Easy distributions for combinatorial optimization problems with probabilistic constraints

Bernard Fortz

Department of Computer Science, Faculté des Sciences,
Université Libre de Bruxelles,
Brussels, Belgium.
CORE, Université catholique de Louvain,
Louvain-la-Neuve, Belgium.

Michael Poss*

Department of Computer Science, Faculté des Sciences,
Université Libre de Bruxelles,
Brussels, Belgium.

February 22, 2010

Abstract

We show how we can linearize probabilistic linear constraints with binary variables when all coefficients are distributed according to either $\mathcal{N}(\mu_i, \lambda\mu_i)$, for some $\lambda > 0$ and $\mu_i > 0$, or $\Gamma(k_i, \theta)$ for some $\theta > 0$ and $k_i > 0$. The constraint can also be linearized when the coefficients are independent and identically distributed if they are, besides, either positive or strictly stable random variables.

Keywords: Stochastic programming, combinatorial optimization, probabilistic constraint.

1 Introduction

Many combinatorial optimization models address problems with parameters which are impossible to predict exactly. Therefore, it is often more accurate to model these parameters with random variables. This modifies the structure of the optimization problems, depending on the times at which decisions are taken and parameters are revealed. In this note we study probabilistic constraints: all decisions must be taken here and now, such that the constraints of the model shall be satisfied with a certain probability. In other words, we aim at maximizing some objective for a given feasibility tolerance.

*Corresponding author. Email: mposs@ulb.ac.be. Adress: Boulevard du Triomphe CP 210 / 01, B-1050 Brussels, Belgium.

Stochastic programs with linear probabilistic constraints are in general non convex non linear optimization problems, see Henrion and Strugarek (2008) among others. If furthermore some variables are integer, they become non convex Mixed Integer Non Linear Problems, see Grossmann (2002), where the non linearities may not be convex. However, when the set of scenarios is finite and only the right-hand side is random, Luedtke et al. (2008) among others, linearize the problem with the help of integer variables, and study some valid inequalities. Although probabilistic constraints have been widely studied for many years, including practical modeling and theoretical results, see Henrion (2004); Shapiro et al. (2009), papers on combinatorial problems are not very numerous. Among them, Dentcheva et al. (2002) study different formulations and bounds for the problem, Beraldi and Ruszczynski (2002) propose a branch-and-bound algorithm when only the right hand side of constraints is random with joint probabilistic constraints, and Klopfenstein (In press) computes valid inequalities for the sets induced by the probabilistic constraints and tests them in a branch-and-cut algorithm.

In this work, we are particularly interested by the case of probabilistic constraints with binary variables while the random variables follow particular continuous distributions, among which normal distributions. Previous results in this direction assume that all random variables are normally distributed. In that case, the probabilistic constraints can be rewritten as quadratic constraints (see Kataoka (1963); Panne and Popp (1963); Prékopa (1995)), convex under some assumption on the confidence level (Parikh (1968)). If all variables are binary, the constraints can be further linearized using classical techniques (Hansen and Meyer (2009)). Further work extends the classical gaussian framework to the more general class of radial distributions, see Calafiore and Ghaoui (2006). The authors show how a probabilistic constraint can be written as a second-order cone convex constraint. The latter constraint can be linearized as well when working with binary variables.

In this note, we always assume that coefficients are independent continuous random variables. We show that a linear probabilistic constraint with binary variables is equivalent to a linear constraint when all coefficients are distributed according to either $\mathcal{N}(\mu_i, \lambda\mu_i)$, for some $\lambda > 0$ and $\mu_i > 0$, or $\Gamma(k_i, \theta)$ for some $\theta > 0$ and $k_i > 0$. The constraint can also be linearized when the coefficients are independent and identically distributed, if they are either positive or strictly stable random variables. The new right hand side may be difficult to compute for general random variables. As a result, we obtain that certain types of chance-constrained knapsack problems are as easy to solve as their deterministic counterpart, which is similar to the result obtained in Fortz et al. (2009) for the stochastic knapsack problem with simple linear recourse.

