1 INTRODUCTION

1.1 Decision under uncertainty

In a vast range of practical applications, the input data necessary for the solution of mathematical programs cannot be precisely determined beforehand. In general, that may happen either because data is inherently random or due to inevitable errors in measurement. In 1955, Dantzig [20] and Beale [6] first recognized that even a relatively small deviation from the values used as input data could compromise the quality of the optimal solution to a problem. Since then, two main methodologies have been developed with the aim of incorporating – into the modeling and solution procedures – the uncertainties which are part of a diverse set of problems: robust optimization and stochastic programming.

1.2 Robust Optimization

The field of robust optimization was founded in 1973 by Soyster's seminal work [54] which proposed the solution to a problem similar to that in standard form ($\min_{x \in X} c^T x \mid Ax \leq b$) with the additional requirement that the optimal solution should be feasible for all elements of the set $\mathcal{A} = \{A_j, \forall j \in J\}$ of technology matrices.

Following the notation of Bertsimas and Sim (2004) [15], let J_i denote the set of coefficients in row i of matrix A which are subject to uncertainty and each element $a_{ij}, (j \in J_i)$ be modeled as a symmetric and bounded random variable with support $[a_{ij} - \hat{a}_{ij}, a_{ij} + \hat{a}_{ij}]$. The formulation proposed by Soyster may be written as:

$$Min \quad c^T x \tag{1.1}$$

$$\text{subject to:} \quad \sum_j a_{ij}x_j + \sum_{j\in J_i} \hat{a}_{ij}y_j \leq b_i \qquad \forall i \tag{1.2}$$

$$-y_j \leq x_j \leq y_j \qquad \forall j \tag{1.3}$$

$$l \leq x \leq u \tag{1.4}$$

$$y \geq 0 \tag{1.5}$$

where l and u are vectors of appropriate dimension which represent, respectively, lower and upper bounds on variables x_j.

Such an approach is shown by Soyster to be equivalent to a worst-case scenario analysis. This extreme conservativeness leads the value of the objective function at the optimal solution to be usually significantly worse than that of the original (or nominal-value) problem and motivated the search for different approaches which could provide a balance between feasibility and optimality.

A quarter of a century after Soyster's work, Ben-Tal and Nemirovksi ([9], [10], [11] and [12]) and El-Ghaoui et al. ([22] and [23]) proposed an alternative way to model the uncertainty by defining "ellipsoidal regions of uncertainty" around the nominal values of the coefficients, inside which one admits that the realization of the unknown parameters will be. The proposed approach results in a modification of the original constraints of the problem which turns it into a second order conic program, thus requiring specific solution procedures (which are, in general, not guaranteed to find the global optimum solution to a problem):

$$Min \quad c^T x \tag{1.6}$$

$$\text{subject to:} \quad \sum_j a_{ij}x_j + \sum_{j\in J_i} \hat{a}_{ij}y_{ij} + \Omega_i \sqrt{\sum_{j\in J_i} \hat{a}_{ij}^2 z_{ij}^2} \leq b_i \qquad \forall i \tag{1.7}$$

(1.8)

$$-y_{ij} \leq x_j - z_{ij} \leq y_{ij} \qquad \forall j \tag{1.9}$$

$$l \leq x \leq u \tag{1.10}$$

$$y \geq 0 \tag{1.11}$$

where Ω_i is a user-defined parameter related to the probability of violation of each constraint – the authors prove that the probability of each constraint i being violated is less or equal to $\exp(-\Omega_i^2/2)$.

Robust optimization was again boosted in 2003 with the publication of [13], [14] and [15] by Bertsimas and Sim. The novel approach assumes a polyhedral uncertainty set and its major advantage is the fact that the formulation of the robust counterpart of a problem does not modify its structure, maintaining all the original properties such as linearity. In summary, the proposed approach introduces a parameter Γ_i that takes values in the interval $[0, |J_i|]$ and determines the maximum number of coefficients in row i which will be allowed to vary from their respective nominal values a_{ij}. The robust counterpart is initially formulated as:

$$Min \quad c^T x \tag{1.12}$$

$$subject\ to: \quad \sum_j a_{ij} x_j + \beta_i(x, \Gamma_i) \le b_i \qquad \forall i \tag{1.13}$$

