

Models and model value in stochastic programming

John R. Birge*

*Department of Industrial and Operations Engineering,
The University of Michigan, Ann Arbor, MI 48109, USA*

Finding optimal decisions often involves the consideration of certain random or unknown parameters. A standard approach is to replace the random parameters by the expectations and to solve a deterministic mathematical program. A second approach is to consider possible future scenarios and the decision that would be best under each of these scenarios. The question then becomes how to choose among these alternatives. Both approaches may produce solutions that are far from optimal in the stochastic programming model that explicitly includes the random parameters. In this paper, we illustrate this advantage of a stochastic program model through two examples that are representative of the range of problems considered in stochastic programming. The paper focuses on the relative value of the stochastic program solution over a deterministic problem solution.

Keywords: Stochastic optimization models, finance, quality control, engineering design.

1. Introduction

Practical decisions often involve the consideration of uncertain or stochastic parameters. Optimization procedures are increasingly helpful as the size and complexity of tractable problems grow. Ignoring fundamentally random characteristics may, however, limit the usefulness of an optimal solution. Stochastic programs that explicitly consider randomness may be much more beneficial than deterministic approximations for actual operations.

The expected objective value advantage of using a stochastic solution over a deterministic program solution is called the *value of the stochastic solution* (VSS). This quantity represents the added value of using a stochastic model. The VSS is related to the expected value of perfect information (EVPI), which measures the potential added value from determining which outcomes might actually occur, but VSS and EVPI can be quite different. In the examples below, we illustrate how they differ and why VSS may be more relevant than EVPI for modelers who cannot obtain further information about the future.

*The author's work was supported in part by the National Science Foundation under Grant DDM-9215921.

The value of a stochastic programming model stems from the ability to represent solutions that hedge against multiple possible future outcomes. In deterministic optimization, in particular, linear programming, optimal solutions tend toward extreme point solutions which rely on a limited set of activities (basic variables) and force a solution to meet critical constraints tightly. Stochastic program solutions, however, allow for broader sets of activities and naturally impose penalties that enable solutions to meet critical constraints with some cushion to avoid costly violations.

These features give stochastic programming solutions a significant practical validation advantage over deterministic model solutions. The deterministic result with few activities and tight critical constraints is often adjusted or rejected because of the practitioners' intuitive feeling that the result is not sufficiently diversified and may lead to large penalties when small deviations violate constraints. The stochastic program solution's inclusion of diversification and constraint violation penalties gives it immediate validity.

This paper explores the ability of stochastic programs to diversify among activities and to weigh constraint violations as they hedge against varying outcomes. We consider two simple examples. In each case, a deterministic model cannot provide adequate solutions. The VSS becomes quite significant.

The first example, in section 2, illustrates the VSS through a financial planning model concerning a child's higher education. The model includes risk aversion in an objective function, multiple period decisions, and linear constraints. The second example, in section 3, concerns the design of a part with manufacturing and end-use variability. The model includes nonlinear constraints and illustrates the dangers of setting constraints that can be violated with small parameter variations.

These simple examples highlight the practical advantages of stochastic programming models over deterministic models. Other stochastic programs may have much more complicated structures, but their added values over deterministic models remain the core hedging advantages of diversified activities and economically based safety margins on critical constraints. Section 4 briefly discusses these advantages in other application areas and discusses the additional costs in terms of model development and computation that must be weighted against the VSS.

2. Financial planning and the value of the stochastic solution

Financial models are natural for illustrating the value of diversification among different potential investments. Investors clearly recognize this value, making financial decision making one of the largest application areas of stochastic programming. Many references can be found in, for example, Golub et al. [11], Mulvey and Vladimirov [20], Zenios [27], and Ziemba and Vickson [28].

We consider a simple example that illustrates basic stochastic programming properties and, in particular, the value of diversification. The random variables reflect uncertain investment yields. The role of the stochastic program is to hedge against

poor outcomes by maximizing an expected objective function that is concave to represent aversion to risk.

For the current problem, suppose we wish to provide for a child's college education Y years from now. We currently have $\$w$ to invest in any of K investments. After Y years, we will have a wealth of $\$W$ that we would like to have exceed the tuition goal of $\$G$. We suppose that we can change investments every y years, so we have $T = Y/y$ investment periods. For our purposes here, we ignore transaction costs and taxes on income, although these considerations would be important in practice. The model can, in fact, easily incorporate them.

In formulating the problem, we must first describe our objective in mathematical terms. Generally, we can expect that the investor has some aversion to risk and that increases in wealth below G are more valuable to the investor than the same increase at a wealth above G . The simplest way to incorporate this form of utility into a model is to assume a piecewise linear utility function. For this, we suppose that exceeding $\$G$ after Y years would be equivalent to an income i of the excess while not meeting the goal would lead to a cost q of the amount short. This gives us the concave utility function in figure 1. Many other forms of nonlinear utility functions are of course possible, such as the popular von Neumann–Morgenstern isoelastic utility function (see, e.g., Ingersoll [14]). See Kallberg and Ziemba [16] for a description of their relevance in financial planning and of the effect of varying utilities on decisions.

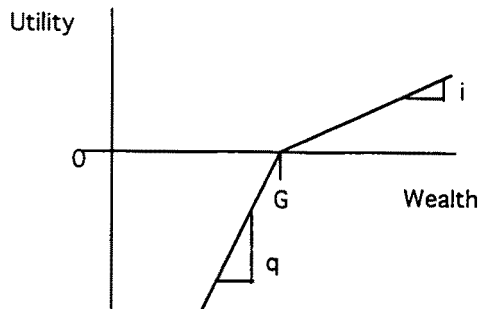


Figure 1. Utility function for value of wealth above and below goal G .

