Introduction to Chance Constrained Programming

René Henrion

1 Introduction

Chance Constrained Programming belongs to the major approaches for dealing with random parameters in optimization problems. Typical areas of application are engineering and finance, where uncertainties like product demand, meteorological or demographic conditions, currency exchange rates etc. enter the inequalities describing the proper working of a system under consideration. The main difficulty of such models is due to (optimal) decisions that have to be taken prior to the observation of random parameters. In this situation, one can hardly find any decision which would definitely exclude later constraint violation caused by unexpected random effects. Sometimes, such constraint violation can be balanced afterwards by some compensating decisions taken in a second stage. For instance, making recourse to pumped storage plants or buying energy on the liberalized market is an option for power generating companies that are faced with unforeseen peaks of electrical load. As long as the costs of compensating decisions are known, these may be considered as a penalization for constraint violation. This idea leads to the important class of ***twostage*** or ***multistage stochastic programs*** [2, 16, 28].

In many applications, however, compensations simply do not exist (e.g., for safety relevant restrictions like levels of a water reservoir) or cannot be modeled by costs in any reasonable way. In such circumstances, one would rather insist on decisions guaranteeing feasibility 'as much as possible'. This loose term refers once more to the fact that constraint violation can almost never be avoided because of unexpected extreme events. On the other hand, when knowing or approximating the distribution of the random parameter, it makes sense to call decisions feasible (in a stochastic meaning) whenever they are feasible with high probability, i.e., only a low percentage of realizations of the random parameter leads to constraint violation under this fixed decision. A generic way to express such a ***probabilistic*** or ***chance constraint*** as an inequality is

$$P(h(x,\xi) \geq 0) \geq p. \tag{1}$$

Here, x and ξ are decision and random vectors, respectively, "$h(x,\xi) \geq 0$" refers to a finite system of inequalities and P is a probability measure. The value $p \in [0,1]$ is called the probability level, and it is chosen by the decision maker in order to model the safety requirements. Sometimes, the probability level is strictly fixed from the very beginning (e.g., $p = 0.95, 0.99$ etc.). In other situations, the decision maker may only have a vague idea of a properly chosen level. Of course, he is aware of that higher values of p lead to fewer feasible decisions x in (1), hence to optimal solutions at higher costs. Fortunately, it turns out that usually p can be increased over quite a wide range without affecting too much the optimal value of some problem, until it closely approaches 1 and then a strong increase of costs becomes evident. In this way, models with chance constraints can also give a hint to a good compromise between costs and safety.

Among the numerous applications of chance constrained programming one may find areas like water resource management, energy production, chip manufacturing, chemical engineering, telecommunication, and finance. Rather than listing all possible work in this context, we refer to the overview contained in [23]. The present document is devoted to a brief introduction to models, algorithmic aspects and theory of chance constrained programming and to motivate the reader to further consult deeper treatments of this subject area such as [21, 28].

2 An Example

2.1 Deterministic Version

As an illustrating example, we consider the following *cash matching problem* taken from [8] and [26]: the pension fund of a company has to make certain payments for the next 15 years which shall be financed by buying three types of bonds. The assumed data are collected in Table 1. An initial capital $K = 250,000$ is given. The amount of cash which is available at the end of the considered time horizon is to be maximized under the condition of meeting the payment restrictions in each year. For $i = 1, \ldots, n := 3$ and $j = 1, \ldots, m := 15$, we put:

$$\begin{aligned}
\alpha_{ij} &:= \text{yield per bond of type } i \text{ in year } j; \\
\beta_j &:= \text{payment required in year } j; \\
\gamma_i &:= \text{cost per bond of type } i; \\
x_i &:= \text{number of bonds of type } i \text{ to be bought.}
\end{aligned}$$

The amount of cash available in the fund at the end of year j is

$$\underbrace{K - \sum_{i=1}^{n} \gamma_i x_i}_{\text{cash after buying bonds}} \quad + \quad \underbrace{\sum_{k=1}^{j} \sum_{i=1}^{n} \alpha_{ik} x_i}_{\text{cumulative yields of bonds}} \quad - \quad \underbrace{\sum_{k=1}^{j} \beta_k}_{\text{cumulative payments}} \quad .$$

Table 1: Data for the cash matching problem

Year	Payments	Yields per bond of type: 1	2	3
1	11,000	0	0	0
2	12,000	60	65	75
3	14,000	60	65	75
4	15,000	60	65	75
5	16,000	60	65	75
6	18,000	1060	65	75
7	20,000	0	65	75
8	21,000	0	65	75
9	22,000	0	65	75
10	24,000	0	65	75
11	25,000	0	65	75
12	30,000	0	1060	75
13	31,000	0	0	75
14	31,000	0	0	75
15	31,000	0	0	1075
		cost per bonds:		
		980	970	1050

Introducing the quantities $a_{ij} := \sum_{k=1}^{j} \alpha_{ik} - \gamma_i$ and $b_j := \sum_{k=1}^{j} \beta_k - K$, the optimization problem may be written as

$$\max \sum_{i=1}^{n} a_{im} x_i \quad \text{subject to} \quad \sum_{i=1}^{n} a_{ij} x_i \geq b_j \quad (j = 1, \ldots, m)\,, \tag{2}$$

where the system of inequality constraints models the condition of positive cash in year j and the objective function corresponds to the final cash (up to a constant).
Allowing for fractional numbers of bonds to be bought, the solution of the linear programming problem (2) with the given data is

