1
Introduction

The field of stochastic programming is mainly concerned with the development of models and algorithms for optimization problems with uncertainty. More often than not the constants of an optimization problem are only approximations of measured quantities that could hardly be known to high accuracy. For instance in [BN], the authors analyze the set of problems of the well-known `NETLIB` library and perform their sensitivity analysis. Using robust optimization techniques, they show that feasibility of the usual optimal solution of linear programs can indeed be heavily affected by small perturbations in problem data.

The first publications in the field of stochastic programming appeared in the 1950's [BEA], [DAN], [CCS]. The subject received moderate attention until the early nineties, when an explosion in the number of publications took place. Stochastic programming is a powerful tool to deal with uncertainty and, unlike other approaches such as robust optimization, models the coefficients as random variables with known joint distribution. From the point of view of applications, such assumption may be quite demanding, specially if there is not enough data to correctly approximate the distribution of parameters. However, in some cases one does not need to have a complete knowledge of the distribution of the parameters. It is enough, instead, to have an algorithm which generates samples from the random variables in the problem. Furthermore, we believe that in most situations even a subjective choice of the joint distribution serves the decision maker well.

There are two main approaches to stochastic programming: two-stage problems with recourse and chance constrained programming. In two-stage models, the decision maker chooses an action in the present without knowing the outcome of future events. After the uncertainty is revealed, he then takes the best possible *recourse* action to (possibly) correct the unwanted consequences of his first decision. Deviations from the goals are penalized by the objective function. Such framework has applications in several fields such as finance [AHM], electricity generation [LS],[PP], hospital budgeting [KQ], production planning [PBK], etc. We refer the reader to [SR] for a

detailed discussion of the theoretical properties and more examples of two-stage problems. Among the efficient algorithms which deal with two-stage problems, one of the most popular is the L-shaped method, which is essentially Benders decompositions applied to the so-called *extensive form* of a two-stage problem (see [HV], [BP]). Other important methods based on sampling are the Stochastic Decomposition ([HS]) and the sample average approximation for two-stage problems ([LSW]).

The second approach, which is the focus of this thesis, is *chance constrained programming*, sometimes referred to as probabilistic programming. The subject was introduced by Charnes, Cooper and Symmonds [CCS] and have been extensively studied since. For a theoretical background we may refer to Prékopa [PREb] where an extensive list of references can be found. Applications include, e.g., water management [DGKS], soil management [ZTSK] and optimization of chemical processes [HLMS],[HM].

In chance constrained programming, the decision maker is interested in satisfying his goal "most of the time", that is, he admits constraint violation for some realizations of the random events. While two-stage problems penalize deviations from goals, chance constrained programming considers only the possibility of infeasibility, regardless of the amount by which the constraints are violated. In other words, the former approach measures risk *quantitatively* while the latter does it *qualitatively*. We consider problems of the form

$$\begin{array}{ll} \underset{x\in X}{\text{Min}} & f(x) \\ \text{s.t.} & \mathsf{Prob}\big\{G(x,\xi)\leq 0\big\}\geq 1-\varepsilon, \end{array} \tag{1-1}$$

where $X\subset\mathbb{R}^n$, ξ is a random vector[1] with probability distribution P supported on a set $\Xi\subset\mathbb{R}^d$, constraints are expressed through $G:\mathbb{R}^n\times\Xi\to\mathbb{R}^m$, $f:\mathbb{R}^n\to\mathbb{R}$ is the objective function and $\varepsilon\in(0,1)$ is the *reliability level*.

Although chance constraints were introduced almost 50 years ago, little progress was made until recently. Even for simple functions $G(\cdot,\xi)$, e.g., linear, problem (1-1) may be extremely difficult to solve numerically. One of the reasons is that for a given $x\in X$ the quantity $\mathsf{Prob}\{G(x,\xi)\leq 0\}$ requires a multi-dimensional integration. Thus, it may happen that the only way to check feasibility of a point $x\in X$ is by Monte-Carlo simulation. Moreover, convexity of X and of $G(\cdot,\xi)$ does not imply the convexity of the feasible set of problem (1-1).

