RUTCOR
RESEARCH
REPORT

Convexity and Solutions of Stochastic Multidimensional Knapsack Problems with Probabilistic Constraints

Kunikazu Yoda [a]   András Prékopa [b]

RRR 30-2012, November 2012

RUTCOR
Rutgers Center for
Operations Research
Rutgers University
640 Bartholomew Road
Piscataway, New Jersey
08854-8003
Telephone: 732-445-3804
Telefax: 732-445-5472
Email: rrr@rutcor.rutgers.edu
http://rutcor.rutgers.edu/∼rrr

[a]RUTCOR, Rutgers University, 640 Bartholomew Road, Piscataway, NJ 08854-8003; kyoda@rutcor.rutgers.edu
[b]RUTCOR, Rutgers University, 640 Bartholomew Road, Piscataway, NJ 08854-8003; prekopa@rutcor.rutgers.edu

Rutcor Research Report
RRR 30-2012, November 2012

Convexity and Solutions of Stochastic Multidimensional Knapsack Problems with Probabilistic Constraints

Kunikazu Yoda   András Prékopa

Abstract. In the multidimensional knapsack problem a set of items, each with a value and a multidimensional size, is given and we want to select a subset of them in such a way that the total value of the selected items is maximized while the total size satisfies some capacity constraint for each dimension. In this paper we assume that the sizes are independent random variables such that each size follows the same type of probability distribution, not necessarily with the same parameter. A joint probabilistic constraint is imposed on the capacity constraints and the objective function is the same as that of the underlying deterministic problem. We showed that the problem is convex, under some condition on the parameters, for special continuous and discrete distributions: gamma, normal, Poisson, and binomial, where the latter two discrete distribution functions are approximated by logconcave continuous distribution functions.

Keywords: stochastic programming; probabilistic constraints; multidimensional knapsack problem; convexity

1 Introduction.

The knapsack problem is one of the most fundamental combinatorial optimization problems with 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 size and a value, 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 & Vygen [22]). One generalization of the problem includes the multidimensional knapsack problem (or multiply constrained knapsack problem), where each item has multiple dimensions of sizes, such as length, volume, weight, etc., and a knapsack has a capacitiy for each dimension. This variant with a fixed dimension (≥ 2) was shown to be $\mathcal{NP}$-complete, and more strongly, has no FPTAS unless $\mathcal{P} = \mathcal{NP}$ by Gens & Levner [13] and Korte & Schrader [21] (see also Kellerer et al. [20]). 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 paper, we study the multidimensional (single) knapsack problem and also the (one-dimensional) multiple knapsack problem where the item sizes 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. [20]. The deterministic model for the multidimensional knapsack problem has been studied extensively since the 1950s (see Fréville [9] and Fréville & Hanafi [10] 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 sizes, which can be considered as a special case for the general stochastic programming problem with probabilistic constraints of linear inequalities. 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 [19], van de Panne & Popp [29], and Prékopa [26]). For the random matrix with more than one row, the first paper where convexity theorems are presented is by Prékopa [25] and an important progress was made by Henrion & Strugarek [17]. A couple of works are known for the probabilistic constrained stochastic one-dimensional knapsack problem. Goyal and Ravi [15] showed a polynomial time approximation scheme via a parametric LP reformulation when the random item sizes are independent and normally distributed. Fortz and Poss [8] showed that the problem can be linearized when the random item sizes 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.

Applications of the multidimensional knapsack problem include, but are not limited to, cargo loading (see Bellman & Dreyfus [2]), cutting stock (see Gilmore & Gomory [14]), capital budgeting (see Lorie & Savage [23] and Weingartner [31]), project selection (see Petersen [24]), resource allocation in distributed data processing (see Gavish & Pirkul [12]), computer systems design (see Ferreira et al. [7]), daily management of a satellite (see Vasquez & Hao [30]), and combinatorial auctions (see de Vries & Vohra [3] and Rothkopf et al. [28]). See also a survey paper by Wilbaut et al. [32].

This paper is organized as follows. In section 2, we formulate the probabilistic constrained stochastic programming models for the multidimensional knapsack problem and the multiple knapsack problem. In section 3, we show convexity of a relaxed feasible set of the stochastic multidimensional knapsack problem for various distributions. Section 4 illustrates computational experiments.

2 Formulations of the problem.

2.1 Stochastic multidimensional knapsack problem with independent random item sizes.

First let us consider the deterministic problem. We are given a set of n items with values $v_1, v_2, \ldots, v_n$. Each item has m dimensions of sizes (such as width, length, weight, etc.) and we denote by $w_{ij} > 0$ the size of dimension i for item j. We have single knapsack with m size capacities $W_i > 0$, $i = 1, \ldots, m$ for the m dimensions, respectively. The goal is to select a subset of items (to be placed into the knapsack) maximizing the total value while keeping the size capacities, which can be formulated as follows:

$$\text{maximize} \sum_{j=1}^{n} v_j x_j \tag{2.1a}$$

$$\text{subject to} \sum_{j=1}^{n} w_{ij} x_j \le W_i \quad \text{for } i = 1, \ldots, m \tag{2.1b}$$

$$x_j \in \{0, 1\} \quad \text{for } j = 1, \ldots, n. \tag{2.1c}$$

Now suppose the item sizes are random variables denoted by ξ_{ij}'s (in place of w_{ij}'s), and formulate the problem as a probabilistic constrained stochastic programming, where constraints (2.1b) are replaced by the following joint probabilistic constraint:

$$\mathbb{P}\left(\sum_{j=1}^{n} \xi_{ij} x_j \le W_i \quad \text{for } i = 1, \ldots, m \right) \ge q. \tag{2.2}$$

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

Assuming the random variables ξ_{ij}'s are independent, the joint probabilistic constraint

