2
Chance constrained programming

In this Chapter we give a brief introduction to chance constrained programming. The goals are to motivate the subject and to give the reader an idea of the related difficulties. All proofs are omitted: we indicate references where rigorous deductions can be found.

2.1
An example

At the risk of being repetitive, we start giving an example from [HV] of a simple chance constraint that illustrates one of the difficulties associated with this formulation. The simplicity of the example makes it essentially unique. For $x_1, x_2 \in \mathbb{R}$, $\varepsilon \in [0, 1]$, let

$$p(x) = \mathsf{Prob}\{\xi x_1 + x_2 \geq 7\} \geq 1 - \varepsilon$$

be a chance constraint, where ξ is uniformly distributed in $[0, 1]$, with cumulative distribution

$$F(t) = \begin{cases} 0, & \text{if } t \in (-\infty, 0), \\ t, & \text{if } t \in [0, 1], \\ 1 & \text{otherwise.} \end{cases}$$

In the general framework defined in (1-1), we have

$$G(x, \xi) = -\xi x_1 - x_2 + 7. \tag{2-1}$$

We are interested in an explicit representation of the set

$$C(\varepsilon) = \{(x_1, x_2) \in \mathbb{R}^2 : p(x) \geq 1 - \varepsilon\}. \tag{2-2}$$

If $x_1 = 0$, we clearly need to have $x_2 \geq 7$. If $x_1 > 0$,

$$p(x) = \mathbb{P}\left(\omega x_1 + x_2 \geq 7\right) = \mathbb{P}\left(\omega \geq \frac{7 - x_2}{x_1}\right) = 1 - F\left(\frac{7 - x_2}{x_1}\right). \tag{2-3}$$

Thus,

$$p(x) \geq 1 - \varepsilon \Leftrightarrow F^{-1}(\varepsilon) x_1 + x_2 \geq 7.$$

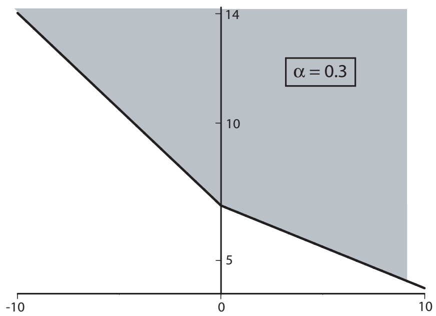

Figure 2.1: The set $C(0.3)$.

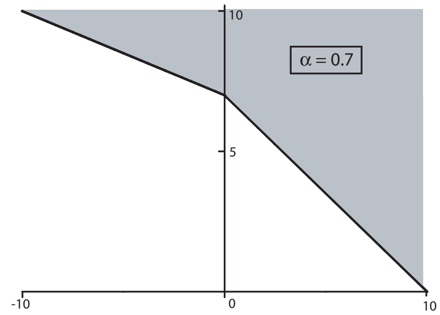

Figure 2.2: The set $C(0.7)$.

Proceeding in an analogous way for the case $x_1 < 0$, we end up with $C(\varepsilon) = C_+(\varepsilon) \bigcup C_0(\varepsilon) \bigcup C_-(\varepsilon)$, where

$$\begin{aligned} C_+(\varepsilon) &= \left\{x \in \mathbb{R}^2 \mid x_1 > 0,\ F^{-1}(\varepsilon)x_1 + x_2 \geq 7\right\}, \\ C_0(\varepsilon) &= \left\{(0, x_2) \in \mathbb{R}^2 \mid x_2 \geq 7\right\}, \\ C_-(\varepsilon) &= \left\{x \in \mathbb{R}^2 \mid x_1 < 0,\ x_1 F^{-1}(1-\varepsilon) + x_2 \geq 7\right\}. \end{aligned}$$

Figures 2.1 and 2.2 show the sets $C(0.3)$ and $C(0.7)$.

Clearly, from Figure 2.2, one cannot expect to have convex feasible sets for chance constrained programs even for linear functions G. Convexity is restored by requiring additional hypothesis, as shown below.