The next section describes precisely the constraints studied herein. Then, in Section 3 we study the case of identically distributed random variables, while in Section 4 we study gaussian and gamma random variables. Finally, Section 5 applies the results from Section 4 to the chance-constrained knapsack problem.

2 Studied constraints

In the following we study mainly the following type of probabilistic constraints,

$$\mathcal{C}_1(x) = P\left(\sum_{i=1}^{n} a_i x_i \leq b\right) \geq p, \tag{1}$$

though our results extend easily to

$$\mathcal{C}_2(x) = P\left(\sum_{i=1}^{n} a_i x_i \leq c_1 y_1 + b\right) \geq p, \tag{2}$$

and

$$\begin{aligned} \mathcal{C}_3(x) \;=\; P\left(\sum_{i=1}^{n} a_i x_i \leq \sum_{j=1}^{m} c_j y_j + b\right) \;&\geq\; p \\ \sum_{j=1}^{m} y_j \;&\leq\; 1, \end{aligned} \tag{3}$$

where $p \in (0,1)$, a_i are independent random variables, c_i and b are fixed coefficients. In addition, we always consider that $x_i, y_j \in \{0,1\}$, for $1 \leq i \leq n$ and $1 \leq j \leq m$. The first constraint (1) is the so-called knapsack constraint, which plays an important role in capacitated problems such as unsplittable multicommodity flow and generalized assignment problems. The second constraint (2) appears when the choice of the capacitated facilities to be built is part of the decision: b denotes the initial capacity and c_1 the capacity provided by the facility. Typical examples are network design problems and facility location problems. Finally, in many technical problems we must choose at most one out of a set of different facilities, for instance, different capacities for a new link to install in a telecommunication network. This is represented by (3).

In the sequel, we say that two constraints $C_1(x) \geq 0$ and $C_2(x) \geq 0$ are equivalent, denoted by $C_1(x) \geq 0 \Leftrightarrow C_2(x) \geq 0$, if the sets $\{x \in \{0,1\}^n \text{ s.t. } C_1(x) \geq 0\}$ and $\{x \in \{0,1\}^n \text{ s.t. } C_2(x) \geq 0\}$ are equal.

3 Identically distributed variables

We first consider (1) for the simple example where a_i are positive random variables identically distributed. We see that

$$P\left(\sum_{i=1}^{m} a_i \leq b\right) \leq P\left(\sum_{i=1}^{m-1} a_i \leq b\right). \tag{4}$$

Thus, the number of x_i that can be equal to 1 can certainly not exceed

$$N(b) = \max_{1 \leq l \leq n} \left\{ l \text{ s.t. } P\left(\sum_{i=1}^{l} a_i \leq b\right) \geq p \right\}. \tag{5}$$

Conversely, if some binary vector x satisfies $\sum_{i=1}^{n} x_i \leq N(b)$, then certainly x satisfies (1) because a_i are identically distributed. Then, considering (2), the previous reasoning holds with $N(b)$ for $x_n = 0$, and with $N(b + c_1)$ for $x_n = 1$. Finally, this reasoning extends to the pair of constraints (3), since at most one of the y_j can be equal to 1. We just proved the following:

Proposition 1. *Consider n independent identically distributed positive random variables a_i, $1 \leq i \leq n$. Then, for $x_i, y_j \in \{0,1\}$, $1 \leq i \leq n$ and $1 \leq j \leq m$, the following constraints are equivalent:*

1. $\mathcal{C}_1(x) \geq p \Leftrightarrow \sum_{i=1}^{n} x_i \leq N(b)$

2. $\mathcal{C}_2(x) \geq p \Leftrightarrow \sum_{i=1}^{n} x_i \leq (N(b+c_1) - N(b))y_1 + N(b)$

3. *If furthermore,* $\sum_{j=1}^{m} y_j \leq 1$*, then* $\mathcal{C}_3(x) \geq p \Leftrightarrow \sum_{i=1}^{n} x_i \leq \sum_{j=1}^{m} (N(c_j) - N(b))y_j + N(b)$

with $N(r)$ *defined in* (5) *for any real* r*.*

In the following, we focus on results of type *1.* since *2.* and *3.* can be deduced from *1.* by the above arguments.

