Operations Research*, 34(6):441–461, 1990.

[58] A. Prékopa. *Stochastic Programming*. Kluwer Academic Publishers, 1995.

[59] A. Prékopa. The use of discrete moment bounds in probabilisticconstrained stochastic programming models. *Annals of Operations Research*, 85(0):21–38, 1999.

[60] A. Prékopa. Probabilistic programming. In A. Ruszczyński and A. Shapiro, editors, *Handbooks in Operations Research and Management Science: Stochastic Programming*, volume 10, chapter 5, pages 267–351. Elsevier, 2003.

[61] A. Prékopa. Conditional mean-conditional variance portfolio selection model. RUTCOR research report 34-2007, Rutgers Center for Operations Research, 2007.

[62] A. Prékopa and L. Gao. Bounding the probability of the union of events by aggregation and disaggregation in linear programs. *Discrete applied mathematics*, 145(3):444–454, 2005.

[63] A. Prékopa, M. Subasi, and E. Subasi. Sharp bounds for the probability of the union of events under unimodality condition. *European Journal of Pure and Applied Mathematics*, 1(1):60–81, 2008.

[64] A. Prékopa, B. Vizvári, and T. Badics. Programming under probabilistic constraint with discrete random variable. In F. Giannessi, T. Rapcsák, and S. Komlósi, editors, *New trends in mathematical programming*, pages 235–255. Kluwer Academic Publishers, 1998.

[65] A. Prékopa, K. Yoda, and M. M. Subasi. Uniform quasi-concavity in probabilistic constrained stochastic programming. *Operations Research Letters*, 39(3):188–192, 2011.

[66] M. H. Rothkopf, A. Pekeč, and R. M. Harstad. Computationally manageable combinational auctions. *Management Science*, 44(8):1131–1147, 1998.

[67] Y. S. Sathe, M. Pradhan, and S. P. Shah. Inequalities for the probability of the occurrence of at least $m$ out of $n$ events. *Journal of Applied Probability*, 17(4):1127–1132, 1980.

[68] J. J. Sylvester. Note sur le théorème de legendre cité dans une note insérée dans les comptes rendus. *Comptes Rendus Acad. Sci. Paris*, 96:463–465, 1883.

[69] T. Szántai. A computer code for solution of probabilistic-constrained stochastic programming problems. *Numerical Techniques for Stochastic Optimization*, pages 229–235, 1988.

[70] L. Takács. On the method of inclusion and exclusion. *Journal of the American Statistical Association*, 62(317):102–113, 1967.

[71] I. Tomescu. Hypertrees and bonferroni inequalities. *Journal of Combinatorial Theory, Series B*, 41(2):209–217, 1986.

[72] C. van de Panne and W. Popp. Minimum-cost cattle feed under probabilistic protein constraints. *Management Science*, 9(3):405–430, 1963.

[73] M. Vasquez and J. K. Hao. A "logic-constrained" knapsack formulation and a tabu algorithm for the daily photograph scheduling of an earth observation satellite. *Computation Optimization and Applications*, 20(2):137–157, 2001.

[74] B. Vizvári. The integer programming background of a stochastic integer programming algorithm of Dentcheva-Prékopa-Ruszczyński. *Optimization Methods and Software*, 17(3):543–559, 2002.

[75] H. M. Weingartner. Capital budgeting of interrelated projects: survey and synthesis. *Management Science*, 12(7):485–516, 1966.

[76] C. Wilbaut, S. Hanafi, and S. Salhi. A survey of effective heuristics and their application to a variety of knapsack problems. *IMA Journal of Management Mathematics*, 19(3):227–244, 2008.

[77] K. J. Worsley. An improved bonferroni inequality and applications. *Biometrika*, 69(2):297–302, 1982.

[78] K. Yoda and A. Prékopa. Optimal portfolio selection based on multiple Value at Risk constraints. RUTCOR Research Report 12-2010, Rutgers Center for Operations Research, 2010.

[79] K. Yoda and A. Prékopa. Convexity and solutions of stochastic multidimensional knapsack problems with probabilistic constraints. *Mathematics of Operations Research*, accepted pending revision.

[80] K. Yoda and A. Prékopa. Improved bounds on the probability of the union of events, some intersections of which are empty. Working paper.

[81] S. Zymler, D. Kuhn, and B. Rustem. Distributionally robust joint chance constraints with second-order moment information. *Mathematical Programming*, 137(1-2):167–198, 2013.

# Vita

## Kunikazu Yoda

**2007-2013** Ph.D. in Operations Research, Rutgers University, New Brunswick, NJ

**1994-1996** M.Eng. in Applied Systems Science, Kyoto University, Kyoto, Japan

**1990-1994** B.Eng. in Applied Mathematics and Physics, Kyoto University, Kyoto, Japan

**2011-2013** Graduate assistant, RUTCOR, Rutgers University, New Brunswick, NJ

**2010-2011** Teaching assistant, Department of Mathematics, Rutgers University, New Brunswick, NJ

**1996-2007** Researcher, IBM Tokyo Research Laboratory, Yamato, Japan

**2013** K. Yoda and A. Prékopa, Improved bounds on the probability of the union of events, some intersections of which are empty (working paper).

**2013** K. Yoda and A. Prékopa, Convexity and solutions of stochastic multidimensional knapsack problems with probabilistic constraints, *Mathematics of Operations Research* (accepted pending revision).

**2011** A. Prékopa, K. Yoda, and M. M. Subasi, Uniform quasi-concavity in probabilistic constrained stochastic programming, *Operations Research Letters* (39) 188–192.

**2010** K. Yoda and A. Prékopa, Optimal portfolio selection based on multiple Value at Risk constraints, *RUTCOR Research Report* 12-2010.