Discontinuous utilities, such as a function which is zero for any wealth below G and one for values above G , can also be used. This utility function corresponds to maximizing the probability of reaching the goal. Constraints on the probability of reaching the goal can also be included. The inclusion of these *chance constraints* (see Charnes and Cooper [5]) or *probabilistic constraints* (see Prékopa [23]) depends on the investor's preferences. The same type of diversification value would still result from these alternative models.

The major uncertainty in this model is the return on each investment k within each period t . To illustrate the effects of including alternative returns, we use a simple example with $K = 2$ possible investments types, stocks (1) and government securities (bonds) (2). We begin by setting Y at fifteen years and allow investment changes every five years so that $T = 3$.

From examining the data, suppose we assume that in each five-year period it is equally likely to have one of two outcomes: (inflation adjusted) returns of 1.25 for stocks and 1.14 for bonds or 1.06 for stocks and 1.12 for bonds. With three periods, the result is a tree with eight branches or scenario paths, corresponding to the possible future scenarios (see figure 2). We first consider the common optimization approach to replace these uncertain quantities by their expectations. Now, if we take expectations, then stocks have an expected return in each five-year period of 1.155 and bonds have an expected return of 1.13.

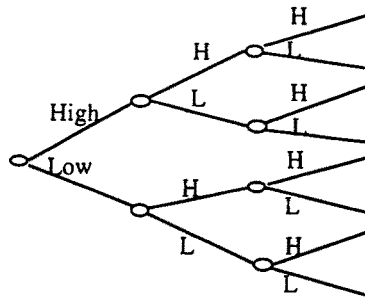


Figure 2. Tree of possible outcomes for high and low stock returns in each period.

The remaining data are the initial wealth $w = 55,000$, the target value $G = 80,000$, the surplus reward $i = 1$, and the shortage penalty $q = 4$. The deterministic *mean value problem* is to maximize the surplus minus penalty over the three periods. The decisions are $x(k, t)$ for investment in k at time t . The resulting linear program is

$$\begin{array}{llll}
 \max & & v & -4s \\
 \text{s.t.} & x(1,1) + x(2,1) & & = 55, \\
 & -1.155x(1,1) - 1.13x(2,1) + x(1,2) + x(2,2) & & = 0, \\
 & & -1.155x(1,2) - 1.13x(2,2) & +x(1,3) + x(2,3) = 0, \\
 & & 1.155x(1,3) + 1.13x(2,3) & -v + s = 80, \\
 & x(1,1), x(2,1) & x(1,2), x(2,2), x(1,3), x(2,3) & v \quad s \geq 0,
 \end{array} \tag{1}$$

which has a solution with $x(1, 1) = 55$, $x(1, 2) = 63.525$, $x(1, 3) = 73.372$, and $v = 4.74$ to yield a value of 4.74. This result is, of course, deceiving because the expected

return will not occur. If this solution is followed (to invest all funds in stock) with the two-outcome distribution we assume, then the result is much different in expectation.

In this case, we consider each of the eight possible outcomes or *scenarios* over three periods and weigh each by a probability of 0.125. Now, the return on stock is 1.953 with probability 0.125, 1.65625 with probability 0.375, 1.4045 with probability 0.375, and 1.191 with probability 0.125. Notice that we need a return of $80/55 = 1.45$ to reach the goal. Investing in stock alone misses this goal with probability 0.5. The expected value of the utility is in fact -3.8 , which we call *EMV* for the *expectation of the mean value solution*.

The other common solution approach is to find the optimal decision under each of the future scenarios. In figure 2, this policy would correspond to making a separate decision for each of the eight separate paths from the root to the leaves in the decision tree. This would be the strategy if we could postpone our decision or knew the future with certainty. This perfect information (or, in stochastic programming, following Madansky [16], the *wait-and-see* solution (*WS*)) is to invest in whatever has the highest yield in each period. In this case, we would invest in stocks if stocks increase 25% in a five-year period (bonds increase 14%) and we would invest in bonds if bonds increase 12% (stocks increase 6%). Thus, with probability 0.125, we receive a return of $(1.25)^3 = 1.953$, with probability 0.375, we receive a return of $(1.12)^3 = 1.405$, with probability 0.375, we receive a return of $1.12 * (1.25)^2 = 1.75$, and, with probability 0.125, we receive a return of $(1.12)^2 * 1.25 = 1.568$. The result is that we would have an expected utility of $WS = 10.5$. It is quite difficult to imagine how this analysis can be used, however. The recommendations depend on the realized returns which, unfortunately, are not known a priori.

Now, let us consider a formulation that considers the uncertainty explicitly in a single mathematical program. We have two alternative methods for writing the formulation in general. They correspond either to taking each outcome in each period separately, or to model the scenarios more explicitly. We describe the general forms of these formulations, denoting the random return by $r(k, t) = r(k, t, \omega)$, where ω is some underlying random element in Ω . We can think of Ω as the set of all possible outcome paths from the initial period to the end of the horizon. Decisions are associated with each path or scenario and denoted by $x(k, t) = x(k, t, \omega)$. The final wealth on each path is $W = W(\omega)$.

A key point about this investment model is that we cannot completely observe the path, the random element ω , when we make our decisions $x(k, t, \omega)$. We can only observe the returns that have already taken place through the branches up to time t . In stochastic programming, we say that we cannot *anticipate* every possible outcome so that our decisions are *nonanticipative* of future outcomes. Before the first period, this restriction corresponds to saying that we must make fixed investments $x(k, 1)$ for all paths $\omega \in \Omega$.