(2.2) can be written as follows:

$$\prod_{i=1}^{m} \mathbb{P}\left(\sum_{j=1}^{n} \xi_{ij} x_j \leq W_i\right) \geq q. \tag{2.3}$$

Here we only need independence on the random vectors $(\xi_{i1}, \xi_{i2}, \ldots, \xi_{in})$, $i = 1, \ldots, m$, that is, the sizes of different dimensions in the same item are independent but the sizes of different items in the same dimension may be dependent. Note that even if we can't assume the random vectors 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. [6]):

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

Then the probabilistic constraint (2.3) ensures the joint probabilistic constaint (2.2). The probabilistic constrained stochastic programming model for the multidimensional knapsack problem with independent random item sizes is formulated as follows:

$$\text{maximize} \sum_{j=1}^{n} v_j x_j \tag{2.4a}$$

$$\text{subject to} \prod_{i=1}^{m} \mathbb{P}\left(\sum_{j=1}^{n} \xi_{ij} x_j \leq W_i\right) \geq q \tag{2.4b}$$

$$x_j \in \{0, 1\} \text{ for } j = 1, \ldots, n. \tag{2.4c}$$

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

$$F_i(\boldsymbol{x}) := \mathbb{P}\left(\sum_{j=1}^{n} \xi_{ij} x_j \leq W_i\right). \tag{2.5}$$

The feasible set of the problem is as follows:

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

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

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

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

Lemma 2.1. *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 decreasing (resp. increasing) logconcave function of $t \in \mathbb{R}$ and $h(\boldsymbol{x})$ is a convex (resp. concave) function of $\boldsymbol{x} \in \mathbb{R}^n$.*

In all of our cases studied in this paper, $F_i(\boldsymbol{x})$ 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 decreasing (or increasing) and logconcave to ensure the logconcavity of $F_i(x)$ and hence the convexity of the set $\mathcal{S}$.

2.2 Stochastic multiple knapsack problem with independent random item sizes.

Here is another generalizaion of the knapsack problem. First let us consider the deterministic problem. We are given a set of n items, each of which has value v_j and size $w_j > 0$ for $j = 1, \dots, n$. We have m knapsacks, each of which has size capacity $W_i > 0$ for $i = 1, \dots, m$. The goal is to select a subset of items (each to be placed into one of the knapsacks) maximizing the total value while keeping the size capacities, which can be formulated as follows:

$$\text{maximize} \ \sum_{i=1}^{m} \sum_{j=1}^{n} v_j x_{ij} \tag{2.6a}$$

$$\text{subject to} \ \sum_{j=1}^{n} w_j x_{ij} \le W_i \quad \text{for } i = 1, \dots, m \tag{2.6b}$$

$$\sum_{i=1}^{m} x_{ij} \le 1 \quad \text{for } j = 1, \dots, n \tag{2.6c}$$

$$x_{ij} \in \{0, 1\} \quad \text{for } i = 1, \dots, m \text{ and } j = 1, \dots, n. \tag{2.6d}$$

Now suppose the item sizes are random variables denoted by ξ_j's (in place of w_j's), and formulate the problem as a probabilistic constrained stochastic programming, where constraints (2.6b) are replaced by the following joint probabilistic constraint:

$$\mathbb{P}\left(\sum_{j=1}^{n} \xi_j x_{ij} \le W_i \quad \text{for } i = 1, \dots, m\right) \ge q. \tag{2.7}$$

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

Assuming the random variables ξ_j's are independent, the joint probabilistic constraint (2.7), where a set of ξ_j's in the sum are disjoint for $i = 1, \dots, m$ because of (2.6c), can be written as follows:

$$\prod_{i=1}^{m} \mathbb{P}\left(\sum_{j=1}^{n} \xi_j x_{ij} \le W_i\right) \ge q.$$

The probabilistic constrained stochastic programming model for the multiple knapsack problem with independent random item sizes is formulated as follows:

$$\begin{aligned}
\text{maximize} \quad & \sum_{i=1}^{m}\sum_{j=1}^{n} v_j x_{ij} \\
\text{subject to} \quad & \prod_{i=1}^{m} \mathbb{P}\left(\sum_{j=1}^{n} \xi_j x_{ij} \le W_i\right) \ge q \\
& \sum_{i=1}^{m} x_{ij} \le 1 \quad \text{for } j = 1, \ldots, n \\
& x_{ij} \in \{0, 1\} \quad \text{for } i = 1, \ldots, m \text{ and } j = 1, \ldots, n.
\end{aligned}$$

Let us denote $\boldsymbol{x} = (x_{11}, \ldots, x_{m1}, \ldots, x_{1n}, \ldots, x_{mn})$ and $F_i(\boldsymbol{x}) := \mathbb{P}\left(\sum_{j=1}^{n} \xi_j x_{ij} \le W_i\right)$. Then the relaxed feasible set of the problem is as follows:

$$\mathcal{S} = \left\{ \boldsymbol{x} \in \mathbb{R}^{mn} \;\middle|\; \prod_{i=1}^{m} F_i(\boldsymbol{x}) \ge q, \right.$$
$$\left. \sum_{i=1}^{m} x_{ij} \le 1 \text{ and } x_{ij} \in [0, 1] \quad \text{for } i = 1, \ldots, m \text{ and } j = 1, \ldots, n. \right\}.$$

The set $\mathcal{S}$ is convex if $F_i(\boldsymbol{x})$ is logconcave for every $i \in \{1, \ldots, m\}$. The function $F_i(\boldsymbol{x})$ has the same form with (2.5) in terms of analyzing logconcavity, and hence we only deal with the stochastic multidimensional knapsack problem.