2.2
Single and joint constraints

There are essentially two ways of writing a chance constrained model. We can have several separated constraints, each one representing one goal. Formulation (1-1) represent the situation of a single separated chance constraint, which is amenable to the SAA approach we discuss later. A general separated chance constrained problem can be written as follows.

$$\begin{array}{ll} \min\limits_{x \in X} & f(x) \\ \text{s.t.} & p_i(x) := \mathsf{Prob}\left\{G_i(x,\xi) \leq 0\right\} \geq 1 - \varepsilon_i, \quad i = 1, \dots, m, \end{array} \tag{2-4}$$

where $\varepsilon_i \in [0,1]$. A point x is feasible to problem (2-4) if it belongs to the set

$$C(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_m) = \bigcap_{i=1}^{m} C_i(\varepsilon_i),$$

where

$$C_i(\varepsilon_i) = \left\{x \in \mathbb{R}^n \,|\, p_i(x) \geq 1 - \varepsilon_i\right\}.$$

Another possibility is having a number of constraints modeled as a single one as follows.

$$\begin{array}{ll} \underset{x \in X}{\text{Min}} & f(x) \\ \text{s.t.} & p(x) := \mathsf{Prob}\big\{G_1(x,\xi) \leq 0, G_2(x,\xi) \leq 0, \ldots, G_m(x,\xi) \leq 0\big\} \geq 1 - \varepsilon, \end{array} \tag{2-5}$$

with $\varepsilon \in [0,1]$. Formulation (2-5) is refereed to as a *joint chance constrained problem* since all the constraints $G_i(x,\xi) \leq 0$ are taken jointly. A point x is feasible to problem (2-5) if it belongs to the set

$$C(\varepsilon) = \big\{x \in \mathbb{R}^n \,|\, p(x) \geq 1 - \varepsilon\big\}.$$

From a modeling point of view, sometimes it makes sense to model all the constraints jointly if they together describe one goal. In [HEN], the author presents a cash matching problem using both separated and joint chance constraints. He compares the robustness of both formulations in the financial context and performs experiments showing the difference between the two approaches. We will have the opportunity to compare both formulations when we discuss the *joint hurdle-race problem* in Chapter 5.

Joint chance constrained problems are usually hard to solve because the joint expression in (2-5) requires a multidimensional integration to be computed. Even checking feasibility for a given candidate solution cannot be done easily and Monte-Carlo is required. There are some algorithms available for those problems such as Szántai's method ([HV]) or the solvers PCSPIOR, PROCON and PROBALL [KM], but they are restricted to multivariate normal distributions. Furthermore, they only deal with linear chance constraints with constant technology matrix [HV]. Other examples of algorithms are the SUMT and the supporting hyperplane method, described in detail in Chapter 5 of [PREa].

There is an interesting result linking the two formulations.

Theorem 1 *Let (2-5) be a joint chance constrained problem with reliability level ε. If we choose reliability levels $\varepsilon_i = 1 - (1-\varepsilon)/m,\ i = 1, \ldots, m$ for the separated problem (2-4), then*

$$\bigcap_{i=1}^{m} C_i\left(1 - \frac{1-\varepsilon}{m}\right) \subset C(\varepsilon),$$

that is, any feasible solution to the separated problem is feasible for the joint problem for a suitable choice of reliability levels ε_i.

Proof. The result follows directly from Bonferroni inequality [HV]. ∎

We can convert any joint chance constrained problem such as (2-5) into the form (1-1) by using the max-function as follows.

$$\begin{array}{ll} \operatorname*{Min}_{x\in X} & f(x) \\ \text{s.t.} & \mathsf{Prob}\big\{ \max_{i=1\dots,n} \{G_i(x,\xi)\} \le 0\big\} \ge 1-\varepsilon. \end{array} \tag{2-6}$$

It is straightforward to check that problems (2-6) and (2-5) are equivalent. Of course in some cases desired properties of the considered functions may be destroyed, but convexity is preserved and if the functions $G_i(\cdot,\xi)$ are linear we still can write (2-6) as a linear program. We will see an explicit example of such operation when we discuss the blending problem.

2.3
Some properties and special cases

The following result gives basic properties of feasible sets of chance constrained problems.

Theorem 2 *a) Let $p(x)$ be defined as in (2-5). We have that $p(x)$ is upper semicontinuous, that is*

$$p(x) \ge \limsup_{y\to x} p(y), \quad x \in \mathbb{R}^n,$$

and thus the set $C(\varepsilon)$ is a closed set for all $\varepsilon \in [0,1]$.

b) The set $C(\varepsilon)$ is nondecreasing: if $0 \le \varepsilon_1 < \varepsilon_2 \le 1$, then $C(\varepsilon_1) \subset C(\varepsilon_2)$. In addition, $C(1) = \mathbb{R}^n$ and $C(\varepsilon) \ne \emptyset$ for all $\varepsilon \in [0,1]$ if and only if the set $C(0)$ is non empty.