After making the initial investment decisions, suppose there are N_1 different possible outcomes in the first investment period. These outcomes correspond to

partitioning the set of all paths Ω into sets $\Omega_1^1, \dots, \Omega_{N_1}^1$ corresponding to these different initial outcomes. Decisions for a given value of ω must be the same for every ω within the set. We can therefore describe our second period decisions just in terms of the Ω_i^1 that occurs in the first period. We write these decision variables as $x(k, 2, i)$, where $i = 1, \dots, N_1$. In figure 2, these variables would be $x(k, 2, 1)$ for Ω_1^1 , the set of the top four paths, and $x(k, 2, 2)$ for the Ω_2^1 , the set of the bottom four paths.

We can continue this process by splitting each set of paths at time $t-1$ into i_t distinct sets of paths that are still combined at time t . Assuming the number of new branches is the same for each t , we can then describe the nonanticipative decisions as $x(k, t+1, i_1, \dots, i_t)$, which depends on the outcomes $r(t, i_1, \dots, i_t)$. In the following, for simplicity, we assume that each set of outcomes i_1, \dots, i_{t-1} up to time $t-1$ leads to N_t outcomes at time t . In all T periods, we would then have $N_1 \cdot N_2 \cdots N_T$ different possible outcomes, corresponding to all the paths in the problem. To illustrate how quickly these problems can grow, note that, with just ten outcomes per period, we would obtain 10^{10} outcomes in 10 periods.

The probabilities of each outcome at time t correspond to the sum of the path probabilities with a given set of paths. We let the probability of the i th outcome in period t given outcomes i_j in periods $j = 1, \dots, t-1$ be $p(t, i_1, \dots, i_t)$. Note that we must have $\sum_{i=1}^{N_t} p(t, i_1, \dots, i_{t-1}, i) = 1$ for all t and i_1, \dots, i_{t-1} .

We can then state a general formulation for our problem. We wish to find

$$\begin{aligned}
 \max \quad z = & \sum_{i_1, \dots, i_T} p(T, i_1, \dots, i_T) (i v(i_1, \dots, i_T) - qs(i_1, \dots, i_T)) \\
 \text{s.t.} \quad & \sum_{k=1}^K x(k, 1) = w, \\
 & \sum_{k=1}^K r(k, t, i_1, \dots, i_{t-1}) x(k, t-1, i_1, \dots, i_{t-2}) \\
 & - \sum_{k=1}^K x(k, t, i_1, \dots, i_{t-1}) = 0, \quad \text{for all } (i_1, \dots, i_{t-1}), \\
 & \hspace{15em} t = 2, \dots, T; \\
 & \sum_{k=1}^K r(k, T, i_1, \dots, i_T) x(k, T, i_1, \dots, i_{T-1}) \\
 & - v(i_1, \dots, i_T) + s(i_1, \dots, i_T) = G, \\
 & \hspace{10em} x(k, t, i_1, \dots, i_{t-1}) \geq 0, \\
 & \hspace{10em} v(i_1, \dots, i_T) \geq 0, \\
 & \hspace{10em} s(i_1, \dots, i_T) \geq 0, \quad \text{for all } 1 \leq i_t \leq N_t; \\
 & \hspace{15em} 1 \leq k \leq K; \\
 & \hspace{15em} 1 \leq t \leq T,
 \end{aligned} \tag{2}$$

where $v(i_1, \dots, i_T)$ represents any amount above the goal and $s(i_1, \dots, i_T)$ represents any shortfall from the target value G .

With the data given above, each period has two new alternative outcomes, so $N_t = 2$ for $t = 1, 2, 3$. Each $p(i_1, i_2, i_3)$ is 0.125. The resulting specification of (2) is

$$\begin{aligned}
 \max z = & \sum_{i_1=1}^2 \sum_{i_2=1}^2 \sum_{i_3=1}^2 0.125(v(i_1, i_2, i_3) - 4s(i_1, i_2, i_3)) \\
 \text{s.t.} \quad & x(1, 1) + x(2, 1) = 55, \\
 & -1.25x(1, 1) - 1.14x(1, 2) + x(1, 2, 1) + x(2, 2, 1) = 0, \\
 & -1.06x(1, 1) - 1.12x(1, 2) + x(1, 2, 2) + x(2, 2, 2) = 0, \\
 & -1.25x(1, 2, 1) - 1.14x(2, 2, 1) + x(1, 3, 1, 1) + x(2, 3, 1, 1) = 0, \\
 & -1.06x(1, 2, 1) - 1.12x(2, 2, 1) + x(1, 3, 1, 2) + x(2, 3, 1, 2) = 0, \\
 & -1.25x(1, 2, 2) - 1.14x(2, 2, 2) + x(1, 3, 2, 1) + x(2, 3, 2, 1) = 0, \\
 & -1.06x(1, 2, 2) - 1.12x(2, 2, 2) + x(1, 3, 1, 2) + x(2, 3, 1, 2) = 0, \\
 & 1.25x(1, 3, 1, 1) + 1.14x(2, 3, 1, 1) - v(1, 1, 1) + s(1, 1, 1) = 80, \quad (3) \\
 & 1.06x(1, 3, 1, 1) + 1.12x(2, 3, 1, 1) - v(1, 1, 2) + s(1, 1, 2) = 80, \\
 & 1.25x(1, 3, 1, 2) + 1.14x(2, 3, 1, 2) - v(1, 2, 1) + s(1, 2, 1) = 80, \\
 & 1.06x(1, 3, 1, 1) + 1.12x(2, 3, 1, 2) - v(1, 2, 2) + s(1, 2, 2) = 80, \\
 & 1.25x(1, 3, 2, 1) + 1.14x(2, 3, 2, 1) - v(2, 1, 1) + s(2, 1, 1) = 80, \\
 & 1.06x(1, 3, 2, 1) + 1.12x(2, 3, 2, 1) - v(2, 1, 2) + s(2, 1, 2) = 80, \\
 & 1.25x(1, 3, 2, 2) + 1.14x(2, 3, 2, 2) - v(2, 2, 1) + s(2, 2, 1) = 80, \\
 & 1.06x(1, 3, 2, 2) + 1.12x(2, 3, 2, 2) - v(2, 2, 2) + s(2, 2, 2) = 80, \\
 & x(k, t, i_1, \dots, i_{t-1}) \geq 0, \quad v(i_1, i_2, i_3) \geq 0, \quad s(i_1, i_2, i_3) \geq 0, \\
 & \text{for all } k, t, i_1, i_2, i_3.
 \end{aligned}$$