3 Convexity of the stochastic multidimensional knapsack problem for various distributions.

3.1 Convexity result for the gamma distribution.

In this section we consider a case where the independent random variable ξ_{ij} in (2.4) has the gamma distribution

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

with shape $p_{ij} > 0$ and scale $\theta_i > 0$ for $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$. Note that $\text{Var}(\xi_{ij}) = \theta_i \, \mathbb{E}(\xi_{ij}) = p_{ij}\theta_i^2$. 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 θ_i and thus has the gamma distribution

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

where we defined $p_i(\boldsymbol{x}) := \sum_{j=1}^n p_{ij}x_j$.

Let $\Gamma(p)$ denote the Gamma function:

$$\Gamma(p) := \int_0^\infty t^{p-1}e^{-t}dt \quad \text{for } p > 0. \tag{3.1}$$

Let $P(p,\lambda)$ denote the lower regularized Gamma function:

$$P(p,\lambda) := \int_0^\lambda \frac{t^{p-1}e^{-t}}{\Gamma(p)}dt \quad \text{for } p \geq 0,\ \lambda \geq 0. \tag{3.2}$$

For any fixed $\lambda > 0$, we define $P(0,\lambda) = \lim_{p\to 0+} P(p,\lambda) = 1$. Since $\mathbb{P}(\sum_{j=1}^n \xi_{ij}x_j \leq W_i) = P(p_i(\boldsymbol{x}), W_i/\theta_i)$, which also holds for $\boldsymbol{x} = 0$, the stochastic multidimensional knapsack problem can be formulated as follows:

$$\text{maximize} \sum_{j=1}^n v_j x_j \tag{3.3a}$$

$$\text{subject to} \prod_{i=1}^m P\left(p_i(\boldsymbol{x}), \frac{W_i}{\theta_i}\right) \geq q \tag{3.3b}$$

$$x_j \in \{0,1\} \quad \text{for } j = 1,\dots,n. \tag{3.3c}$$

A relaxed feasible set of the problem is expressed as follows:

$$\mathcal{S} = \left\{ \boldsymbol{x} \in \mathbb{R}^n \;\middle|\; \prod_{i=1}^m P(p_i(\boldsymbol{x}), W_i/\theta_i) \geq q,\ x_j \in [0,1] \ \text{ for } j = 1,\dots,n \right\}. \tag{3.4}$$

The following theorem together with Lemma 2.1 ensures the convexity of $\mathcal{S}$.

Theorem 3.1. *For any fixed $\lambda > 0$, the function $P(p,\lambda)$ defined by* (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 [26] 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. Simple calculation shows that for $p > 0$

$$\begin{aligned}\frac{d^2}{dp^2}\ln P(p,\lambda) = {} & \frac{\int_0^\lambda t^{p-1}(\ln t)^2 e^{-t}dt}{\int_0^\lambda t^{p-1}e^{-t}dt} - \left(\frac{\int_0^\lambda t^{p-1}(\ln t)e^{-t}dt}{\int_0^\lambda t^{p-1}e^{-t}dt}\right)^2 \\ & - \left[\frac{\int_0^\infty t^{p-1}(\ln t)^2 e^{-t}dt}{\int_0^\infty t^{p-1}e^{-t}dt} - \left(\frac{\int_0^\infty t^{p-1}(\ln t)e^{-t}dt}{\int_0^\infty t^{p-1}e^{-t}dt}\right)^2\right].\end{aligned} \tag{3.5}$$

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 (3.5) can be written as

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

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 [27] that

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

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

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

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

Corollary 3.1. *The relaxed feasible set $\mathcal{S}$ defined by* (3.4) *of the stochastic multidimensional knapsack problem* (3.3) *is convex.*

3.2 Convexity result for the normal distribution.

In this section we consider a case where the independent random variable ξ_{ij} in (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

$$\frac{4W_i}{\lambda_i} \ge -\gamma_i^2 + \left[\frac{1 + \sqrt{1 + \gamma_i(\gamma_i + \varphi(\gamma_i)/\Phi(\gamma_i))}}{\gamma_i + \varphi(\gamma_i)/\Phi(\gamma_i)}\right]^2 \tag{3.6}$$

where

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

for $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$. Here we denoted by $\varphi(t)$ and $\Phi(t)$, the p.d.f. and the c.d.f., respectively, of the standard normal distribution:

$$\varphi(t) := \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right), \quad \Phi(t) := \int_{-\infty}^{t} \varphi(u) du. \tag{3.7}$$

Note that $\mathrm{Var}(\xi_{ij}) = \lambda_i \, \mathbb{E}(\xi_{ij}) = \lambda_i \mu_{ij}$. The condition (3.6) is satisfied if ν_i is smaller than the threshold determined by W_i and λ_i (see (3.21)). 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:

$$F_i(\mu) := \Phi\left(\frac{W_i - \mu}{\sqrt{\lambda_i \mu}}\right) \quad \text{for } \mu \geq 0. \tag{3.8}$$

We define $F_i(0) = \lim_{\mu \to 0+} F_i(\mu) = 1$. Since $\mathbb{P}(\sum_{j=1}^{n} \xi_{ij} x_j \leq W_i) = F_i(\mu_i(\boldsymbol{x}))$, which also holds for $\boldsymbol{x} = 0$, the stochastic multidimensional knapsack problem can be formulated as follows:

$$\text{maximize} \sum_{j=1}^{n} v_j x_j \tag{3.9a}$$

$$\text{subject to} \prod_{i=1}^{m} F_i(\mu_i(\boldsymbol{x})) \geq q \tag{3.9b}$$

$$x_j \in \{0, 1\} \quad \text{for } j = 1, \ldots, n. \tag{3.9c}$$

A relaxed feasible set of the problem is expressed as follows:

$$\mathcal{S} = \left\{ \boldsymbol{x} \in \mathbb{R}^n \;\middle|\; \prod_{i=1}^{m} F_i(\mu_i(\boldsymbol{x})) \geq q, \; x_j \in [0, 1] \;\text{ for } j = 1, \ldots, n \right\}. \tag{3.10}$$

The following lemma about $\varphi(x)/\Phi(x)$ is required to prove the logconcavity of $F_i(\mu)$. Most of 1 and 2 are well-known, but we present them for completeness.