That led to two somewhat different directions of research. One consists of discretizing the probability distribution P and solving the related combinato-

[1]We use the same notation ξ to denote a random vector and its particular realization. Which of these two meanings will be used will be clear from the context.

rial problem (see, e.g., [DPR], [LAN]). Another approach is to employ convex approximations of chance constraints ([NS]). As active members in this lines of research, A. Shapiro and S. Ahmed have been working recently in the theory of sampling and simulation applied to chance constrained programming. The *sample average approximation* (SAA) studied in this thesis is a sampling method for joint chance constrained problems or problems with a single chance constraint. The approach is natural, and, as will be seen, it is a flexible tool which can alleviate several difficulties such as non-convexity and the intractability of the probabilistic constraint.

In the third year of my doctorate in Atlanta, Ahmed and Shapiro introduced me to some theoretical aspects of SAA, and we proceeded to clarify foundations in order to advance to interesting applications. Ahmed's previous supervision of J. Luedtke [LA] gave rise to convergence results of SAA on specific scenarios, which were then used on probabilistic versions of the set covering and transportation problems. In this text we continue this path with the following contributions: first, the theoretical results vindicate the numerical approximations; then we provide further empirical evidence on how to choose the parameters involved in SAA and how to use it to solve chance constrained problems. Part of this material may be found in [PAS].

Chapter 2 contains some basic results about chance constrained problems, with emphasis on hypotheses leading to the convexity of the feasible set. In Chapter 3 we provide theoretical background and present the main results about sample approximations of (1-1). We state and prove convergence results and describe how to construct bounds for the optimal value of chance constrained programs.

In Chapter 4, we apply SAA to two rather simple problems, which allow for verification of our methods. The first is a linear portfolio selection problem with 10 assets, in the spirit of Markowitz ([MAR]). We consider two very distinct situations: the distribution of the returns of the assets is either multivariate normal or lognormal. In the first case, the explicit solution is well known: we use it as a benchmark to our numerics. The second problem is a simple blending problem modeled as a *joint* chance constrained problem, for which, again, the explicit solution is known.

In Chapter 5 we turn our attention to a more realistic problem arising from actuarial sciences, the *hurdle-race problem*, proposed in [VDGK]. It consists of a decision maker who needs to determine the current capital (provision) required to meet future obligations. Furthermore, for each period separately he needs to keep his capital above given thresholds, the *hurdles*, with high probability. In [VDGK], the authors make use of *comonotonicity*

([DDGa], [DDGb]) to obtain candidate solutions to the problem.

I contacted S. Vanduffel, the corresponding author of [VDGK], and proposed a variant of the *hurdle-race problem*, the *joint hurdle-race problem*. Instead of separate hurdles, the decision maker has to pass the whole collection of hurdles with high probability. This phrasing is clearly more adequate from an actuarial point of view. Comonotonicity cannot be easily applied to the joint version, but SAA yields good candidate solutions to the problem. Both models are compared in the original hurdle-race format, and numerical evidence of the robustness of the joint formulation is provided in Section 5.2.1.

In addition, we extend the formulation to include stochastic hurdles, so that the hurdles itself are not known at (known) future times and depend on discounted values of futures obligations at the risk free rate. Although the model becomes more involved, SAA handles the extension by essentially the same computational cost.

Chapter 6 concludes the thesis with a summary of the results and future directions of research.

We use the following notation throughout the text. The integer part of number $a \in \mathbb{R}$ is denoted by $\lfloor a \rfloor$. By $\Phi(z)$ we denote the cumulative distribution function (cdf) of standard normal random variable and by z_ε the corresponding $\varepsilon-$quantile, i.e., $\Phi(z_\varepsilon) = 1-\varepsilon$, for $\varepsilon \in (0,1)$. The cdf $B(k;p,N)$ of the binomial distribution is

$$B(k;p,N) := \sum_{i=0}^{k} \binom{N}{i} p^i (1-p)^{N-i}, \quad k = 0, ..., N. \tag{1-2}$$

For sets $A, B \subset \mathbb{R}^n$ we denote by

$$\mathbb{D}(A,B) := \sup_{x \in A} \operatorname{dist}(x,B) \tag{1-3}$$

the *deviation* of set A from set B.