Another formulation approach to multistage problems is to consider decisions as dependent on each complete path or scenario ω . We must still maintain non-anticipativity, but this time we do so explicitly in the formulation via constraints. This corresponds to splitting the tree from figure 2 into the form in figure 3. In that case, all the paths are separate but are grouped together in every period where paths have the same history. The result is an alternative formation:

$$\begin{aligned}
 \max z = & \sum_{\omega} p(\omega)(iv(\omega) - qs(\omega)) \\
 \text{s.t.} \quad & \sum_{k=1}^K x(k, 1, \omega) = w, \quad \forall \omega \in \Omega;
 \end{aligned}$$

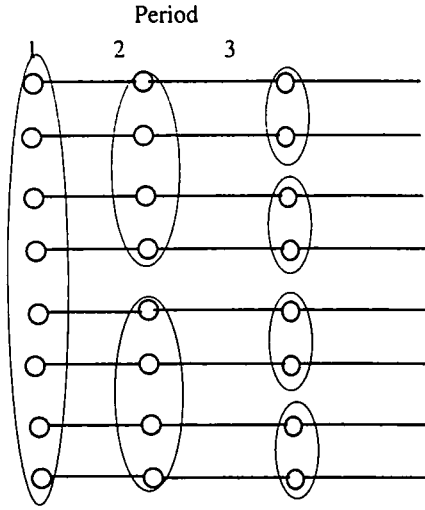


Figure 3. Separate paths with linking nonanticipativity constraints.

$$\begin{aligned}
 \sum_{k=1}^K r(k, t, \omega) x(k, t-1, \omega) - \sum_{k=1}^K x(k, t, \omega) &= 0, \quad \forall \omega \in \Omega; \\
 & \quad t = 2, \dots, T; \\
 \sum_{k=1}^K r(k, T, \omega) x(k, T, \omega) - v(\omega) + s(\omega) &= G; \\
 \left(\sum_{\omega' \in \Omega'_{I(\omega, t)}} p(\omega') x(k, t, \omega') \right) - \left(\sum_{\omega' \in \Omega'_{I(\omega, t)}} p(\omega') \right) x(k, t, \omega) &= 0, \quad \forall 1 \leq k \leq K; \\
 & \quad \forall 1 \leq t \leq T; \quad \forall \omega \in \Omega; \\
 x(k, t, \omega) \geq 0, & \quad v(\omega) \geq 0, & \quad s(\omega) \geq 0; \\
 \forall 1 \leq k \leq K; & \quad \forall 1 \leq t \leq T; & \quad \forall \omega \in \Omega;
 \end{aligned} \tag{4}$$

where $I(\omega, t) = \{i_1, \dots, i_{t-1}\}$ such that $\omega \in \Omega'_{i_1, \dots, i_{t-1}}$. Note that the last equality constraint indeed forces all decisions within the same group at time t to be the same. We use the form above because it represents a projection constraint that is both computationally useful (see, e.g. [20]) and allows simple representation of optimality conditions (see Dempster [8]). Formulation (4) has a special advantage for the problem here because these *nonanticipativity* constraints are the only constraints linking the separate scenarios. Without them, the problem would decompose into a separate problem for each ω , maintaining the structure of that problem.

In modeling terms, the nonanticipativity constraint makes it relatively easy to move from a deterministic model to a stochastic model of the same problem. The

addition of the scenario indicators and nonanticipativity constraints are the only additions to a deterministic model. Given the ease of this modeling effort, standard optimization procedures can be simply applied to this problem. However, as we noted above, the number of scenarios can become extremely large. The model in (4) also has more constraints than (2) to compensate for a loosening of the restrictions among scenarios with similar pasts. Standard methods may not be able to solve the problem in any reasonable fashion, necessitating other techniques.

In this financial problem, it is particularly worthwhile to try to exploit the underlying structure of the problem without the nonanticipativity constraints. This relaxed problem may be solved as a generalized network that allows the use of efficient network techniques (see Mulvey and Vladimirou [20]).

With either formulation (2) or (4), in completing the model, some decisions must be made about the possible set of outcomes or scenarios and the coarseness of the period structure, i.e. the number of periods T allowed for investment. We must also find probabilities to attach to outcomes within each of these periods. These probabilities are often approximations (see Birge and Wets [4] and Kall et al. [15]). A key observation is that the important step is to include stochastic elements at least approximately in order to obtain any form of a hedging policy. Deterministic solutions most often give misleading results, as illustrated with the example here.

Solving (3) yields an optimal expected utility value of -1.52 . We will call this value RP for the expected *recourse problem* solution value. The optimal solution (in thousands of dollars) appears in table 1. We have labeled the scenario paths in figure 2 from 1 to 8 (top to bottom).

Table 1

Optimal solution with a 3-period stochastic program.