Lemma 3.1. *Let us denote $\rho(x) := \varphi(x)/\Phi(x)$. For $x \in \mathbb{R}$, we have the following:*

1. *$\rho(x)$ is positive, strictly decreasing, strictly logconcave, and strictly convex.* $\lim_{x\to-\infty} \rho(x) = \infty$. $\lim_{x\to\infty} \rho(x) = 0$.

2. *$x + \rho(x)$ is positive and strictly increasing.* $\lim_{x\to-\infty}(x + \rho(x)) = 0$. $\lim_{x\to\infty}(x + \rho(x)) = \infty$.

3. *$\rho(x)(x + \rho(x)) \in (0, 1)$ is strictly decreasing.* $\lim_{x\to-\infty} \rho(x)(x + \rho(x)) = 1$. $\lim_{x\to\infty} \rho(x)(x + \rho(x)) = 0$.

4. *$1/(x + \rho(x))$ is positive, strictly decreasing, and strictly convex.*

Proof. 1 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 2 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 3 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.

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

$$x + \rho(x) = \frac{x\Phi(x) + \varphi(x)}{\Phi(x)} = \frac{\int_{-\infty}^{x}(t\Phi(t) + \varphi(t))'dt}{\Phi(x)} = \frac{\int_{-\infty}^{x}\Phi(t)dt}{\Phi(x)} = \frac{1}{\left(\ln \int_{-\infty}^{x}\Phi(t)dt\right)'}. \tag{3.11}$$

Since $\int_{-\infty}^{x}\Phi(t)dt$ is an integral of a strictly logconcave distribution function, it is strictly logconcave. Hence the denominator of (3.11) is a strictly decreasing function. The proof of the limits are easily derived by l'Hôpital's rule.

3 First we prove

$$0 < \rho(x)(x + \rho(x)) < 1. \tag{3.12}$$

The left inequality is obvious from the positivities in 1 and 2. Since $x + \rho(x)$ is strictly increasing by 2, 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:

$$\begin{aligned} f(x) &:= -\frac{1}{\rho(x)}(\rho(x)(x + \rho(x)))' \\ &= (x + \rho(x))^2 + \rho(x)(x + \rho(x)) - 1 > 0 \quad \text{for } x \in \mathbb{R}. \end{aligned} \tag{3.13}$$

Either $f(x) > 0$ for all $x \in \mathbb{R}$ or there exists $x_0 \in \mathbb{R}$ such that $f(x_0) \le 0$. We show the latter case doesn't hold by contradiction. Assume it holds. By the definition of $f(x)$ in (3.13) we have the following implication:

$$\begin{aligned} & f'(x) = 2(x+\rho(x))(1-\rho(x)(x+\rho(x))) - \rho(x)((x+\rho(x))^2 + \rho(x)(x+\rho(x)) - 1) \le 0 \\ & \Rightarrow \ f(x) \ge \frac{2}{\rho(x)}(x+\rho(x))(1-\rho(x)(x+\rho(x))) > 0 \quad (\because \text{positivities in 1 and 2, (3.12)}), \end{aligned} \tag{3.14}$$

which implies that $f'(x_0) > 0$. Since $f(x)$ is continuously differentiable, there exists $x_1 < x_0$ such that $f(x_1) < f(x_0) \le 0$. 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 (3.14).

4 The positivity and the decreasing property follows directly from 2. To prove the strict convexity we show the following inequality:

$$\begin{aligned} g(x) &:= (x+\rho(x))^3 \left(\frac{1}{x+\rho(x)}\right)'' \\ &= 2(1-\rho(x)(x+\rho(x)))^2 - \rho(x)(x+\rho(x))f(x) \\ &= 2(1-\rho(x)\eta(x))^2 - \rho(x)\eta(x)f(x) \ > 0 \quad \text{for } x \in \mathbb{R}, \end{aligned} \tag{3.15}$$

where we defined $\eta(x) := x + \rho(x)$ and used $f(x) = \eta(x)^2 + \rho(x)\eta(x) - 1$ defined in (3.13). Either $g(x) > 0$ for all $x \in \mathbb{R}$ or there exists $x_0 \in \mathbb{R}$ such that $g(x_0) \le 0$. We show the latter case doesn't hold by contradiction. Assume it holds. By the definition of $g(x)$ in (3.15) we have the following implication:

$$\begin{aligned} & g' = \rho\left[4(1-\rho\eta)f + f^2 - 2\eta^2(1-\rho\eta) + \rho\eta f\right] \le 0 \\ & \Rightarrow \ g \ge 2(1-\rho\eta)^2 + 4(1-\rho\eta)f - 2\eta^2(1-\rho\eta) + f^2 \\ & \qquad = 2(1-\rho\eta)f + f^2 > 0 \quad (\because (3.12), (3.13)), \end{aligned} \tag{3.16}$$

which implies that $g'(x_0) > 0$. Since $g(x)$ is continuously differentiable, there exists $x_1 < x_0$ such that $g(x_1) < g(x_0) \le 0$. 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 (3.16). □

The following theorem together with Lemma 2.1 ensures the convexity of $\mathcal{S}$.

Theorem 3.2. *The function $F_i(\mu)$ defined by* (3.8) *is strictly decreasing for $\mu \ge 0$. It is logconcave for $\mu \in [0, \nu_i]$ if and only if the condition* (3.6) *is satisfied.*

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

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

where $a, b, \nu > 0$ for simplicity of the notation. We define $F(0) = \lim_{x\to0+} F(x) = 1$. Note that $F_i(\mu) = F(\mu)$ with $a = W_i/\sqrt{\lambda_i}$, $b = 1/\sqrt{\lambda_i}$, and $\nu = \nu_i$.