$$-y_j \le x_j \le y_j \qquad \forall j \tag{1.14}$$

$$l \le x \le u \tag{1.15}$$

$$y \ge 0 \tag{1.16}$$

where:

$$\beta(x, \Gamma_i) = Max \quad \sum_{j \in J_i} \hat{a}_{ij} |x_j| z_{ij} \tag{1.17}$$

$$subject\ to: \quad \sum_{j \in J_i} z_{ij} \le \Gamma_i \qquad \forall i \tag{1.18}$$

$$0 \le z_{ij} \le 1 \qquad \forall j \in J_i \tag{1.19}$$

As shown in [15], this is equivalent to the linear formulation presented below:

$$Min \quad c^T x \tag{1.20}$$

$$\textit{subject to:} \quad \sum_j a_{ij}x_j + z_i\Gamma_i + \sum_{j\in J_i} p_{ij} \le b_i \qquad \forall i \qquad (1.21)$$

$$z_i + p_{ij} \ge \hat{a}_{ij}y_j \qquad \forall i, j \in J_i \qquad (1.22)$$

$$-y_j \le x_j \le y_j \qquad \forall j \qquad (1.23)$$

$$l \le x \le u \qquad (1.24)$$

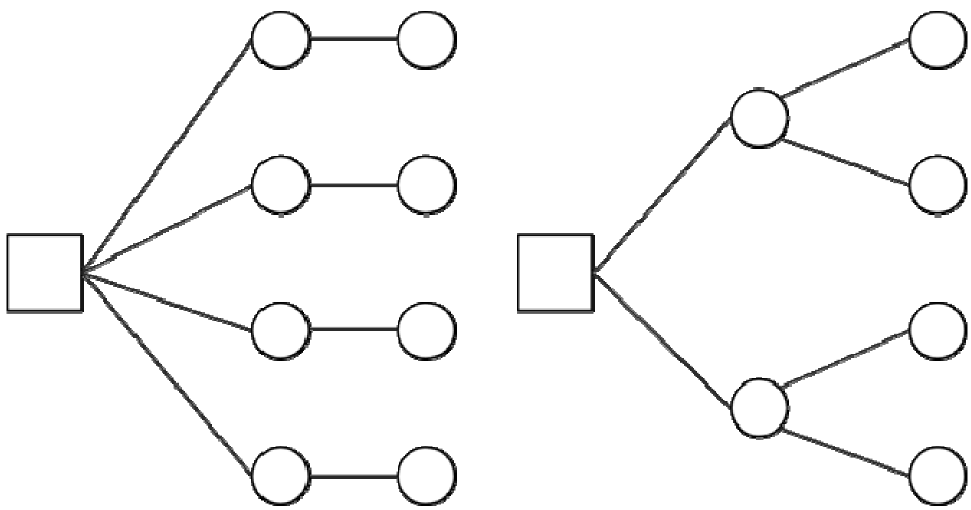
$$y, z, p \ge 0 \qquad (1.25)$$

1.3
Stochastic programming

The stochastic programming approach relies on the assumption – which is perfectly reasonable in various settings – that one might be able to know or estimate the probability distribution of the unknown parameters. Generally speaking, the objective of stochastic programming models is to determine a solution that is feasible for all possible data realizations (or for a given percentage of them) and that minimizes the expected value of a function of the decision and random variables.

The objective of this Section is not to provide a comprehensive overview on the subject – which the interested reader may find in Birge and Loveaux (1997) [16], Kall and Wallace (1994) [40], Ruszczynski and Shapiro (2003) [52], Shapiro, Dentcheva and Ruszczynski (2009) [55] and Haneveld and van der Vlerk (2005) [33] – but to introduce the topic so that the reader may grasp the basic difference between standard stochastic programming models in the literature and the one studied in this thesis. In addition to the basic references just mentioned, the state-of-the-art in various applications may be found in Wallace and Fleten (2003) [67] (energy), Dupacova, Hurt and Stepan (2002) [21] (finance), Poojari, Lucas and Mitra (2006) [49] (supply chain and logistics) and Gaivoronski (2005) [26] (telecommunications).