Period, Scenario(s)	Stocks	Bonds
1, 1-8	41.5	13.5
2, 1-4	65.1	2.17
2, 5-8	36.7	22.4
3, 1,2	83.8	0.0
3, 3,4	0.0	71.4
3, 5,6	0.0	71.4
3, 7,8	64.0	0.0
Scenario	Above G (v)	Below G (s)
1	24.8	0.0
2	8.87	0.0
3	1.42	0.0
4	0.0	0.0
5	1.42	0.0
6	0.0	0.0
7	0.0	0.0
8	0.0	12.2

In this solution, the initial investment is heavily in stocks (\$41,500), with only \$13,500 in government securities. Notice the reaction to first-period outcomes, however. In the case of scenarios 1–4, stocks are even more prominent, while scenarios 5–8 reflect a more conservative bond portfolio. In the last period, notice how the investments are either completely in stocks or completely in bonds. This is a general trait of one-period decisions. The optimization leads to an extreme point in which only one type of investment is made. The main advantage of the multiperiod stochastic formulation is indeed to enable a hedging solution involving multiple investment types.

The solution in table 1 shows only stock investment in scenarios 1, 2 because there is no risk of missing the target. In scenarios 3–6, stock investments may cause one to miss the target, so they are avoided. In scenarios 7, 8, the only hope of reaching the target is through stocks. The stock investment is not, therefore, a monotonic function of wealth. A general observation is that any monotonic decision rule based on the current state is suboptimal.

We compare the results of table 1 to the deterministic model, where all investments would be in stock in each period because of its high expected return. If this policy is followed in each period, we would realize an expected utility of $EMV = -3.84$, while the stochastic program value $RP = -1.52$. The difference between these quantities is the *value of the stochastic solution* (introduced in Birge [1]),

$$VSS = RP - EMV = -1.52 - (-3.84) = 2.32. \quad (5)$$

Another comparison is in terms of the probability of reaching the goal, which might be an alternative objective. Notice that the stochastic program solution reaches the goal 87.5% of the time, while the mean value solution reaches the goal only 50% of the time. In this case, the value of the stochastic solution is even more significant.

Some distinction should be made between the value of the stochastic solution and the *expected value of perfect information*, which compares the recourse problem value (or maximum expected utility RP) to the expectation of scenario solution values that would be obtained if the future was known perfectly (WS). Here, the expected value of perfect information is

$$EVPI = WS - RP = 10.5 - (-1.52) = 12.0.$$

In this case, $EVPI > VSS$, but in many cases (see Birge [1]), we may have $VSS > EVPI$. In fact, since $WS > RP > EMV$, VSS and $EVPI$ are only assured of being the same when $WS = EMV$. Either could be zero, while the other is positive.

Another option in practice is to formulate a simpler model by limiting the horizon. This may seem quite reasonable since indeed we are only interested now in the first-period decision. Consider then a two-period, 10-year model. In this case, we need to determine conditions on the end of the horizon to make our shorter time horizon model reflect the original model as closely as possible. This is the problem

Table 2

Optimal solution with a 2-period, 10-year stochastic program.

Period, Scenario(s)	Stocks	Govt secs
1, 1-4	9.8	45.2
2, 1,2	0.0	63.8
2, 3,4	17.1	44.0
Scenario	Above G	Below G
1	1.28	0.0
2	0.0	0.0
3	0.0	0.0
4	0.0	4.12

of mitigating *end effects* (see Grinold [12]). It seems reasonable here to choose a target value that will ensure our achieving the original \$80,000 target value at year 15. We therefore make our 10-year target equal to \$71,400. Solving this two-period, 10-year problem with all other data as in the original three-period model, we obtain the optimal solution in table 2.

Notice how different the 10-year solution is from the 15-year solution. We now predominantly invest in bonds in the first period because we have not fully considered the chances of later recovering from initial poor stock investments. In this case, in five years, with probability one half, we have \$63,800 and, with probability one half, we have \$61,100. Again, we may solve the model with two five-year periods and from these wealth levels. The result is an expected utility of 1.87 in the first case and -5.77 in the second case. The overall expected utility is -3.9 . This expected utility is quite close to that obtained by the completely deterministic model. It demonstrates the importance of the horizon.

In general, the middle period in the first stochastic programming model in (3) is useful because it allows for a type of steady state or turnpike policy (see Birge and Dempster [3]) while the first and last periods have start-up or shut-down characteristics. A general result is that for stochastic programs over a long horizon, much of the solution in middle periods is relatively independent of the beginning and ending boundary conditions. This part of the solution is like a turnpike that would be followed in going between major cities. The beginning and ending period solutions give directions from the origin to the turnpike and from the turnpike to the destination, but the turnpike directions remain unchanged for most changes in origin or destination. To allow for this type of turnpike behavior, it is often necessary to have at least a three-period model. Two-period solutions as in table 2 tend to reflect starting or ending behavior that may be far from optimal.

In closing this section, note that the mathematical form of (2) or (4) actually represents a broad class of control problems. In fact, it is basically equivalent to any control problem governed by a linear system of differential or difference equations.

We have merely taken a discrete time approach to this problem. The approach can be applied to the control of a wide variety of electrical, mechanical, chemical, and economic systems. We merely redefine state variables (now, wealth) in each time period and the controls (investment levels). The random gain or loss is reflected in the return coefficients. Typically, these types of control problems would have nonlinear (e.g. quadratic) costs associated with the control in each time period. This presents no complication for our purposes, so we may include any of these problems as potential applications.

In the current problem, we have a limited set of possible outcomes in each period. In practical problems, we could not expect such a small finite set of realizations. One of the main steps in stochastic program modeling is how to select such sets of scenarios and how to compare their use with what may actually happen. One goal is to create approximations with a small number of realizations that bound the expected objective value from above and below (see Birge and Wets [4]). Another goal may be to obtain asymptotic convergences through sampling (see Dantzig and Glynn [7], Ermoliev [9], and Higle and Sen [13]).