Let us denote

$$g(x;a,b) := ax^{-1/2} - bx^{1/2} \quad \text{and} \quad \rho(z) := \frac{\varphi(z)}{\Phi(z)}. \tag{3.17}$$

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 $x \in (0,\nu]$,

$$\begin{aligned} 0 \geq (\ln F(x))'' = \left(\frac{\varphi(g)g'}{\Phi(g)}\right)' = (\rho(g)g')' = \rho(g)\left[g'' - (g+\rho(g))(g')^2\right] \\ \Leftrightarrow g + \rho(g) \geq \frac{g''}{(g')^2} = \frac{x^{\frac{1}{2}}(3a+bx)}{(a+bx)^2}. \end{aligned} \tag{3.18}$$

By solving the quadratic equation

$$g = ax^{-1/2} - bx^{1/2} \text{ (in (3.17))} \;\Leftrightarrow\; b(x^{1/2})^2 + gx^{1/2} - a = 0$$

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

$$x^{\frac{1}{2}} = \frac{\sqrt{g^2+4ab} - g}{2b}. \tag{3.19}$$

By using (3.19), the condition (3.18) 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

$$\sqrt{g^2+4ab} \leq \frac{1 - \sqrt{1 + g(g+\rho(g))}}{g+\rho(g)} \tag{3.20a}$$

or

$$\sqrt{g^2+4ab} \geq \frac{1 + \sqrt{1 + g(g+\rho(g))}}{g+\rho(g)} \tag{3.20b}$$

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 3.1 3. No a, b exists under (3.20a), for instance when $g(g+\rho(g)) > 0$, which holds for sufficiently small $x > 0$. Since both sides of (3.20b) are positive, the condition (3.20b) can be expressed as

$$4ab \geq -g^2 + \left[\frac{1}{g+\rho(g)} + \sqrt{h(g)}\right]^2 = -g^2 + \frac{1}{(g+\rho(g))^2} + \frac{2\sqrt{h(g)}}{g+\rho(g)} + h(g). \tag{3.21}$$

Here we defined

$$\begin{aligned} h(z) &:= \frac{1+z(z+\rho(z))}{(z+\rho(z))^2} = 1 + \frac{1-\rho(z)(z+\rho(z))}{(z+\rho(z))^2} \quad (>0 \text{ by (3.12) in lemma 3.1 3)} \\ &= 1 - \left(\frac{1}{z+\rho(z)}\right)', \end{aligned}$$

which is decreasing for $z \in \mathbb{R}$ by lemma 3.1 4. The sum $-g^2 + 1/(g+\rho(g))^2$ in (3.21) is decreasing in g because

$$\begin{aligned} -\frac{(z+\rho(z))^3}{2}\left(-z^2 + \frac{1}{(z+\rho(z))^2}\right)' &= (z+\rho(z))^4 - \rho(z)(z+\rho(z))^3 + \{1-\rho(z)(z+\rho(z))\} \\ &> (z+\rho(z))^4 - \{1-\rho(z)(z+\rho(z))\}^2 \quad (\because \text{(3.15) in lemma 3.1 4)} \\ &> 0. \quad (\because \text{(3.12),(3.13) in lemma 3.1 2,3)} \end{aligned}$$

By lemma 3.1 2, $g + \rho(g)$ is increasing in g. Thus the right-hand side of (3.21) is decreasing for $g \in \mathbb{R}$ and hence it is increasing for $x > 0$, regardless of a and b. So the condition (3.18) for a and b is equivalent to

$$4ab \geq -g(\nu;a,b)^2 + \left[\frac{1+\sqrt{1+g(\nu;a,b)\{g(\nu;a,b)+\rho(g(\nu;a,b))\}}}{g(\nu;a,b)+\rho(g(\nu;a,b))}\right]^2. \tag{3.22}$$

Therefore $F(x)$ is logconcave for $x \in [0,\nu]$ under the condition for a and b in (3.22). □

Corollary 3.2. *The relaxed feasible set $\mathcal{S}$ defined by* (3.10) *of the stochastic multidimensional knapsack problem* (3.9) *is convex.*

3.3 Convexity result for the Poisson distribution.

In this section we consider a case where the independent random variable ξ_{ij} in (2.4) has the Poisson distribution

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

with parameter $\lambda_{ij} > 0$ for $i \in \{1,\ldots,m\}$ and $j \in \{1,\ldots,n\}$. Note that $\text{Var}(\xi_{ij}) = \mathbb{E}(\xi_{ij}) = \lambda_{ij}$. 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

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

where we defined $\lambda_i(\boldsymbol{x}) := \sum_{j=1}^n \lambda_{ij} x_j$.

Since $\mathbb{P}(\sum_{j=1}^n \xi_{ij} x_j \le W_i) = \sum_{k=0}^{\lfloor W_i \rfloor} ([\lambda_i(\boldsymbol{x})]^k / k!) \exp(-\lambda_i(\boldsymbol{x}))$, which also holds for $\boldsymbol{x} = 0$, the stochastic multidimensional knapsack problem can be formulated as follows:

$$\begin{aligned} \text{maximize } & \sum_{j=1}^n v_j x_j \\ \text{subject to } & \prod_{i=1}^m \sum_{k=0}^{\lfloor W_i \rfloor} \frac{[\lambda_i(\boldsymbol{x})]^k}{k!} e^{-\lambda_i(\boldsymbol{x})} \ge q \\ & x_j \in \{0,1\} \quad \text{for } j = 1, \dots, n. \end{aligned}$$

Let $Q(p, \lambda)$ denote the upper regularized gamma function:

$$Q(p,\lambda) := 1 - P(p,\lambda) = \int_\lambda^\infty \frac{t^{p-1} e^{-t}}{\Gamma(p)} dt \quad \text{for } p \ge 0,\ \lambda \ge 0. \tag{3.23}$$