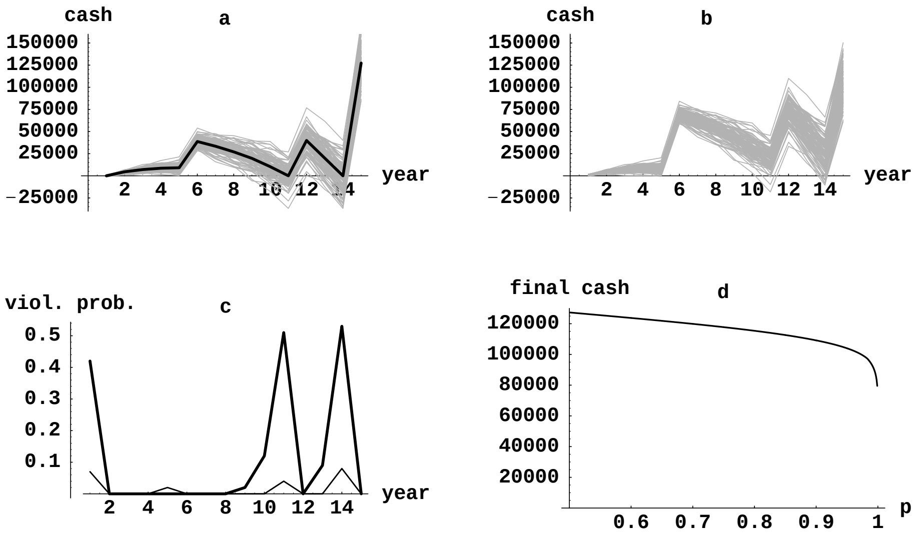
$$x_1 = 31.1,\ x_2 = 55.5,\ x_3 = 147.3, \text{ final amount of cash: } 127,332.$$

The amount of cash as a function of time is illustrated by the thick line in Figure 1 a.

2.2 Stochastic Version: individual chance constraints

As one can see from Figure 1 a (thick line), the cash profile of the fund reaches zero several times. Formally, this does not contradict feasibility of the solution,

Figure 1: Illustration of the cash matching problem (deterministic version versus individual chance constraints). For details see text.



but it strongly depends on the exactness of payment data from Table 1. It is not hard to believe that slight changes of these data would result in infeasible cash profiles under the solution found above. Indeed, there is good reason to assume random payment data in our problem due to demographic uncertainty in a future time period. Therefore, let us assume now, that the β_j introduced above are not the (deterministic) payments themselves but rather expected values of random payments $\tilde{\beta}_j$. To be more precise, we suppose that the $\tilde{\beta}_j$ are independently normally distributed with expectation β_j and standard deviation $\sigma_j := 500 \cdot j$ for $j = 1, \ldots, m$ (growing uncertainty over time). When simulating 100 payment profiles according to these assumptions and applying the deterministic solution from above, one arrives at 100 cash profiles illustrated in Figure 1 a (thin lines). As was to be expected, many of the cash profiles fall below zero in particular at the 'sharp times' when the deterministic profile reached zero. Evidently, the deterministic solution is no longer satisfactory as it violates the constraint of positive cash with high risk. A clearer view is obtained when plotting for each year the percentage of simulated cash profiles yielding constraint violation (Fig. 1 c (thick line)). Several times, this value approaches 50%. In order to arrive at a more robust solution, one could impose the restriction that with each year fixed, the probability of having positive cash exceeds say $p = 0.95$. Passing to the stochastic counterpart $\xi_j := \sum_{k=1}^{j} \tilde{\beta}_k - K$ of the deterministic quantity b_j introduced above, problem (2) then turns into an optimization problem with *individual chance constraints*:

$$\max \sum_{i=1}^{n} a_{im} x_i \quad \text{subject to } P\left(\sum_{i=1}^{n} a_{ij} x_i \geq \xi_j\right) \geq p \quad (j = 1, \ldots, m) . \tag{3}$$

The term 'individual' relates to the fact that each of the (stochastic) constraints $\sum_{i=1}^{n} a_{ij} x_i \geq \xi_j$ is transformed into a chance constraint individually. A different - and more realistic - version will be presented later. Before solving (3), we note that, due to the distribution assumptions on $\tilde{\beta}_j$, the random vector ξ has a multivariate normal distribution with mean vector b and with covariance matrix

$$\begin{pmatrix} \sigma_1^2 & \sigma_1^2 & \cdots & \sigma_1^2 \\ \sigma_1^2 & \sigma_1^2 + \sigma_2^2 & \cdots & \sigma_1^2 + \sigma_2^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_1^2 & \sigma_1^2 + \sigma_2^2 & \cdots & \sigma_1^2 + \cdots + \sigma_m^2 \end{pmatrix} . \tag{4}$$