In this financial model, we used a piecewise linear utility function to obtain a linear program. This form allows very large-scale computations (see Wets [26]) and can lead to great efficiencies over standard linear programming techniques (see Birge [2] and Gassmann [10]). Linear programming structure is not, however, necessary in stochastic programs. Large-scale nonlinear stochastic programming problems (see Nielsen and Zenios [21]) can now be solved quite efficiently. The next section gives one example of the range of nonlinear models in stochastic programming.

3. Design for manufacturing quality

This section illustrates a common engineering design problem that we model as a stochastic program. The problem demonstrates nonlinear functions in stochastic programming and provides further evidence of the importance of the stochastic solution. The key advantage of the model in this situation is that the stochastic program does not force an extremal solution that may have significant consequences due to minor variations that violate important constraints.

Consider a designer deciding various product specifications to achieve some measure of product cost and performance. The specifications do not, however, completely determine the characteristics of each manufactured product. Key characteristics of the product are often random. For example, every item includes variations due to machining or other processing. Each consumer also does not use the product in the same way. Cost and performance characteristics thus become random variables.

Deterministic methods may yield costly results that are only discovered after production has begun. From this experience, designing for quality and consideration of variable outcomes has become an increasingly important aspect of modern manufacturing (see, for example, Taguchi et al. [25]). In this development, the methods

of Taguchi have been widely used (see also Taguchi [24]). These approaches can, in fact, be seen as examples of stochastic programming, although they are not often described this way.

In this section, we give a small example of the uses of stochastic programming in manufacturing design and show how the general stochastic programming approach can be applied. We base our analysis on actual performance measures (or predictions), whereas the Taguchi methods generally attach surrogate penalties to deviations from nominal parameter values. The important characteristic in both approaches, however, is that they lead to designs that protect against violating critical constraints with small variations.

We consider the design of a simple axle assembly for a bicycle cart. The axle has the general appearance shown in figure 4.

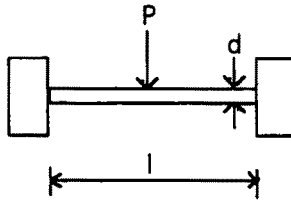


Figure 4. An axle of length l , diameter d , with a central load P .

The designer must determine the specified length l and diameter d of the axle. We use inches to measure these quantities and assume that other dimensions are fixed. Together, these quantities determine the performance characteristics of the product. The goal is to determine a combination that will give the greatest expected profit.

The initial costs are for manufacturing. We assume a single process with no alternative technologies although, in practice, several processes might be available.

Due to machining variations, the actual dimensions of a finished product are not exactly those that are specified. For our example, we suppose that the length l can be produced exactly but that the diameter d is a random variable $d(\bar{d})$ that depends on a specified mean value or machine setting \bar{d} . We assume a triangular distribution for $d(\bar{d})$ on $[0.9\bar{d}, 1.1\bar{d}]$. This distribution has a density

$$f_{\bar{d}}(d) = \begin{cases} \frac{100}{\bar{d}^2} (d - 0.9\bar{d}) & \text{if } 0.9\bar{d} \leq d < \bar{d}; \\ \frac{100}{\bar{d}^2} (1.1\bar{d} - d) & \text{if } \bar{d} \leq d \leq 1.1\bar{d}; \\ 0 & \text{otherwise.} \end{cases} \quad (6)$$

The decision is then to determine l and \bar{d} , subject to certain limits $l \leq l^{\max}$ and $\bar{d} \leq d^{\max}$, in order to maximize expected profits. For revenues, we assume that, if the product is profitable, we sell as many as we can produce. This amount is fixed by labor and equipment regardless of the size of the axle. We therefore only wish to determine the maximum selling price that generates enough demand for all production. From marketing studies, we determine that this maximum selling price depends on the length, and is expressed as

$$s(1 - e^{-0.1l}), \quad (7)$$

where s is a maximum possible for any such product.

Our production costs for labor and equipment are assumed fixed, so that only material cost is variable. This cost is proportional to the mean values of the specified dimensions since material is acquired before the actual machining process and producing many units leads to low unit material cost variance. For c , the cost of a single axle material unit, the total manufacturing cost for an item is then

$$c \left(\frac{l\pi\bar{d}^2}{4} \right). \quad (8)$$

In this simplified model, we assume that no quantity discounts apply in the production process.

Other costs are incurred after the product is made due to warranty claims and potential future sales losses from product defects. These costs are often called quality losses. In stochastic programming terms, these are the recourse costs. Here, the product may perform poorly if the axle becomes bent or broken due to excess stress of deflection. The stress limit, assuming a steel axle and 100-pound maximum central load, is

$$\frac{l}{d^3} \leq 39.27. \quad (9)$$

For deflection, we use a maximum 2000 rpm speed to obtain

$$\frac{l^3}{d^4} \leq 63169. \quad (10)$$

When either of these constraints is violated, the axle generally deforms. The cost of violation may vary according to the exposure time beyond the limits and the actual load placed on the axle. The main direct cost will be replacement if the axle fails, but other costs due to damages (injury insurance claims, for example) and loss of future market may push the cost of violation higher. For a single axle, we put these costs together into an expected cost for not meeting the stress and deflection constraints. We assume the cost is proportional to the square of the violation, written as

$$Q(l, \bar{d}, d) := \min_y \left\{ wy^2 \text{ s.t. } \frac{l}{d^3} - y \leq 39.27, \frac{l^3}{d^4} - 300y \leq 63169 \right\}, \quad (11)$$

where y is, therefore, the maximum of stress violation and (to maintain similar units) $1/300$ of the deflection violation.