Here $\Gamma(p)$ and $P(p, \lambda)$ are defined by (3.1) and (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. Thus we can rewrite the problem as follows:

$$\text{maximize } \sum_{j=1}^n v_j x_j \tag{3.24a}$$

$$\text{subject to } \prod_{i=1}^m Q(\lfloor W_i \rfloor + 1, \lambda_i(\boldsymbol{x})) \ge q \tag{3.24b}$$

$$x_j \in \{0,1\} \quad \text{for } j = 1, \dots, n. \tag{3.24c}$$

A relaxed feasible set of the problem is expressed as follows:

$$\mathcal{S} = \left\{ \boldsymbol{x} \in \mathbb{R}^n \;\middle|\; \prod_{i=1}^m Q(\lfloor W_i \rfloor + 1, \lambda_i(\boldsymbol{x})) \ge q,\ x_j \in [0,1],\ j = 1, \dots, n \right\}. \tag{3.25}$$

The following theorem together with Lemma 2.1 ensures the convexity of $\mathcal{S}$.

Theorem 3.3. *For any fixed $p \ge 1$, the function $Q(p, \lambda)$ defined by* (3.23) *is strictly decreasing and logconcave for $\lambda \ge 0$.*

Proof. First we prove the decreasing property. We have for $\lambda > 0$,

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

Thus $Q(p, \lambda)$ 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 [26] 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 $\lambda \geq 0$. □

Corollary 3.3. *The relaxed feasible set* $\mathcal{S}$ *defined by* (3.25) *of the stochastic multidimensional knapsack problem* (3.24) *is convex.*

3.4 Convexity result for the binomial distribution.

In this section we consider a case where the independent random variable ξ_{ij} in (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 $i \in \{1, \ldots, m\}$ and $j \in \{1, \ldots, n\}$. 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 $n_i(\boldsymbol{x}) := \sum_{j=1}^n n_{ij} x_j$.

Since $\mathbb{P}\left(\sum_{j=1}^n \xi_{ij} x_j \leq W_i\right) = \sum_{k=0}^{\min(\lfloor W_i \rfloor, n_i(\boldsymbol{x}))} \binom{n_i(\boldsymbol{x})}{k} {p_i}^k (1 - p_i)^{n_i(\boldsymbol{x}) - k}$, which also holds for $\boldsymbol{x} = 0$, the stochastic multidimensional knapsack problem can be formulated as follows:

$$\begin{aligned} \text{maximize} \; & \sum_{j=1}^n v_j x_j \\ \text{subject to} \; & \prod_{i=1}^m \sum_{k=0}^{\min(\lfloor W_i \rfloor, n_i(\boldsymbol{x}))} \binom{n_i(\boldsymbol{x})}{k} {p_i}^k (1 - p_i)^{n_i(\boldsymbol{x}) - k} \geq q \\ & x_j \in \{0, 1\} \quad \text{for } j = 1, \ldots, n. \end{aligned}$$

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:

$$J(z; c, p) := \begin{cases} I(1-p; z-c, c+1) & \text{for } z > c \\ 1 & \text{for } z \le c \end{cases}, \tag{3.26}$$

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

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

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. Thus we can rewrite the problem as follows:

$$\text{maximize} \sum_{j=1}^{n} v_j x_j \tag{3.27a}$$

$$\text{subject to} \prod_{i=1}^{m} J(n_i(\boldsymbol{x}); \lfloor W_i \rfloor, p_i) \ge q \tag{3.27b}$$

$$x_j \in \{0, 1\} \quad \text{for } j = 1, \ldots, n. \tag{3.27c}$$

A relaxed feasible set of the problem is expressed as follows:

$$\mathcal{S} = \left\{ \boldsymbol{x} \in \mathbb{R}^n \;\middle|\; \prod_{i=1}^{m} J(n_i(\boldsymbol{x}); \lfloor W_i \rfloor, p_i) \ge q,\ x_j \in [0, 1],\ j = 1, \ldots, n \right\}. \tag{3.28}$$

The following theorem together with Lemma 2.1 ensures the convexity of $\mathcal{S}$.

Theorem 3.4. *For any fixed $c \ge 0$ and $p \in (0, 1)$, the function $J(z; c, p)$ defined by (3.26) 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:

$$\begin{aligned} \frac{dJ(z; c, p)}{dz} &= \frac{d}{dz} \frac{\int_p^1 f(y, z; c)dy}{\int_0^1 f(y, z; c)dy} \\ &= J(z; c, p) \left(\frac{\int_p^1 f(y, z; c) \ln(1-y)dy}{\int_p^1 f(y, z; c)dy} - \frac{\int_0^1 f(y, z; c) \ln(1-y)dy}{\int_0^1 f(y, z; c)dy} \right). \end{aligned} \tag{3.29}$$

The derivative with respect to p of the first term in the parenthesis in (3.29) is

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

Thus the first term in the parenthesis in (3.29) 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:

$$\begin{aligned}
&\frac{d^2}{dz^2}\ln J(z;c,p) =\\
&\frac{\int_p^1 f(y,z;c)\left\{\ln(1-y)\right\}^2 dy}{\int_p^1 f(y,z;c)dy} - \left(\frac{\int_p^1 f(y,z;c)\ln(1-y)dy}{\int_p^1 f(y,z;c)dy}\right)^2\\
&-\left[\frac{\int_0^1 f(y,z;c)\left\{\ln(1-y)\right\}^2 dy}{\int_0^1 f(y,z;c)dy} - \left(\frac{\int_0^1 f(y,z;c)\ln(1-y)dy}{\int_0^1 f(y,z;c)dy}\right)^2\right].
\end{aligned} \tag{3.30}$$