Remark that computing the value of $N(b)$ requires, in general, the solution of a multivariate integral that must be solved using efficient packages for numerical integration, see Prékopa (1995). For some distributions, this computational burden can be avoided. For instance, if all a_i are uniformly distributed between 0 and 1, their sum is distributed according to (see for instance Grinstead and Snell (1997))

$$f(z) = \frac{1}{n!} \sum_{k=0}^{n+1} (-1)^k \begin{pmatrix} n+1 \\ k \end{pmatrix} [(z-k)_+]^n \,.$$

The uniform distributions are not the only distributions which sum up nicely. Stable distributions satisfy interesting summation properties too. Recall that if a_i are n independent copies of a stable random variable a, then for any constants x_i the random variable $\sum_{i=1}^{n} x_i a_i$ has the same distribution as $v_n\, a + w_n$ with some constants $v_n = n^{1/\alpha}$ for some $\alpha \in (0, 2]$, and w_n. Moreover, a is said strictly stable if $w_n = 0$ in the relation above. For instance, the Levy distribution, with density function equal to $f(z; c) = \sqrt{\frac{c}{2\pi}} \frac{e^{-c/2x}}{x^{3/2}}$ for $z \geq 0$, is stable and positive, so that it satisfies the hypothesis of Proposition 1. We refer to Nolan (2010) for a good introduction on stable distributions.

In general, the support of stable distributions intersect negative reals. For instance normal and Cauchy distributions always have negative tails. We show next that property (4) still holds for strictly stable distributions. By definition

$$\sum_{i=1}^{n} a_i \sim n^{1/\alpha} a_1 \qquad \alpha \in (0, 2],$$

so that

$$P\left(\sum_{i=1}^{n} a_i \leq b\right) = P(n^{1/\alpha} a_1 \leq b) = P(a_1 \leq b n^{-1/\alpha}).$$

If $b \geq 0$, the function $n \mapsto bn^{-1/\alpha}$ is non increasing, implying (4). We obtain the following:

Proposition 2. *Consider* n *independent identically distributed strictly stable random variables* a_i*,* $1 \leq i \leq n$*, and* $b \geq 0$*. Then, if* $x_i \in \{0, 1\}$ *for each* $1 \leq i \leq n$*, the following constraints are equivalent:*

$$\mathcal{C}_1(x) \geq p \Leftrightarrow \sum_{i=1}^{n} x_i \leq N(b),$$

with $N(b)$ *defined in* (5)*.*

An example of strictly stable distribution with $\alpha = 1$ is the Cauchy distribution, with density function $f(z; z_0, \gamma) = \frac{1}{\pi}\left(\frac{\gamma}{(z-z_0)^2+\gamma^2}\right)$ for some location parameter $z_0 \in \mathbb{R}$ and scale parameter $\gamma > 0$.

4 Non identically distributed variables

A well known stable distribution is the gaussian distribution. In fact, for gaussian and gamma random variables we are able to derive stronger results, allowing for the random variables to be distributed differently, as long as some regularity condition holds. Consider independent gaussian random variables, $a_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$, $1 \le i \le n$. Then, $\mathcal{C}_1(x) \ge p$ can be rewritten (see for instance Prékopa (1995))

$$\sum_{i=1}^{n} \mu_i x_i + \Phi^{-1}(p)\sqrt{\sum_{i=1}^{n} \sigma_i^2 x_i^2} \le b, \tag{6}$$

where Φ is the cumulative distribution of the standard normal distribution $\mathcal{N}(0,1)$. When $x \in \{0,1\}^n$, (6) can be linearized introducing additional continuous variables, see Hansen and Meyer (2009). However, these linearizations contain more variables and provide looser bounds than the direct linearization from Proposition (3) below.