The cost in (11) is for a single axle with known dimensions. The full expected recourse costs, given l and \bar{d} , is found by integrating over d as

$$Q(l, \bar{d}) = \int_d Q(l, \bar{d}, d) f_{\bar{d}}(d) dd, \tag{12}$$

which can also be written here as

$$Q(l, \bar{d}) = w \int_{0.9\bar{d}}^{1.1\bar{d}} (100/\bar{d}^2) \min\{d - 0.9\bar{d}, 1.1\bar{d} - d\} \left[\max \left\{ 0, \left(\frac{l}{d^3} \right) - 39.27, \left(\frac{l^3}{300d^4} \right) - 210.36 \right\} \right]^2 dd. \tag{13}$$

Note that, if the cost calculation in (11) and the distribution model for axle diameters are correct, then the expected value in (13) is quite close to the actual future unit cost for a large number of sales. The risk-neutral approach becomes quite appropriate in this case (in contrast to the financial planning model) because the risk or variance in the objective becomes quite small with large production quantities. Other models with large numbers of realizations corresponding to individual customers or products would also fit this risk-neutral form, while models with single random realizations in each period would generally include some form of risk aversion as in financial planning objectives.

The overall problem is to find

$$\begin{aligned} &\max (\text{total revenue per item} - \text{manufacturing cost per item} \\ &\quad - \text{expected future cost per item}). \end{aligned} \tag{14}$$

Mathematically, we write this as

$$\begin{aligned} \max z(l, \bar{d}) &= s(1 - e^{-0.1l}) - c \left(\frac{l\pi\bar{d}^2}{4} \right) - Q(l, \bar{d}) \\ \text{s.t. } &0 \leq l \leq l^{\max}, \quad 0 \leq \bar{d} \leq d^{\max}. \end{aligned} \tag{15}$$

In stochastic programming terms, this formulation gives the deterministic equivalent problem to the stochastic program for minimizing the current value for the design decision plus future reactions to deviations in the axle diameter. Standard optimization procedures can be used to solve this problem. Assuming maximum values of $l^{\max} = 35$, $d^{\max} = 1.25$, a maximum sales price of \$10 ($s = 10$), a material cost of \$0.025 per cubic inch ($c = 0.025$), and a unit penalty $w = 1$, an optimal

solution is found at $l^* = 33.6$, $\bar{d}^* = 1.038$, and $z^* = z(l^*, \bar{d}^*) = 8.94$. In this solution, the stress constraint is only violated when $0.9\bar{d} = 0.934 \leq d \leq 0.949 = (l/39.27)^{1/3}$.

We next consider the expected value problem where random variables are replaced with their means to obtain a deterministic problem. For this problem, we would obtain

$$\begin{aligned} \max z^{Det}(l, \bar{d}) &= s(1 - e^{-0.1l}) - c \left(\frac{l\pi\bar{d}^2}{4} \right) \\ &- \left[\max \left\{ 0, \left(\frac{l}{\bar{d}^3} \right) - 39.27, \left(\frac{l^3}{300\bar{d}^4} \right) - 210.36 \right\} \right]^2 \quad (16) \\ \text{s.t. } &0 \leq l \leq l^{\max}, \quad 0 \leq \bar{d} \leq d^{\max}. \end{aligned}$$

Using the same data as above, an optimal solution to (16) is $l^{Det} = 35.0719$, $\bar{d}^{Det} = 0.963$, and $z^{Det}(l^{Det}, \bar{d}^{Det}) = 9.07$. In this case, the deterministic solution makes the stress constraint active. As in any deterministic model, this solution pushes against one of the constraints. The result is that any randomness leads to frequent constraint violations and warranty or quality costs.

At first glance, it appears that the deterministic solution obtains a higher unit profit than the stochastic problem solution. However, this deterministic problem again paints an overly optimistic picture of the actual situation. The deterministic objective is (in the case of concave maximization) always an *overestimate* of the actual expected profit. As is often seen in practice, prediction based on mean value optimization problems are always biased. In this case, the true *expected* value of the deterministic solution is $z(l^{Det}, d^{Det}) = 5.88$. This problem then has a value of the stochastic solution equal to the difference between the expected value of the stochastic solution and the expected value of the deterministic solution of $z^* - z(l^{Det}, d^{Det}) = 3.06$. In other words, solving the stochastic program yields an expected profit increase of $3.06/5.88 = 52\%$ over solving the deterministic problem.

This problem is another example of how stochastic programming can be used. It has nonlinear functions and a future penalty cost on constraint violation, called *simple recourse* structure. In other problems, decisions may also be taken after the observation of the outcome to reduce this penalty. For example, we could inspect and then decide whether to sell the product. This often leads to tolerance settings and is indeed the focus of much of quality control.

The general stochastic program provides a framework for uniting design and quality control. Many loss functions can be used to measure performance degradation to help to improve designs in their initial stages. These functions may include the stress performance here, the Taguchi-type of quadratic loss, or methods based on reliability characterizations.

Most traditional approaches assume some form for the distribution as we have done here. This situation rarely matches practice, however. Approximations can

nevertheless be used that obtain bounds on the actual solution value so that robust decisions may be made without complete distributional information (see, for example, [4] and [15]).

5. Conclusions

This paper presented a brief discussion of stochastic programming models. We began with a simple example in financial planning that illustrated the value of the stochastic solution over that of a deterministic model solution. We also showed how this quantity is different from the expected value of perfect information and noted how the model could take advantage of linear programming methods.

This first model illustrated the advantage of a stochastic programming solution in terms of hedging with investments in two types of securities. Any deterministic approximation with only a single outcome in each period can only lead to one security investment. Diversification is never optimal. A stochastic programming model is indeed necessary to demonstrate the utility of this fundamental aspect of investment planning.

We then described a problem with highly nonlinear functions as an example of using stochastic programming for designing to meet quality goals. In this example, we also demonstrated the value of the stochastic solution in protecting against costly constraint violations. In contrast to the financial planning model, this model used a risk-neutral objective due to anticipated large production quantities.