In particular, ξ_j has variance $\tilde{\sigma}_j^2 = \sum_{k=1}^{j} \sigma_k^2$. Now, we pass to the standardized random variables

$$\eta_j := \tilde{\sigma}_j^{-1} (\xi_j - b_j) ,$$

which are normally distributed with zero mean and unit standard deviation. Then,

$$P\left(\sum_{i=1}^{n} a_{ij} x_i \geq \xi_j\right) = P\left(\tilde{\sigma}_j^{-1} \left(\sum_{i=1}^{n} a_{ij} x_i - b_j\right) \geq \eta_j\right) .$$

Denoting by q_p the p-quantile of the standard normal distribution (i.e., $q_{0.95} = 1.65$), we have the following equivalence:

$$P\left(\sum_{i=1}^{n} a_{ij} x_i \geq \xi_j\right) \geq p \Longleftrightarrow \sum_{i=1}^{n} a_{ij} x_i \geq \tilde{b}_j := b_j + \tilde{\sigma}_j q_p .$$

In other words, the chance constrained problem (3) becomes a simple linear programming problem again, but now with inequality constraints having higher right hand side values than in (2). The difference $\tilde{\sigma}_j q_p$ may be interpreted as a safety term. The solution of (3) with $p = 0.95$ is

$$x_1 = 62.8,\ x_2 = 72.6,\ x_3 = 101.1, \text{ final amount of cash: } 103,925.$$

This solution is evidently more in favour of short term bonds and it realizes a smaller final amount of cash. On the other hand, applying this solution to the same 100 payment profiles used above, it leads to cash profiles with quite rare constraint violations (see Fig. 1 b). The thin line in Fig. 1 c shows that the maximum percentage of negative cash is around 5% in good coincidence with the chosen probability level. The gain in reliability may be considered as substantial when compared to the loss of final cash. More insight is provided here by Fig. 1 d, where the optimal value of problem (3) is plotted as a function of the probability level p. The decrease of the optimal value is quite moderate over a wide range of p, and it is only beyond 0.95 that a further gain in safety is bought by drastic losses in final cash.

When dealing with stochastic constraints, one could be led to the simple idea of replacing the random parameter by its expectation. In our example, this approach would reduce to the deterministic problem (since the deterministic payments were taken as expectations for the stochastic payments). Therefore, the solution based on chance constraints illustrates at the same time the superiority with respect to the reliability/costs ratio over the expected value solution.

2.3 Stochastic Version: joint chance constraints

The model with individual chance constraints is appealing for its simple solvability. It has, however, the substantial drawback of not reflecting the proper safety requirements. It guarantees that, for each year fixed, the probability of negative cash is small. Yet, the probability of having negative cash at least once in the considered period may remain high. To illustrate this argument, Figure 2 plots again the 100 simulated cash profiles with those profiles being emphasized as dark lines which become negative at least one time. In case of the expected value solution (Figure 2 a) there are 82 profiles with occasional negative cash. This is an even more striking argument against the use of expected value solutions than the one of the preceding section. For the solution of individual chance constraints (Figure 2 b) it turns out that 16 profiles fall below zero at certain times. In other terms: the probability of maintaining positive cash over the whole period is around 0.84 and certainly significantly lower than the level of 0.95 chosen for the indivdual chance constraints. Obviously different subsets of profiles violate the constraint at different times. Of course, in practice, one has to make sure that - at high probability - cash remains positive over the whole time interval. In order to take care of what one could call 'uniform safety', one would have to replace the m individual chance constraints in (3) by one single *joint chance constraint* of the form

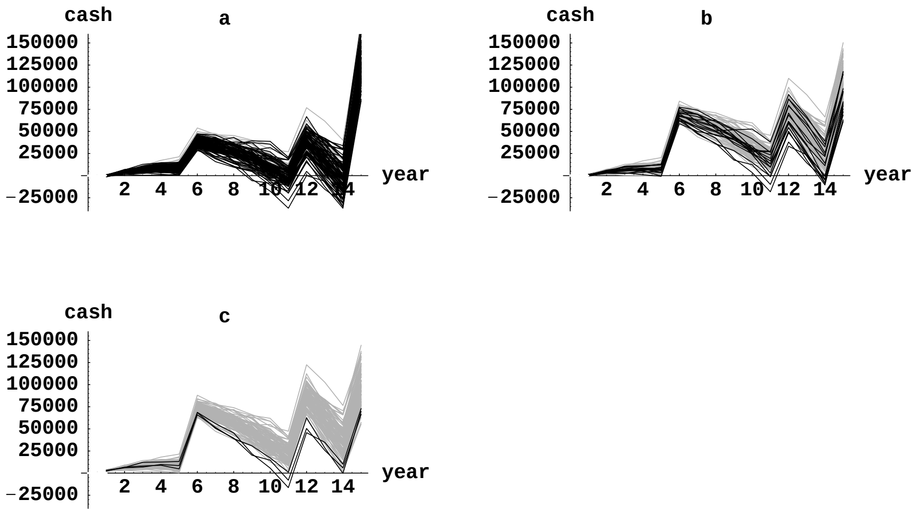
$$P\left(\sum_{i=1}^{n} a_{ij}x_i \geq \xi_j \quad (j = 1, \ldots, m)\right) \geq p. \tag{5}$$