The majority of research and applications of stochastic programming is done on the so-called two-stage stochastic programming linear models, although multistage stochastic programs are also the subject of great interest – a graphical depiction of the conceptual difference between two-stage and multistage models is presented in Figure 1-1. In the former case, one usually seeks to determine a first stage decision which is then succeeded by the realization of a random event that affects the outcome of the action taken. Recourse actions may then be taken in the second stage so as to compensate for potential damages caused by the realization of the random variable(s). While in the second stage there might be a different set of corrective decisions for each scenario, according the possible outcomes of the random event, first stage decisions for all scenarios are required to be the same – a condition usually referred to as non-anticipativity.



(A) (B)

Figure 1-1 – Two-stage (A) and multistage (B) scenario-tree structure of stochastic programming models

The general formulation of a two-stage stochastic program is presented next:

$$Min \quad c^T x + \mathbb{E}\{Q(x,\xi)\} \tag{1.26}$$

$$subject\ to: \quad Ax \leq b \tag{1.27}$$

$$x \in X \tag{1.28}$$

where $Q(x,\xi)$ is defined as the value of the optimal solution of the second stage problem:

$$Min \quad q(\xi)^T y \tag{1.29}$$

$$\textit{subject to:} \quad T(\xi)x + W(\xi)y \leq h(\xi) \tag{1.30}$$

$$y \in Y \tag{1.31}$$

The actions to be taken before the random parameters are known are determined by the vector of first stage decision variables x, whose feasible region is defined by the set of constraints $Ax \leq b$ and by the set X – which may include integrality constraints. The vector of second stage decision variables is denoted by y and the the vector of coefficients of the objective function q, technology matrices T and W and the right-hand side vector h may all depend on the vector of random variables ξ.

Difficulties in evaluating multi-dimensional integrals imply that the determination of a numerical solution to these problems usually require the enumeration of a finite number S of possible outcomes for the vector $\xi = \{\xi_1, \xi_2, \dots, \xi_S\}$. Each one of these outcomes is called a scenario, to which there must also be an associated probability of occurrence $p = \{p_1, p_2, \dots, p_S\}$. This discretization allows the expression for the expected value in equation (1.22) to be written as:

$$\mathbb{E}\{Q(x,\xi)\} = \sum_{s \in S} p_s \cdot Q(x, \xi_s) \tag{1.32}$$

Finally, problems (1.26) – (1.28) and (1.29) – (1.31) may now be jointly re-written as follows:

$$Min \quad c^T x + \textstyle\sum_{s \in S} p_s q_s y_s \tag{1.33}$$

$$\textit{subject to:} \quad Ax \leq b \tag{1.34}$$

$$T_s x + W_s y_s \leq h_s \qquad \forall s \in S \tag{1.35}$$

$$x \in X, y \in Y \tag{1.36}$$

1.4
Motivation and related bibliography

A common hypothesis concerning the two approaches discussed above is that the realization of the uncertain parameters is independent of the decision variables, as illustrated in Figure 1-2. This conjecture is valid in a variety of applications, such as portfolio optimization, hydrothermal scheduling for electricity generation, communication network planning under demand uncertainty, etc. Not surprisingly, the vast majority of the body of work both in robust optimization and in stochastic programming deals with problems in which this hypothesis is satisfied and the uncertainty is said to be exogenous.

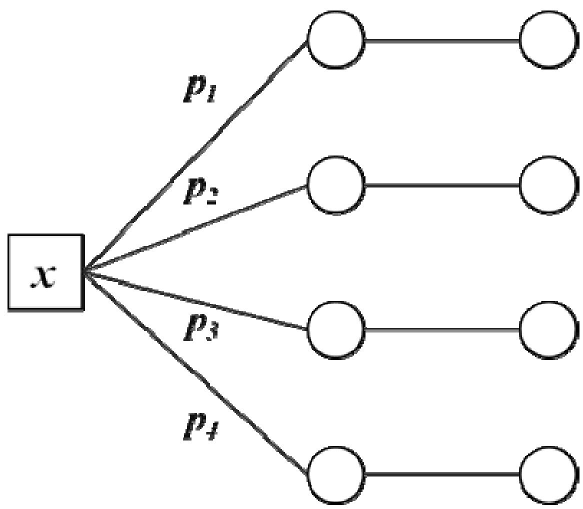


Figure 1-2 – Stochastic programming model with exogenous uncertainty – probabilities $\boldsymbol{p_1}$, $\boldsymbol{p_2}$, $\boldsymbol{p_3}$ and $\boldsymbol{p_4}$ are independent of decision $\boldsymbol{x}$

