A. Ruszczyński and A. Shapiro, Eds., *Handbooks in OR & MS, Vol. 10*

Chapter 1

Stochastic Programming Models

Andrzej Ruszczyński

Department of Management Science and Information Systems, Rutgers University, 94 Rockefeller Rd, Piscataway, NJ 08854, USA

Alexander Shapiro

School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA

Abstract

In this introductory chapter we discuss some basic approaches to modeling of stochastic optimization problems. We start with motivating examples and then proceed to formulation of linear, and later nonlinear, two stage stochastic programming problems. We give a functional description of two stage programs. After that we proceed to a discussion of multistage stochastic programming and its connections with dynamic programming. We end this chapter by introducing robust and min–max approaches to stochastic programming. Finally, in the appendix, we introduce and briefly discuss some relevant concepts from probability and optimization theories.

Key words: Two stage stochastic programming, expected value solution, stochastic programming with recourse, nonanticipativity constraints, multistage stochastic programming, dynamic programming, chance constraints, value at risk, scenario tree, robust stochastic programming, mean–risk models.

1 Introduction

1.1 Motivation

Uncertainty is the key ingredient in many decision problems. Financial planning, airline scheduling, unit commitment in power systems are just few examples of areas in which ignoring uncertainty may lead to inferior or simply wrong decisions. Often there is a variety of ways in which the uncertainty can be formalized and over the years various approaches to optimization under uncertainty were developed. We discuss a particular approach based on probabilistic models of uncertainty. By averaging possible outcomes or considering probabilities of events of interest we can define the objectives and the constraints of the corresponding mathematical programming model.

To formulate a problem in a consistent way, a number of fundamental assumptions need to be made about the nature of uncertainty, our knowledge of it, and the relations of decisions to the observations made. In order to motivate the main concepts let us start by discussing the following classical example.

Example 1 (Newsvendor Problem). A newsvendor has to decide about the quantity x of newspapers which he purchases from a distributor at the beginning of a day at the cost of c per unit. He can sell a newspaper at the price s per unit and unsold newspapers can be returned to the vendor at the price of r per unit. It is assumed that $0 \le r < c < s$. If the demand D, i.e., the quantity of newspapers which he is able to sell at a particular day, turns out to be greater than or equal to the order quantity x, then he makes the profit $sx - cx = (s-c)x$, while if D is less than x, his profit is $sD + r(x-D) - cx = (r-c)x + (s-r)D$. Thus the profit is a function of x and D and is given by

$$F(x, D) = \begin{cases} (s-c)x, & \text{if } x \le D, \\ (r-c)x + (s-r)D, & \text{if } x > D. \end{cases} \tag{1.1}$$

The objective of the newsvendor is to maximize his profit. We assume that the newsvendor is very intelligent (he has Ph.D. degree in mathematics from a prestigious university and sells newspapers now), so he knows what he is doing. The function $F(\cdot, D)$ is a continuous piecewise linear function with positive slope $s - c$ for $x < D$ and negative slope $r - c$ for $x > D$. Therefore, if the demand D is known, then the best decision is to choose the order quantity $x^* = D$. However, in reality D is not known at the time the order decision has to be made, and consequently the problem becomes more involved.

Since the newsvendor has this job for a while he collected data and has quite a good idea about the probability distribution of the demand D. That is, the demand D is viewed now as a *random variable* with a known, or at least well estimated, probability distribution measured by the corresponding cumulative distribution function (cdf) $G(w) := \mathbb{P}(D \le w)$. Note that since the demand cannot be negative, it follows that $G(w) = 0$ for any $w < 0$. By the Law of Large Numbers the average profit over a long period of time tends to the expected value

$$\mathbb{E}[F(x, D)] = \int_0^{\infty} F(x, w) \, \mathrm{d}G(w).$$