Observe the difference between (5) and the constraint in (3) which is given by the position of the the counter $j = 1, \ldots, m$. At first glance, dealing with a single rather than many inequalities, seems to be a simpler task. However, the opposite is true as the calculation of (5) requires dealing with multidimensional distributions (see Section 3 below). The solution of the cash matching problem with joint chance constraints is

$$x_1 = 65.8,\ x_2 = 83.7,\ x_3 = 86.2, \text{ final amount of cash: } 98{,}160.$$

Comparing this with the previous solutions, once again there is some shift from long term to short term bonds, and the final amount of cash has slightly decreased again. On the other hand, robustness is significantly improved: Just 3

Figure 2: Illustration of the cash matching problem (deterministic version versus individual and joint chance constraints). For details see text.



out of 100 cash profiles fall below zero, i.e. 97% of the profiles stay positive over the whole period (see Figure 2 c). This is in good coincidence with the chosen probability level $p = 0.95$ and illustrates the difference with individual chance constraints.

Although the cash matching problem is a prototype example for models with joint chance constraints, there are certain situations where the use of individual chance constraints is appropriate. This applies, in particular, when proper working of some system (power generation, telecommunication) is not affected as long as only a sufficiently small number of components of that system fails. Then it does not make much sense to insist on proper working of a large majority of components over the whole period considered (joint chance constraints). Rather, at each time, a possibly varying large subset of components should be out of failure.

Finally, we emphasize that our presentation of the cash matching problem is just a first approximation to a real life model. Better models would include the possibility of short term borrowing cash in case of constraint violation. This leads to a mixture of probabilistic constraints and multistage programs which is very challenging but goes beyond the purpose of illustration here. Also, to justify the simplified setting, one could imagine that, for some reason, borrowing is excluded or future conditions of borrowing are extremely uncertain. It has to be mentioned that our model has a similar mathematical structure as many problems in engineering dealing with so called storage level constraints [14, 21, 24].

3 Algorithmic Challenges and Theory

Not surprisingly, there does not exist a unique solver for chance constrained programs. The choice strongly depends on how random and decision variables interact in the constraint model. Sometimes a linear programming solver will do the job (e.g., cash matching problem with individual chance constraints, see Section 2.2). In other models, one has to have access to values and gradients of multidimensional distribution functions (e.g., cash matching problem with joint chance constraints, see Section 2.3). Of particular interest is the application of algorithms from convex optimization. Convexity, of chance constraints, however, does not only depend on convexity properties of the constraint function h in (1) but also of the distribution of the random parameter ξ. The question of whether this distribution is continuous or discrete is another crucial aspect for algorithmic treatment. The biggest challenges from the algorithmic and theoretical points of view arise in chance constraints where random and decision variables cannot be decoupled.

In contrast to conventional optimization problems, those including chance constraints introduce the presence of two kinds of approximations: first, the distribution of the random parameter is almost never known exactly and has to be estimated from historical data. Secondly, even a given multivariate distribution (such as multivariate normal) can not be calculated exactly in general but has to be approximated by simulations or bounding arguments. Both types of imprecision motivate the discussion of stability in programs with chance constraints. All issues discussed up to now illustrate the close tie between algorithmic, structural and stability aspects. Some of these shall be briefly presented in the following sections. For the numerical solution of problems including chance constraints with random right hand side, we refer to the SLP-IOR model management system (see [18] and link section). If more general models are considered, specially designed algorithms are required.

3.1 Structural Aspects

Starting from a system of stochastic inequalities $h_j(x,\xi) \geq 0$, we have introduced before individual chance constraints as

$$\alpha_j(x) \geq p \quad (j=1,\ldots,m), \quad \text{where} \quad \alpha_j(x) := P(h_j(x,\xi) \geq 0) \quad (j=1,\ldots,m)$$

and joint chance constraints as

$$\alpha(x) \geq p, \quad \text{where} \quad \alpha(x) := P(h_j(x,\xi) \geq 0 \quad (j=1,\ldots,m)).$$

Formally, in both cases, one arrives at constraints on the decision variables as in usual optimization problems. However, α or the α_j are not given by an explicit formula but rather defined as probabilities of some implicitly defined regions in the space of the random parameter ξ. Calculating values, gradients and possibly Hessians of these functions remains the main challenge in chance constraint programing. The degree of difficulty in doing so, will mainly depend on the following items:

- Model of chance constraints (individual or joint)
- Assumptions on the random vector (e.g., continuous or discrete distribution, independent components)
- Type of stochastic inequalities (e.g., linear, convex, random right hand side)

The most favorable situation arises when $h_j(x, \xi) = g_j(x) - \xi_j \quad (j = 1, \ldots, m)$ for some functions g_j. In that case, the stochastic inequalities take the form $g_j(x) \geq \xi_j$ so that the random parameter appears on the right hand side. Using the distribution function $F_\eta(z) := P(\eta \leq z)$ for some random variable η, we may then write the individual chance constraints as

$$\alpha_j(x) = P(g_j(x) \geq \xi_j) = F_{\xi_j}(g_j(x)) \quad (j = 1, \ldots, m).$$