On the other hand, the literature on problems where the knowledge of the probability of occurrence of random events depends on the decisions taken (i.e., when the uncertainty is said to be endogenous) is very limited. According to Goel and Grossmann (2006) [31], out of the 4300+ works in the Stochastic Programming Bibliography compiled by van der Vlerk [65], only 8 ([48], [66], [2], [39], [36], [30], [31] and [61]) involve the case of endogenous uncertainty (references [54] and [47] are other works on the subject, not yet included in the database).

The work on stochastic programs with endogenous uncertainty may be further sub-divided into two categories with respect to the particular way in which decisions affect the knowledge of the probability distributions.

The first group involves problems where the probability distribution of the random variables is not directly affected but, rather, uncertainty may be partially resolved depending on actions performed by the decision-maker. This is essentially related to the timing of information discovery and to an anticipation or delay of the moment at which more accurate information is revealed. Such situation is pictured in Figure 1-3 below, in which the dashed line represents a possible relaxation of non-anticipativity constraints between scenarios related to first-stage decisions.

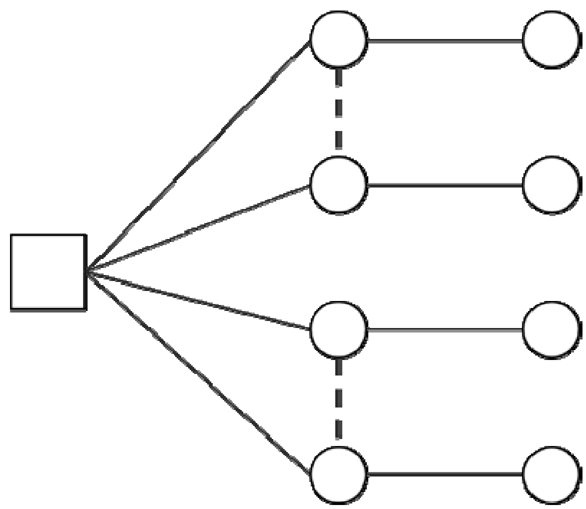


Figure 1-3 – Endogenous uncertainty related to the time of information discovery

This group includes the work of Jonsbraten (1998) [39], Goel and Grossmann (2004, 2006) [30][31], Held (2003) [36] and Senay (2007) [54]. The type of uncertainty dealt with in these works is exemplified by that studied in [39] and [30] where an oil and gas exploration company must choose among different testing and probing methods in order to try and find the size and quality of reserves – the installation of a facility does not change the likelihood of the company actually finding oil, but may provide evidence as to what are the most probable scenarios. Other examples lie in the areas of project management [54] and network interdiction.

Finally, the second group of stochastic programs with endogenous uncertainty refers to those in which decisions directly affect the probability distribution of the random parameters i.e., the actions performed at a given stage may change the probability of occurrence of future events – as conceptually illustrated in Figure 1-4.