These models illustrate the type of advantages possible in stochastic programs. They show the typical benefits of stochastic programs in terms of hedging and safety against constraint violations. The models each have a form of action before observation, observation, new action and, perhaps, repetition. This characteristic sequence fits most stochastic programming models.

While the advantages in stochastic programming are often quite clear, constructing stochastic programs usually requires information that has not been routinely gathered. Distributions and basic parameter values might not be known. Approximations deal with these difficulties by constructing models that use whatever information is known. The goal in these approximations and stochastic programming in general is always to use all information that is available and to make decisions as well informed as possible.

As the other papers in this volume indicate, there are many more examples of stochastic programming formulations. We should mention the extensive work in energy and power planning (see Dantzig and Glynn [6], Louveaux [18], and Noël and Smeers [22]) as just one additional example. Several other references appear in the survey by King [17]. Many more applications are open to stochastic programming as powerful new solution and modeling techniques become increasingly available.

References

- [1] J.R. Birge, The value of the stochastic solution in stochastic linear programs with fixed recourse, *Math. Progr.* 24(1982)314–325.

- [2] J.R. Birge, Decomposition and partitioning methods for multi-stage stochastic linear programs, *Oper. Res.* 33(1985)989–1007.
- [3] J.R. Birge and M.A.H. Dempster, Optimality conditions for match-up strategies in stochastic scheduling, Technical Report 92-58, Department of Industrial and Operations Engineering, The University of Michigan, Ann Arbor (1992).
- [4] J.R. Birge and R.J-B Wets, Designing approximation schemes for stochastic optimization problems, in particular, for stochastic programs with recourse, *Math. Prog. Study* 27(1986)54–102.
- [5] A. Charnes and W.W. Cooper, Chance-constrained programming, *Manag. Sci.* 5(1959)73–79.
- [6] G.B. Dantzig and P. Glynn (eds.), *Resource Planning under Uncertainty for Electric Power Systems* (National Science Foundation, Washington, DC, 1989).
- [7] G.B. Dantzig and P. Glynn, Parallel processors for planning under uncertainty, *Ann. Oper. Res.* 22(1990)1–21.
- [8] M.A.H. Dempster, On stochastic programming II: Dynamic problems under risk, *Stochastics* 25(1988)15–42.
- [9] Y. Ermoliev, Stochastic quasigradient methods and their applications to systems optimization, *Stochastics* 9(1983)1–36.
- [10] H.I. Gassmann, MSLiP: A computer code for the multistage stochastic linear programming problem, *Math. Progr.* 47(1990)407–423.
- [11] B. Golub, M. Holmer, R. McKendall, L. Pohlman and S.A. Zenios, Stochastic programming models for money management, *Euro. J. Oper. Res.*, to appear.
- [12] R.C. Grinold, Model building techniques for the correction of end effects in multistage convex programs, *Oper. Res.* 31(1983)407–431.
- [13] J. Hige and S. Sen, Stochastic decomposition: An algorithm for two-stage linear programs with recourse, *Math. Oper. Res.* 16(1991)650–669.
- [14] J.E. Ingersoll, Jr., *Theory of Financial Decision Making* (Rowman and Littlefield Publishers, 1986).
- [15] P. Kall, A. Ruszczyński and K. Frauendorfer, Approximations in stochastic programming, in: *Numerical Techniques for Stochastic Optimization*, ed. Y. Ermoliev and R.J-B Wets (Springer, Berlin, 1988) pp. 33–64.
- [16] J.G. Kallberg and W.T. Ziemba, Comparison of alternative utility functions in portfolio selection problems, *Manag. Sci.* 29(1983)1257–1276.
- [17] A. King, Stochastic programming problems: Examples from the literature, in: *Numerical Techniques for Stochastic Optimization*, ed. Y. Ermoliev and R.J-B Wets (Springer, Berlin, 1988) pp. 543–567.
- [18] F.V. Louveaux, A solution method for multistage stochastic programs with recourse with application to an energy investment problem, *Oper. Res.* 28(1980)889–902.
- [19] A. Madansky, Inequalities for stochastic linear programming problems, *Manag. Sci.* 6(1960) 197–204.
- [20] J. Mulvey and H. Vladimirou, Stochastic network optimization models for investment planning, *Ann. Oper. Res.* 20(1989)187–217.
- [21] S.S. Nielsen and S.A. Zenios, A massively parallel algorithm for nonlinear stochastic network problems, *Oper. Res.* 41(1993)319–337.
- [22] M.C. Noël and Y. Smeers, Nested decomposition of multistage nonlinear programs with recourse, *Math. Progr.* 37(1987)131–152.
- [23] A. Prékopa, Numerical solution of probabilistic constrained programming problems, in: *Numerical Techniques for Stochastic Optimization*, ed. Y. Ermoliev and R.J-B Wets (Springer, Berlin, 1988) pp. 123–140.
- [24] G. Taguchi, *Introduction to Quality Engineering* (Asian Productivity Center, Tokyo, 1986).
- [25] G. Taguchi, E.A. Alsayed and T. Hsiang, *Quality Engineering in Production Systems* (McGraw–Hill, New York, 1989).
- [26] R.J-B Wets, Large-scale linear programming techniques in stochastic programming, in: *Numerical Techniques for Stochastic Optimization*, ed. Y. Ermoliev and R.J-B Wets (Springer, Berlin, 1988) pp. 65–94.
- [27] S.A. Zenios (ed.), *Financial Optimization* (Cambridge University Press, Cambridge, 1993).
- [28] W.T. Ziemba and R.G. Vickson (eds.), *Stochastic Optimization Models in Finance* (Academic Press, New York, 1975).