We recall that, for a one-dimensional random variable η, the following relation holds true with respect to the $p-$ quantile q_p of F_η:

$$F_\eta(t) \geq p \Longleftrightarrow t \geq q_p := \inf\{\tau | F_\eta(\tau) \geq p\}.$$

Consequently, the individual chance constraints can be rewritten as

$$\alpha_j(x) \geq p \Longleftrightarrow g_j(x) \geq q_p^{(j)} \quad (j = 1, \ldots, m),$$

where $q_p^{(j)}$ is the $p-$ quantile of F_{ξ_j}. In other words: individual chance constraints with random right hand side inherit their structure from the underlying stochastic constraint. If the latter was linear (as in the cash matching problem) then the induced individual chance constraints will be linear too. Unfortunately, this observation does not generalize to joint chance constraints. For random right hand side, the constraint function may still be written as a composition

$$\alpha(x) = P(g_j(x) \geq \xi_j \quad (j = 1, \ldots, m)) = F_\xi(g(x)), \tag{6}$$

where $g = (g_1, \ldots, g_m)$ and F_ξ is the distribution function of the random vector ξ. However, now the distribution function is multidimensional and no quantile argumentation applies any longer. Evidently, in order to calculate values and derivatives of α one has to be able to do so for mulitdimensional distribution functions (as g is given by an analytic formula in general). From the structural point of view, the composition formula $\alpha = F_\xi \circ g$ allows to transfer analytic properties like continuity, (local or global) Lipschitz continuity or smoothness from F_ξ and g to α. As far as convexity is concerned, we refer to Section 3.3.

If the random vector ξ has independent components, then the calculation of α breaks down to one dimensional distribution values again:

$$\alpha(x) = F_{\xi_1}(g_1(x)) \cdots F_{\xi_m}(g_m(x)).$$

Although the constraint $\alpha(x) \geq p$ cannot be further simplified to an explicit constraint invovling just the g_j (as was the case for individual chance constraints), one may benefit from the fact that one dimensional distributions are usually easy to calculate. On the other hand, the assumption of independent components is not justified in many applications. For instance, in the cash matching problem, the original random data (payments) were assumed to be independent, yet the cumulative data - which appeared in the chance constraint models - can definitely not be independent. This becomes evident from the nondiagonal covariance matrix in (4).

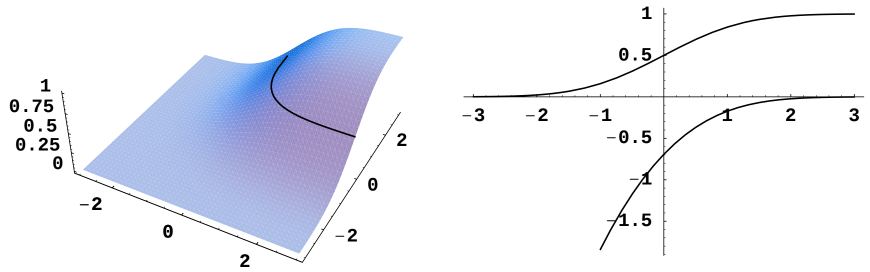
3.2 Random right hand side with nondegenerate multivariate normal distribution

Perhaps the most important special case in practical applications arises from joint chance constraints with random right hand side having a nondegenerate multivariate normal distribution. This assumption was made for instance, in the cash matching problem of Section 2.3. According to (6) the constraint takes the form $F_\xi(g(x)) \geq p$ then. As g is explicitly given by some formula, in general, the evaluation of such constraints by optimization algorithms requires the calculation of $F_\xi, \nabla F_\xi \ldots$, i.e., of values and (higher) derivatives of some nondegenerate multivariate normal distribution function. Fortunately, gradients of such distribution functions can be reduced analytically to some lower dimensional multivariate normal distribution functions (see [21], p. 204). Thus, proceeding by induction for higher order derivatives, the whole optimization issue hinges upon the evaluation of nondegenerate normal distribution functions in this situation. Much progress has been made here in moderate dimension. Two basic approaches can be distinguished: the first one is a combination of simulation and bounding techniques. Efficient bounds for intersection probabilities (special case: distribution functions) are based on improvements of classical Bonferroni bounds and on sophisticated graph theoretical derivations (e.g., [3, 4, 19]). Starting from simulated samples, these bounds may be used for constructing a variety of unbiased estimators for the probability to be approximated. An optimal convex combination of these estimators produces a final estimator having much less variance and, hence, yielding much more precise approximations than the original ones (e.g., [31]). A simulation method specially designed for multivariate normal distributions has been proposed in [6]. The second approach, which seems to be more precise in low dimensions, relies on numerical integration (e.g., [11, 29]). A Fortran code is available at [10]. A hybrid method combining the two approaches is discussed in [9].