Proposition 3. *Consider n random variables $a_i \sim \mathcal{N}(\mu_i, \lambda\mu_i)$, $1 \le i \le n$, for $\lambda > 0$ and $\mu_i > 0$. Then, if $x_i \in \{0,1\}$ for each $1 \le i \le n$, the following constraints are equivalent:*

$$\mathcal{C}_1(x) \ge p \Leftrightarrow \sum_{i=1}^{n} \mu_i x_i \le \mu^*,$$

where μ^* is the unique root of the equation $b - \mu = \Phi^{-1}(p)\sqrt{\lambda\mu}$.

Proof. Recall that if $a_1, \dots, a_n$ are independent Gaussian with mean μ_i and variance σ_i^2, and x_i are real numbers, then $a := \sum_{i=1}^n x_i a_i \sim \mathcal{N}(\mu(x), \sigma^2(x))$, with $\mu(x) = \sum_{i=1}^n x_i\mu_i$ and $\sigma^2(x) = \sum_{i=1}^n x_i^2\sigma_i^2$. Thus, because $x_i \in \{0,1\}$ and $\sigma_i^2 = \lambda\mu_i$ for each $1 \le i \le n$, we have $\sigma^2(x) = \lambda\mu(x)$. Then,

$$P\left(\sum_{i=1}^{n} a_i x_i \le b\right) = P\left(\mathcal{N}(0,1) \le \frac{b-\mu(x)}{\sqrt{\lambda\mu(x)}}\right),$$

so that $\mathcal{C}_1(x) \ge p$ is equivalent to

$$\frac{b-\mu(x)}{\sqrt{\lambda\mu(x)}} \ge \Phi^{-1}(p). \tag{7}$$

The left hand side of (7) is decreasing in $\mu(x)$, and thus $\mathcal{C}_1(x) \ge p$ is equivalent to $\mu(x) \le \mu^*$, where μ^* is the unique root of the equation $b - \mu = \Phi^{-1}(p)\sqrt{\lambda\mu}$. □

For gamma random variables, our result is the first to present a simple deterministic equivalent to a probabilistic constraint.

Proposition 4. *Consider n random variables $a_i \sim \Gamma(k_i, \theta)$, $1 \le i \le n$, for some $\theta > 0$ and $k_i > 0$, and assume that $b > 0$. Then, if $x_i \in \{0, 1\}$ for each $1 \le i \le n$, the following constraints are equivalent:*

$$\mathcal{C}_1(x) \ge p \Leftrightarrow \sum_{i=1}^{n} k_i x_i \le k^*,$$

where k^* is the unique solution of $\frac{\int_0^b z^{k-1} e^{\frac{-z}{\theta}} dz}{\Gamma(k)\theta^k} = p$ and the gamma function is defined by $\Gamma(k) = \frac{\int_0^\infty z^{k-1} e^{\frac{-z}{\theta}} dz}{\theta^k}$.

Proof. Gamma distributions satisfy also some kind of summation property, although weaker than the property satisfied by normal distributions. Recall that if $a_1, \dots, a_n$ are independent Gamma with shape k_i and a common scale θ, then $a := \sum_{i=1}^n a_i \sim \Gamma(k, \theta)$, with $k = \sum_{i=1}^n k_i$. Thus, if x_i are binary numbers, we have also that $a := \sum_{i=1}^n x_i a_i \sim \Gamma(k(x), \theta)$, with $k(x) = \sum_{i=1}^n k_i x_i$. Thus, for binary x_i, $\mathcal{C}_1(x)$ is equivalent to $P(\Gamma(k(x), \theta) \le b)$ defined by

$$\mathcal{K}(k(x)) := \frac{\int_0^b z^{k(x)-1} e^{\frac{-z}{\theta}} dz}{\Gamma(k(x))\theta^{k(x)}}.$$