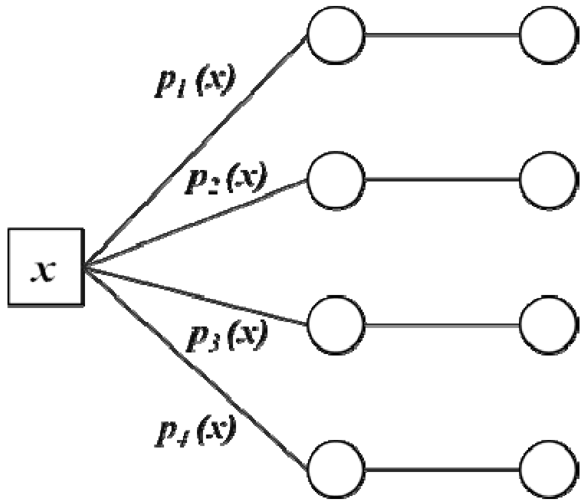


Figure 1-4 – Endogenous uncertainty and decision-dependent probabilities

Pflug (1990) [48] was the first to address this issue by discussing an application in stochastic queuing networks – decisions affect the arrival and service rates of each element in the queue – and proposing a stochastic quasigradient algorithm which requires repeated simulations of the system's functioning for each fixed first-stage solution. Talluri and Ryzin (2004) [61] worked on a revenue management problem from the point of view of an airline who must choose which combination of fares to offer at each moment in time preceding the departure of a flight. Under some assumptions regarding consumer behavior, they developed a dynamic programming algorithm to determine the pricing policy which results in the maximum total expected revenue. In 2000, Ahmed [2] presented some examples related to network design, server selection and facility location. These problems were formulated under a hyperbolic programming framework and a specialized algorithm was developed. An application to the stochastic PERT (Program Evaluation and Review Technique) problem is developed by Plambeck et al. in [47] where one seeks to minimize two conflicting objectives: a project's cost and its completion time. A sample-path algorithm is proposed and results are presented under the assumption of uniform distributions with a fixed spread around the mean. Viswanath et al. (2004) [66]

studied the humanitarian logistics problem – briefly described in Section 1.5 below and then again discussed in Chapter 2 in a more detailed fashion – and proposed an approximation to the objective function which allows the simplification of the problem down to an ordinary knapsack problem.

Given the diminished amount of research on the topic, it is expected that there should be many questions to be answered. In the next section a brief description of the specific problem to be tackled is given, along with a characterization of a more general class of problems for which the results obtained in the thesis are also valid.

1.5
Objective and contributions

This thesis will focus on the second group of stochastic programs with endogenous uncertainty discussed above and, in this sense, the humanitarian logistics problem (as defined in Viswanath et al. [66]) will be used as the main motivating example.

A detailed description of the problem is provided in Chapter 0 but, essentially, it refers to the problem of determining the optimal set of investments on the reinforcement of the links of a network which are subject to random failures – the decision to reinforce a link increases the probability that it will be available afterwards.

The results presented in the thesis, although discussed in the context of the humanitarian logistics problem, should also hold for a more general class of problems, including some of those discussed above – namely the ones related to stochastic queuing networks, stochastic PERT and revenue management. The general formulation of such problem class is given by:

$$Min \quad c^T x + \mathbb{E}_x\{Q(x, \xi(x))\} \tag{1.37}$$

$$subject\ to{:} \quad Ax \leq b \tag{1.38}$$

$$x \in X \tag{1.39}$$

where the function $Q(x, \xi(x))$ is now defined as the optimal solution of the following second stage problem:

$$Min \quad q(\xi(x))^T y \tag{1.40}$$

$$subject\ to: \quad W(\xi(x))y \leq h(\xi(x)) \tag{1.41}$$

$$y \in Y \tag{1.42}$$

It is important to observe that the coupling between the first and second stages is not given by the existence of the term Tx as in the set of constraints (1.35) of problem (1.33) – (1.36) but by the dependence of the probability distribution of the random variables with respect to first stage decision variables x – evidenced by the subscript x in the expression $\mathbb{E}_x\{Q(x, \xi(x))\}$.

The methodology proposed in the thesis will allow the determination of provably optimal solutions to instances of problems much larger then those currently solved in the literature. Specifically, the contributions of the thesis are:

1) **Reformulation scheme** which avoids the non-linearities due to products of first and second stage variables and due to the calculation of scenarios probabilities.

2) **Provably finite cut generation algorithm** that overcomes a potential pitfall of the proposed linearization technique and allows the solution of moderately-sized instances for a given error tolerance level;

3) **Incorporation of importance sampling concepts** into the stochastic programming framework. This overcomes the problem of not knowing the probability distribution of the random variables beforehand and allows the solution of large sample-based instances of the problem.

1.6 Outline

The remainder of this work is organized as follows: Chapter 2 describes the humanitarian logistics problem in detail, with a special emphasis on the difficulties that arise out of its formulation; Chapter 3 presents the re-formulation scheme which solves the obstacles related to existing non-linearities; Chapter 4 introduces the approximation algorithm based on cut generation and Chapter 5 extends this algorithm into a statistical framework in order to consider instances of the problem that are not amenable to complete scenario enumeration; Chapter 6 presents computational results, Chapter 7 concludes and discusses future work alternatives and how the developments presented in the previous chapters may be extended to other contexts.