3.3 Convexity

Convexity is a basic issue for theory (structure, stability) and algorithms (convergence towards global solutions) in any optimization problem. In chance constrained programming, the first question one could deal with is convexity of the

Figure 3: Bivariate normal distribution function (left) and standard normal distribution and its Logarithm (right).



feasible set defined say by a very simple probabilistic constraint of the type

$$\{x|P(\xi \leq x) \geq p\} = \{x|F_\xi(x) \geq p\}. \tag{7}$$

It is well known that such a set is convex if F_ξ is a quasiconcave function. Although distribution functions can never be concave or convex (due to being bounded by zero and one) it turns out that many of them are quasiconcave. The left plot of Figure 3 shows the graph of the bivariate normal distribution function with independent components. It is neither concave nor convex, but all of its upper level sets are convex (the boundary of the upper level set corresponding to the level $p = 0.5$ is depicted by a curve on the graph). For algorithmic purposes it is often desirable to know that the function defining an inequality constraint of type '$\geq$' is not just quasiconcave but actually concave. As mentioned above, this cannot hold for inequalities of type (7). However, a suitable transformation might do the job. Indeed, it turns out that most of the prominent multivariate distribution functions (e.g., multivariate normal, uniform distribution on convex compact sets, Dirichlet, Pareto, etc.) share the property of being log-concave, i.e., $\log F_\xi$ is concave (an illustration for the one-dimensional normal distribution and its log is given in the right plot of Figure 3). The key for verifying such a nontrivial property for the distribution function is to check the same property of log-concavity for the density of F_ξ, if it exists. The latter task is easy in general. For instance, a nondegenerate normal density is proportional to the exponential of a concave function, hence multivariate normal distributions are logarithmically concave. The mentioned result is a consequence of a famous theorem due to Prekopá [21]. Now, when F_ξ is log-concave, (7) may be equivalently rewritten as a concave inequality constraint $\{x|\log F_\xi(x) \geq \log p\}$.

Nothing changes of course for more general constraints of the type $F_\xi(Ax) \geq p$ where A is a matrix. This allows, for instance to solve the cash matching problem with joint chance constraints by means of convex optimization algorithms, for instance the supporting hyperplane method. Care has to be taken here, however, of the imprecise nature of F_ξ which is calculated by the methods mentioned in Section 3.2. For details of adapted algorithms, we refer to [18, 21, 30].

A challenging question with importance both for algorithms and stability (see Section 3.4) is to characterize strong concavity of the log of distribution functions. Examples of distributions sharing this property are the uniform distribution on rectangles [13] and the multivariate normal distribution with independent components [15]. It is interesting to observe that uniform distributions on arbitrary polytopes may lack strong log-concavity (e.g., on $\mathrm{conv}\,\{(0,1),(1,1)(1,0)\}$). The problem of normal distributions with correlated components is open.

3.4 Stability

In stochastic programming models, it is usually assumed that the probability distribution of the random parameter is known. In practice, however, this is rarely the case and one is faced with the need to approximate this distribution by parameter estimation or empirically. Solving the problem with approximating distributions yields solutions and optimal values that differ from those of the "true" problem. The main concern then is whether small approximation errors may lead to large deviations between solutions, or, expressed the other way around, whether it pays to spend large efforts in obtaining good approximations in order to arrive at solutions of high precision. The example of the 'value-at-risk' (see below) confirms that even the most simple chance constrained problems may fail to have stable solutions. In multistage stochastic programming this issue is central, for instance, in the context of scenario reduction/construction. [25] offers a general introduction to stability of stochastic programming problems. Recent results on stability in chance constrained programming can be found in [13, 15, 17]. To give a (strongly simplified) idea of these results, consider a program of type

$$\min\{f(x)|F_\xi(Ax) \geq p\}. \tag{8}$$

In particular the cash matching problem fits to this setting upon rewriting (5) in terms of the distribution function F_ξ. We denote by $v(\xi)$ the optimal value associated with that program, where ξ is considered as a varying random parameter close to some fixed $\bar{\xi}$ (the theoretical random vector). The following result holds:

Theorem 1 *In (8), let f be convex and let $\bar{\xi}$ have a log-concave distribution function $F_{\bar{\xi}}$. If there exists some $\bar{x}$ such that $F_{\bar{\xi}}(A\bar{x}) > p$, and if the solution set for (8) at $\xi = \bar{\xi}$ is nonempty and bounded, then one has a local Lipschitz estimate of the type*

$$\left|v(\xi) - v(\bar{\xi})\right| \leq L \sup_{z \in \mathbb{R}^s} \left|F_\xi(z) - F_{\bar{\xi}}(z)\right|. \tag{9}$$

At the same time, the solution set mapping is upper semicontinuous at $\bar{\xi}$.

To achieve quantitative results (Hölder- or Lipschitz stability) for solution sets, one has to impose further structural conditions (strong log-concavity, strict complementarity, $\mathcal{C}^{1,1}$-differentiability, see [15]).

Let us apply the theorem above to the well-known p-value-at-risk ($p \in [0,1]$) which is defined for a one dimensional random variable ξ as the quantity

$$VaR_p(\xi) = \inf\{x | P(x \geq \xi) \geq p\}.$$

This is clearly a special instance of (8) with $f(x) = x$, $A = 1$, so VaR_p is the optimal value of a special chance constrained program. In particular, if $\bar{\xi}$ has a log-concave distribution and $p \in (0,1)$, then the assumptions of the stability result in Section 3.4 are met, and we may derive from (9) the following Lipschitz continuity result for VaR_p:

$$\left|VaR_p(\xi) - VaR_p(\bar{\xi})\right| \leq L \sup_{z \in \mathbb{R}^s} \left|F_\xi(z) - F_{\bar{\xi}}(z)\right|. \tag{10}$$

Note that, in general, the value-at-risk does not depend continuously on ξ (e.g., for discrete distributions). Hence, (refvar) strongly depends on the assumption of $\bar{\xi}$ having a log-concave distribution.