Then, assuming that $\mathcal{K}(k)$ is a strictly decreasing function of k, the constraint $\mathcal{K}(k(x)) \ge p$ is equivalent to the constraint $k(x) \le k^*$, with $k^* = \mathcal{K}^{-1}(p)$ which proves $\mathcal{C}_1(x) \ge p \Leftrightarrow \sum_{i=1}^n k_i x_i \le k^*$. Note that $\mathcal{K}^{-1}$ is well defined for any $p \in (0, 1)$ because $\mathcal{K}$ is continuous, strictly decreasing, $\lim_{k \to 0^+} \mathcal{K}(k) = 1$ and $\lim_{k \to +\infty} \mathcal{K}(k) = 0$.

We are left to prove that $\mathcal{K}(k)$ is a strictly decreasing function of $k > 0$:

$$\begin{aligned}
\frac{d\mathcal{K}}{dk}(k) &= \theta \frac{d}{dk} \frac{\int_0^b z^{k-1} e^{-z} dz}{\int_0^\infty v^{k-1} e^{-v} dv} \\
&= \frac{\theta}{\Gamma^2(k)} \left(\int_0^b \ln(z) z^{k-1} e^{-z} dz \int_0^\infty v^{k-1} e^{-v} dv - \int_0^b z^{k-1} e^{-z} dz \int_0^\infty \ln(v) v^{k-1} e^{-v} dv \right) \\
&= \frac{\theta}{\Gamma^2(k)} \int_0^b dz \int_0^\infty dv \left(z^{k-1} v^{k-1} e^{-z-v} (\ln(z) - \ln(v) \right) \\
&= \frac{\theta}{\Gamma^2(k)} \int_0^b dz \int_b^\infty dv \left(z^{k-1} v^{k-1} e^{-z-v} \ln \frac{z}{v} \right),
\end{aligned}$$

which is strictly negative because $\ln \frac{z}{v} < 0$ for $(z, v) \in [0, b] \times (b, \infty)$. ☐

When $b \le 0$, we see from the definition of $\hat{\mathcal{C}}(x)$ that the probabilistic constraint is equivalent to $\sum_{i=1}^n x_i \le 0$.

5 Application to the chance-constraint knapsack

Let us apply Proposition 3 and 4 to a classical problem with a unique probabilistic constraint, studied by Klopfenstein and Nace (2008) and Goyal and Ravi (2009) among others.

Corollary 1. *Let a_i be independent random variables, $1 \leq i \leq n$, and define the chance-constrained knapsack problem as follows:*

$$\begin{array}{ll} \max & \sum_{i=1}^{n} p_i x_i \\ \textit{s.t.} & P\left(\sum_{i=1}^{n} a_i x_i \leq b\right) \geq p \\ & x \in \{0,1\}^n. \end{array}$$

The following hold:

1. *If $a_i \sim \mathcal{N}(\mu_i, \lambda\mu_i)$, for some $\lambda > 0$ and a positive integer vector μ, the problem can be solved in $O\left(\sum_{i=1}^{n} \mu_i\right)$.*

2. *If $a_i \sim \Gamma(k_i, \theta)$, for some $\theta > 0$ and a positive integer vector k, the problem can be solved in $O\left(\sum_{i=1}^{n} k_i\right)$.*

Proof. Consider case *2.*, when $a_i \sim \Gamma(k_i, \theta)$. Using Proposition 4, we can replace the probabilistic constraint by a linear one with integer coefficients, making the problem a knapsack problem with complexity $O\left(\sum_{i=1}^{n} \mu_i\right)$. The only trouble may happen with the computation of the new capacity, k^*.

Because k is integer and x binary, $\sum_{i=1}^{n} k_i x_i \leq k^*$ is equivalent to $\sum_{i=1}^{n} k_i x_i \leq \lfloor k^* \rfloor$ so that we only need to compute $C := \lfloor k^* \rfloor$. Moreover if $k^* \geq \sum_{i=1}^{n} k_i$, we can again replace the capacity by $C := \sum_{i=1}^{n} k_i$.