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

$$g(x) = \frac{(1-e^{-x})^c e^{-x(z-c)}}{\int_0^\infty (1-e^{-y})^c e^{-y(z-c)}dy} \quad \text{for } x \geq 0, \tag{3.31}$$

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 (3.30) can be written as

$$\begin{aligned}
&\frac{\int_{\ln 1/(1-p)}^\infty x^2 g(x)dx}{\int_{\ln 1/(1-p)}^\infty g(x)dx} - \left(\frac{\int_{\ln 1/(1-p)}^\infty xg(x)dx}{\int_{\ln 1/(1-p)}^\infty g(x)dx}\right)^2 - \left[\frac{\int_0^\infty x^2 g(x)dx}{\int_0^\infty g(x)dx} - \left(\frac{\int_0^\infty xg(x)dx}{\int_0^\infty g(x)dx}\right)^2\right]\\
&= \mathbb{E}(X^2|X \geq \ln 1/(1-p)) - \mathbb{E}^2(X|X \geq \ln 1/(1-p)) - \left[\mathbb{E}(X^2) - \mathbb{E}^2(X)\right]\\
&= \mathbb{E}(X^2|X \geq v) - \mathbb{E}^2(X|X \geq v) - \left[\mathbb{E}(X^2) - \mathbb{E}^2(X)\right]. \text{ (We denote } v := \ln 1/(1-p))
\end{aligned}$$

Since the random variable X has a continuous and logconcave p.d.f. $g(x)$, it follows by theorem 2.1 in Prékopa [27] 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 have proved that for any $\mu \in (0,1)$, the following inequality for logconcave functions:

$$J(\mu z_1 + (1-\mu)z_2; c,p) \geq J(z_1;c,p)^{\mu} J(z_2;c,p)^{1-\mu} \tag{3.32}$$

holds for any $z_1, z_2 \geq c$. We have two other cases for z_1 and z_2. For any $z_1, z_2 < c$, the equality clearly holds (both sides are 1) for (3.32). 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 $z \in \mathbb{R}$. □

Corollary 3.4. *The relaxed feasible set $\mathcal{S}$ defined by* (3.28) *of the stochastic multidimensional knapsack problem* (3.27) *is convex.*

4 Computational examples.

First we carried out experiments to measure the CPU times to compute solutions of the problems, which are convex mixed-integer nonlinear programming (MINLP) problems. There are number of open source and commercial solvers for MINLP available today. We used GAMS [11] as a modeling system with DICOPT [16] as a MINLP solver. We selected CPLEX [18] as a mixed-integer linear programming (MILP) solver and CONOPT [4, 5] 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 has also support for derivatives of those functions used by NLP solvers. DICOPT is an implementation of the outer approximation method primary designed to solve convex MINLP, which is suitable for our problems. The experiments were run on an Intel Core-i3 processor clocked at 3.06 GHz with 4 GB RAM.

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 model. Each instance in the original data consists of:

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

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

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

as those in the original data and the common probability level $q = 0.9$ for all instances. In preprocessing we set every w_{ij} that is 0 in the original data to 0.1 to avoid singularities in nonlinear functions and we changed the scales of w_{ij} and W_i down in the case of gamma distribution to avoid errors for large derivatives. The parameters for the distributions of random item size ξ_{ij} were generated from w_{ij} as follows.

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

The mean $\mathbb{E}[\xi_{ij}]$ is equal to w_{ij} (approximately in the case of binomial). The variance $\text{Var}[\xi_{ij}]$ is 5 to 10% of the mean. For each original instance we generated four test instances corresponding to gamma, normal, Poisson, and binomial, where all items follow the same type of distribution in each test instance. The result is shown in Table 1. The first column shows the names of the original instances and the last four columns show the CPU times and the iteration counts (the median of three runs) reported by the MINLP solver for the four test instances.

Next 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 ξ_{1j} consumed for the resource 1 follows the gamma distribution with shape p_{1j} and scale θ_1. The random amount ξ_{2j} consumed for the resource 2 follows the normal distribution with mean μ_{2j} and variance $\lambda_2 \mu_{2j}$. The random amount ξ_{3j} consumed for the resource 3 follows the Poisson distribution with parameter λ_{3j}. The random amount ξ_{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 ξ_{ij} follow the same type of distribution in the same resource but follow different types of distributions in different resources. All random amounts ξ_{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 2. The parameters μ_{2j}, λ_2, and W_2 satisfy the condition (3.6). We want to find a subset of the projects that maximizes the total estimated profit while keeping the capacity constraints with a high probability. With

Table 1: CPU time for computing optimal solution

Instance name	n	m	CPU time (in seconds) / iteration count			
			gamma	normal	Poisson	binomial
mknap1 #1	6	10	0.109/ 37	0.094/ 20	0.093/ 40	0.810/ 3004
mknap1 #2	10	10	0.155/ 219	0.078/ 24	0.109/ 47	0.015/ 35
mknap1 #3	15	10	0.157/ 339	0.155/ 346	0.124/ 104	0.109/ 105
mknap1 #4	20	10	0.124/ 146	0.095/ 53	0.140/ 212	0.141/ 247
mknap1 #5	28	10	0.141/ 168	0.126/ 92	0.094/ 126	0.125/ 295
mknap1 #6	39	5	0.219/ 627	0.156/ 521	0.250/ 1393	0.358/ 1199
mknap1 #7	50	5	0.141/ 362	0.204/ 979	0.455/ 1617	0.155/ 423
mknap2 WEISH01	30	5	0.171/ 217	0.109/ 31	0.156/ 182	0.172/ 443
mknap2 WEISH06	40	5	0.125/ 108	0.094/ 32	0.125/ 208	0.124/ 195
mknap2 WEISH10	50	5	0.095/ 44	0.125/ 70	0.110/ 119	0.047/ 154
mknap2 PB4	29	2	0.157/ 245	0.077/ 99	0.094/ 48	0.093/ 58
mknap2 PB5	20	10	1.404/ 34684	0.218/ 704	4.790/ 114932	3.822/ 86126
mknap2 PB6	40	30	0.250/ 1178	0.376/ 3152	0.453/ 2475	0.437/ 1832
mknap2 HP1	28	4	0.312/ 1257	0.124/ 312	0.140/ 181	0.141/ 475
mknap2 HP2	35	4	0.404/ 2141	0.156/ 621	0.234/ 1214	0.110/ 513