3.5 Beyond the classical setting

The setting of joint chance constraints with random right hand side and non-degenerate multivariate normal distribution enjoys a lot of pleasant features such as differentiability or convexity (via log-concavity). Of course, other settings may have practical importance too. For instance, the distribution of the random right hand side could be other than normal. The case of multivariate Gamma or Dirchlet distributions is discussed in [21], Section 6.6. Here, log-concavity keeps being a valid tool.

Things become different, however, when passing to discrete distributions. These are of interest for at least two reasons: first, the problem to be solved could have been directly modeled by discrete random variables (see, e.g., [1]). Second, there may be a need to approximate continuous distributions (e.g., multivariate normal) by discrete ones, for instance when treating probabilistic constraints in two stage models with scenario formulation [27]. The key issue in discrete chance constrained programming is finding the so called p-efficient points (introduced in [20]) of the distribution function F_ξ of ξ. These are points z such that $F_\xi(z) \geq p$ and the relations $F_\xi(y) \geq p$, $y \leq z$ (partial order of vectors) imply that $y = z$. One easily observes that all the information about the p-level set of F_ξ is contained in these points because

$$\{y | F_\xi(y) \geq p\} = \bigcup_{z \in P} (z + \mathbb{R}^s_+),$$

where P is the set of p-efficient points and $\mathbb{R}^s_+$ is the positive orthant in the space of the random vector. In the case of ξ having integer-valued components and $p \in (0,1)$, P is a finite set (see Theorem 1 in [7]). Algorithms for enumerating or generating p-efficient points are described, for instance, in [1, 7, 22, 23]. It is interesting to note that the log-concavity concept, even if not directly applicable, can be adapted with useful consequences to discrete distributions as well [7].

When aiming at the more general case of probabilistic constraints, where the random parameter ξ does not necessarily appear on the right hand side of the inequalities, it is mainly the convex or polyhedral case under normal distribution which has some good chance of efficient treatment. More precisely, the function $h(x, \xi)$ is assumed to be quasiconcave in ξ so that the probability of the convex set $\{\xi || h(x, \xi) \geq 0\}$ is to be calculated then. This can be done, for each given x, by the already mentioned method proposed in [6] which is designed to calculate normal probabilities of convex sets. However, unlike the case of normal distribution functions, the derivative of $P(h(\cdot, \xi) \geq 0)$ no longer reduces analytically to values of the distribution function itself. A general derivative formula for probabilities of parameter dependent inequalities has been established in [32], but its practical use may be limited.

There is more hope in the specific polyhedral case $h(x, \xi) = A(x)\xi - b(x)$, where $A(x)$ and $b(x)$ are matrix and vector functions, respectively. For each x fixed, (1) amounts to the calculation of the probability of some polyhedron. Apart from treating polyhedra as special convex sets and applying [6] again, one could alternatively pass to the transformed random vector $\eta_x := -A(x)\xi$ so that (1) can be equivalently written in terms of the distribution function

$$F_{\eta_x}(-b(x)) \geq p.$$

Then, apparently, one seems to be back to the classical setting discussed in Section 3.2. For instance, if ξ had a multivariate normal distribution, then so does η_x (as a linear transform). However, in many important applications (like capacity expansion of networks with stochastic demands, see [21], section 14.3), the number of inequalities in the system $A(x)\xi \geq b(x)$ will typically exceed the dimension of ξ even after elimination of redundance. As a consequence, F_{η_x} will have a ***singular*** normal distribution which can no longer be calculated by the methods mentioned in Section 3.2. An algorithm for calculating singular normal distributions is proposed in [12]. A different approach, which is more efficient in moderate dimension, relies on directly calculating (regular) normal distributions of polyhedra. Based on a specific inclusion-exclusion formula for polyhedra proved in [5], the problem can be reduced to the calculation of a sum of regular normal distribution functions. At the same time, this offers the possibility of obtaining gradients too.

3.6 Concluding remarks

Chance constraints offer a way to model reliability in optimization problems. This brief introduction could neither provide a survey on the whole subject nor give a representative list of references. There are many topics of interest not covered by this paper. Research on theory and algorithms of chance constraints is quickly progressing with a focus on risk aversion (e.g., integrated chance constraints or stochastic dominance) which is caused by finance applications. This in turn introduces additional interesting structures like "semi-infinite chance constraints" (infinite number of continuously indexed inequalities).