Therefore, we describe next how to compute C in $O\left(\log \sum_{i=1}^{n} k_i\right)$. First, if $\mathcal{K}\left(\sum_{i=1}^{n} k_i\right) \geq p$, then $k^* \geq \sum_{i=1}^{n} k_i$ and we can set $C := \sum_{i=1}^{n} k_i$. Otherwise, compute $\lfloor k^* \rfloor$ by a dichotomic search based on the sign of $\mathcal{K}(k) - p$. Case *1.* is proved similarly. □

Note that similar results could be obtained from Propositions 1 and 2, although the complexity of computing $P\left(\sum a_i \leq b\right)$ depends on the specific distribution of a.

Acknowledgements

This research is supported by an "Actions de Recherche Concertées" (ARC) projet of the "Communautée française de Belgique". Michael Poss is a research fellow of the "Fonds pour la Formation à la Recherche dans l'Industrie et dans l'Agriculture" (FRIA).

References

Beraldi, P. and Ruszczynski, A.: 2002, A branch and bound method for stochastic integer problems under probabilistic constraints, *Optimization Methods and Software* **17**, 359–382.

Calafiore, G. and Ghaoui, L. E.: 2006, Distributionally robust chance-constrained linear programs with applications, *Journal of Optimization Theory and Applications* **130**(1), 1–22.

Dentcheva, D., Prékopa, A. and Ruszczynski, A.: 2002, Bounds for probabilistic integer programming problems, *Discrete Applied Mathematics* **124**(1-3), 55 – 65.

Fortz, B., Labbé, M., Louveaux, F. V. and Poss, M.: 2009, The knapsack problem with gaussian weights, *Technical Report 592*, GOM, Université Libre de Bruxelles. http://www.ulb.ac.be//di/gom/publications/technical/2009/Fortz-Labbe-Louveaux-Poss.html.

Goyal, V. and Ravi, R.: 2009, Chance constrained knapsack problem with random item sizes. Submitted to Operations Research Letters.

Grinstead, C. and Snell, J.: 1997, *Introduction to Probability, 2nd edn*, American Mathematical Society, Providence.

Grossmann, I. E.: 2002, Review of nonlinear mixed-integer and disjunctive programming techniques, *Optimization and Engineering* **3**, 227–252(26).

Hansen, P. and Meyer, C.: 2009, Improved compact linearizations for the unconstrained quadratic 0–1 minimization problem, *Discrete Applied Mathematics* **157**, 1267–1290.

Henrion, R.: 2004, Introduction to chance-constrained programming. Tutorial paper for the Stochastic Programming Community Home Page.

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

Kataoka, S.: 1963, A stochastic programming model, *Econometrica* **31**(1/2), 181–196.

Klopfenstein, O.: In press, Solving chance-constrained combinatorial problems to optimality, *Computational Optimization and Applications* .

Klopfenstein, O. and Nace, D.: 2008, A robust approach to the chance-constrained knapsack problem, *Operations Research Letters* **36**(5), 628–632.

Luedtke, J., Ahmed, S. and Nemhauser, G.: 2008, An integer programming approach for linear programs with probabilistic constraints, *Mathematical Programming* **122**(2), 247–272.

Nolan, J. P.: 2010, *Stable Distributions - Models for Heavy Tailed Data*, Birkhäuser, Boston. In progress, Chapter 1 online at academic2.american.edu/∼jpnolan.

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

Parikh, S.: 1968, Lecture notes on stochastic programming. University of California, Berkeley, CA.

Prékopa, A.: 1995, *Stochastic Programming*, Kluwer.

Shapiro, A., Dentcheva, D. and Ruszczynski, A.: 2009, *Lectures on Stochastic Programming: Modeling and Theory*, SIAM, Philadelphia.