Table 2: Parameters for the projects

Project j	Profit v_j	Resource 1 Gamma(p_{1j}, θ_1)		Resource 2 $\mathcal{N}(\mu_{2j}, \lambda_2\mu_{2j})$		Resource 3 Pois(λ_{3j})	Resource 4 $B(n_{4j}, p_4)$	
		p_{1j}	θ_1	μ_{2j}	λ_2	λ_{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		$W_1 = 150.0$		$W_2 = 203.0$		$W_3 = 78$	$W_4 = 10$	

the probability level $q = 0.9$, we can formulate the stochastic multidimensional knapsack problem as follows:

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

where functions P,Φ,Q,J are defined by (3.2),(3.7),(3.23),(3.26), 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 $\boldsymbol{x} = (1,0,0,1,0)^T$. So the best choice is to select the projects 1 and 4.

References

[1] J. E. Beasley, *OR-Library: Distributing test problems by electronic mail*, The Journal of the Operational Research Society **41** (1990), no. 11, 1069–1072.

[2] R. E. Bellman and S. E. Dreyfus, *Applied dynamic programming*, Princeton University Press, Princeton, NJ, 1962.

[3] S. de Vries and R. V. Vohra, *Combinatorial auctions: A survey*, INFORMS Journal on Computing **15** (2003), no. 3, 284–309.

[4] A. S. Drud, *CONOPT*, ARKI Consulting and Development A/S, Bagsvaerd, Denmark, Available in `http://www.gams.com`.

[5] ———, *CONOPT – A large-scale GRG code*, ORSA Journal on Computing **6** (1994), no. 2, 207–216.

[6] J. D. Esary, F. Proschan, and D. W. Walkup, *Association of random variables, with applications*, The Annals of Mathematical Statistics **38** (1967), no. 5, 1466–1474.

[7] C. E. Ferreira, M. Grötschel, R. Weismantel A. Martin, S. Kiefl, and L. Krispenz, *Some integer programs arising in the design of main frame computers*, Mathematical Methods of Operations Research **38** (1993), no. 1, 77–100.

[8] B. Fortz and M. Poss, *Easy distributions for combinatorial optimization problems with probabilistic constraints*, Operations Research Letters **38** (2010), no. 6, 545–549.

[9] A. Fréville, *The multidimensional 0-1 knapsack problem: An overview*, European Journal of Operational Research **155** (2004), no. 1, 1–21.

[10] A. Fréville and S. Hanafi, *The multidimensional 0-1 knapsack problem–bounds and computational aspects*, Annals of Operations Research **139** (2005), no. 1, 195–227.

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

[12] B. Gavish and H. Pirkul, *Allocation of databases and processors in a distributed computing system*, Management of Distributed Data Processing (North Holland, Amsterdam) (J. Akoka, ed.), 1982, pp. 215–231.

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

[14] P. C. Gilmore and R. E. Gomory, *The theory and computation of knapsack functions*, Operations Research **14** (1966), no. 6, 1045–1074.

[15] V. Goyal and R. Ravi, *A PTAS for the chance-constrained knapsack problem with random item sizes*, Operations Research Letters **38** (2010), no. 3, 161–164.

[16] 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`.

[17] R. Henrion and C. Strugarek, *Convexity of chance constraints with independent random variables*, Computational Optimization and Applications **41** (2008), no. 2, 263–276.

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

[19] S. Kataoka, *A stochastic programming model*, Econometrica **31** (1963), no. 1–2, 181–196.

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

[21] B. H. Korte and R. Schrader, *On the existence of fast approximation schemes*, Nonlinear Programming (O. L. Mangasarian, R. R. Meyer, and S. M. Robinson, eds.), vol. 4, Academic press, 1981, pp. 415–437.

[22] B. H. Korte and J. Vygen, *Combinatorial optimization: Theory and algorithms*, ch. 17, Springer-Verlag, 2008.

[23] J. H. Lorie and L. J. Savage, *Three problems in rationing capital*, The Journal of Business **28** (1955), no. 4, 229–239.

[24] C. C. Petersen, *Computational experience with variants of the Balas algorithm applied to the selection of R&D projects*, Management Science **13** (1967), no. 9, 736–750.

[25] A. Prékopa, *Programming under probabilistic constraints with a random technology matrix*, Mathematische Operationsforschung und Statistik, Ser. Optimization **5** (1974), no. 2, 109–116.

[26] ______, *Stochastic Programming*, Kluwer Academic Publishers, Dordrecht, 1995.

[27] ______, *Conditional mean-conditional variance portfolio selection model*, RUTCOR research report 34-2007, Rutgers Center for Operations Research, 2007.

[28] M. H. Rothkopf, A. Pekeč, and R. M. Harstad, *Computationally manageable combinational auctions*, Management Science **44** (1998), no. 8, 1131–1147.

[29] C. van de Panne and W. Popp, *Minimum-cost cattle feed under probabilistic protein constraints*, Management Science **9** (1963), no. 3, 405–430.

[30] 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** (2001), no. 2, 137–157.

[31] H. M. Weingartner, *Capital budgeting of interrelated projects: survey and synthesis*, Management Science **12** (1966), no. 7, 485–516.

[32] 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** (2008), no. 3, 227–244.