3.7 Links

- Stochastic Programming E-Print Series (SPEPS)

  `www.speps.info/`

- Stochastic Programming Bibliography

  `mally.eco.rug.nl/biblio/SPlist.html`

- Stochastic Programming Information

  `users.iems.nwu.edu/~ jrbirge//html/dholmes/stopro.html`

- SLP-IOR

  `www.research-projects.unizh.ch/oek/unit36100/area38/p70.htm`

- FORTRAN codes for regular and singular multivariate normal distribution functions

  `www.sci.wsu.edu/math/faculty/genz/homepage`

- A tutorial on Chance Constrained Programming by R. Henrion

  `www-iam.mathematik.hu-berlin.de/~ romisch/SP01/Henrion.pdf`

- Handbook on Stochastic Programming

  `www1.elsevier.com/homepage/sae/orms/horms/vol10/contents.sht`

References

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

[2] J.R. Birge and F. Louveaux. *Introduction to Stochastic Programming.* Springer, New York, 1997.

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

[4] J. Bukszár and T. Szántai. Probability bounds given by hypercherry trees. *Optimization Methods and Software*, 17:409-422, 2002.

[5] J. Bukszár, R. Henrion, M. Hujter and T. Szántai. Polyhedral inclusion-exclusion. Preprint No. 913, Weierstrass Institute Berlin, 2004, submitted.

[6] I. Deák. Computing probabilities of rectangles in case of multidimensional distribution. *Journal of Statistical Computation and Simulation*, 26:101-114, 1986.

[7] D. Dentcheva, A. Prékopa and A. Ruszczyński. Concavity and efficient points for discrete distributions in stochastic programming, *Mathematical Programming*, 89:55-79, 2000.

[8] D. Dentcheva, B. Lai and A. Ruszczyński. Efficient point methods for probabilistic optimization problems. Stochastic Programming Eprint Series (SPEPS), No. 25, 2003.

[9] H.I. Gassmann, I. Deák and T. Szántai. Computing multivariate normal probabilities: A new look. *Journal of Computational and Graphical Statistics*, 11:920-949, 2002.

[10] A. Genz. `http://www.sci.wsu.edu/math/faculty/genz/homepage`

[11] A. Genz. Numerical computation of the multivariate normal probabilities. *Journal of Computational and Graphical Statistics*, 1:141-150, 1992.

[12] A. Genz and K. Kwong. Numerical evaluation of Singular Multivariate Normal Distributions, *Journal of Statistical Computation and Simulation*, 68:1-21, 2000.

[13] R. Henrion and W. Römisch. Metric regularity and quantitative stability in stochastic programs with probabilistic constraints. *Mathematical Programming*, 84:55-88, 1999.

[14] R. Henrion and A. Möller. Optimization of a continuous distillation process under random inflow rate. *Computers&Mathematics with Applications*, 45:247-262, 2003.

[15] R. Henrion and W. Römisch. Hölder and Lipschitz stability of solution sets in programs with probabilistic constraints. To appear in *Mathematical Programming*, 100:589-611, 2004.

[16] P. Kall and S. Wallace. *Stochastic Programming.* Wiley, New York, 1994. Online available at `http://stoprog.org`

[17] V. Kaňková. A note on multifunctions in stochastic programming. *Lecture Notes in Economics and Mathematical Systems*, 458:154-168, 1998.

[18] J. Mayer. *Stochastic Linear Programming Algorithms.* Gordon and Breach, Amsterdam, 1998.

[19] A. Prékopa. Boole-Bonferroni inequalities and linear programing. *Operations Research*, 36:145-162, 1988.

[20] A. Prékopa. Sharp bound on probabilities using linear programming. *Operations Research*, 38:227-239, 1990.

[21] A. Prékopa. *Stochastic Programming.* Kluwer, Dordrecht, 1995.

[22] A. Prékopa, B. Vizvári and T. Badics. Programming under probabilistic constraint with discrete random variable. In (F. Giannessi et al. eds.): *New Trends in Mathematical Programming*, pp. 235-255, Kluwer, Dordrecht, 1998.

[23] A. Prékopa. Probabilistic programming. In: [28] (Chapter 5).

[24] A. Prékopa and T. Szántai. Flood control reservoir system design using stochastic programming. *Mathematical Programming Study*, 9:138-151, 1978.

[25] W. Römisch. Stability of Stochastic Programming Problems. In: [28] (Chapter 8).

[26] A. Ruszczyński. `http://www.rusz.rutgers.edu`

[27] A. Ruszczyński. Probabilistic programming with discrete distributions and precedence constrained knapsack polyhedra. *Mathematical Programming*, 93:195-215, 2002.

[28] A. Ruszczyński and A. Shapiro. *Stochastic Programming.* Handbooks in Operations Research and Management Science, Vol. 10. Elsevier, Amsterdam, 2003.

[29] M. Schervish. Multivariate normal probabilities with error bound. *Applied Statistics*, 33:81-87, 1984.

[30] T. Szántai. A computer code for solution of probabilistic-constrained stochastic programming problems. In (Y. Ermoliev and R.J.-B. Wets eds.): *Numerical Techniques for Stochastic Optimization*, pp. 229-235, Springer, 1988.

[31] T. Szántai and A. Habib. On the k-out-of-r-from-n:F System with Unequal Element Probabilities. In (F. Giannessi et al. eds.): *New Trends in Mathematical Programming*, pp. 289-303, Kluwer, Dordrecht, 1998.

[32] S. Uryasev. Derivatives of probability functions and integrals over sets given by inequalities. *Journal of Computational and Applied Mathematics* 56:197–223, 1994.