Part b) is trivial. A proof of part a) can be found in [HV].

As shown in Section 2.1, the feasible set of a chance constrained problem in general is not convex. However, there are results establishing convexity under certain hypothesis on the function G and on the density function of the random vector ξ. The most important is due to Prékopa and Borell ([PREa]) and is stated without proof.

Theorem 3 *Assume the random vector ξ has a continuous probability distribution with density function f. The following statements hold:*

a) If $\log f$ is concave (with $\log 0 = -\infty$), or

b) If $f^{-1/m}$ is convex (with $0^{-1/m} = \infty$),

then the cumulative distribution function F is quasi-concave and hence $C(\varepsilon)$ is a convex set for all $\varepsilon \in [0,1]$.

Fortunately, there are several important distributions that satisfy the hypothesis of Theorem 3. We give two examples:

- Uniform distribution. Let D be a convex subset of $\mathbb{R}^n$ with finite measure $|D|$. The probability density function is given by

$$f(x) = \begin{cases} \frac{1}{|D|} & \text{if } x \in D, \\ 0 & \text{if } x \notin D. \end{cases}$$

- Normal distribution. The probability density function is defined by

$$f(x) = \frac{1}{\sqrt{|\Sigma|}(2\pi)^{\frac{n}{2}}} \exp^{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)}, \quad x \in \mathbb{R}^n,$$

  where μ is the expectation vector, Σ the covariance matrix of the distribution and $|\Sigma|$ is the determinant of Σ.

Other examples are the (multivariate) Beta and Gamma distributions. More examples can be found in [PREa].

In the case the random variable ξ is discretely distributed, we can formulate problem (1-1) as a mixed-integer linear program. Let us assume $\mathsf{Prob}\{\xi = \xi^k\} = p_k,\ k = 1, \dots, K$. Problem (1-1) becomes

$$\begin{array}{ll} \min\limits_{x \in X} & f(x) \\ \text{s.t.} & \sum_{k=1}^K p_k \mathbb{1}_{(-\infty,0)}(G(x,\xi^k)) \geq 1 - \varepsilon, \end{array} \tag{2-7}$$

or, equivalently,

$$\begin{array}{ll} \min\limits_{x \in X} & f(x) \\ \text{s.t.} & \sum_{k=1}^K p_k \mathbb{1}_{(0,\infty)}(G(x,\xi^k)) \leq \varepsilon, \end{array} \tag{2-8}$$

Consider the following mixed-integer program.

$$\begin{array}{ll} \min\limits_{x \in X} & f(x) \\ \text{s.t.} & G(x,\xi^k) - Mz_k \leq 0, \quad k = 1, \dots, K, \\ & \sum_{k=1}^K p^k z_k \leq \varepsilon, \\ & z \in \{0,1\}^K, \end{array} \tag{2-9}$$

where M is a sufficiently large constant. We claim that (2-8) and (2-9) are equivalent. Indeed, let $(x, z_1, \dots, z_k)$ be a solution of problem (2-9). The first constraint of (2-9) tells us that $z_k \geq \mathbb{1}_{(0,\infty)}(G(x,\xi^k))$. From the second constraint of (2-9) we have

$$\varepsilon \geq \sum_{k=1}^K p^k z_k \geq \sum_{k=1}^K p_k \mathbb{1}_{(0,\infty)}(G(x,\xi^k)),$$

which implies that x is feasible for (2-8), with same objective value. Conversely, let x be feasible for problem (2-8) and define $z_k := \mathbb{1}_{(0,\infty)}(G(x, \xi^k))$. We have that $(x, z_1, \ldots, z_k)$ is feasible for (2-9) with same objective value, and thus both problems are equivalent.

We conclude with a convexity result for discrete distributions. A proof can be found in [KW].

Proposition 4 *Consider problem (2-5) with discrete distribution, that is, let* $p^k = \mathsf{Prob}\{\xi = \xi^k\}, k = 1, \ldots, K$. *Then for*

$$\varepsilon < \min_{k=1\ldots,K} \{p_k\}$$

the feasible set $C(\varepsilon)$ *is convex.*