Math Meth Oper Res (1999) 50:245–270

Mathematical Methods
of Operations Research
© Springer-Verlag 1999

# Portfolio optimization via stochastic programming: Methods of output analysis

**Jitka Dupačová**

Department of Probability and Mathematical Statistics, Charles University, Sokolovská 83, CZ-186 75 Prague, Czech Republic (e-mail: dupacova@karlin.mff.cuni.cz)

**Abstract.** Solutions of portfolio optimization problems are often influenced by errors or misspecifications due to approximation, estimation and incomplete information. Selected methods for analysis of results obtained by solving stochastic programs are presented and their scope illustrated on generic examples – the Markowitz model, a multiperiod bond portfolio management problem and a general strategic investment problem. The approaches are based on asymptotic and robust statistics, on the moment problem and on results of parametric optimization.

**Key words:** Portfolio optimization, stochastic programming, stability, postoptimality, worst-case analysis

## 1 Some early contributions

The main feature of the investment and financial problems is the necessity to make *decisions under uncertainty* and *over more than one time period*. The uncertainties concern the future level of interest rates, yields of stock, exchange rates, prepayments, external cashflows, inflation, future demand, liabilities, etc. There exist various stochastic models describing or explaining these parameters and they represent an important part of various procedures used to generate the input for decision models.

To build a decision model, one has to decide first about the purpose or goal; this includes identification of the uncertainties or risks one wants to hedge, of the hard and soft constraints, of the time horizon and its discretization, etc. The next step is the formulation of the model and generation of the data input. An algorithmic solution concludes the first part of the procedure. The subsequent interpretation and evaluation of the results may lead to model changes and, consequently, to a new solution or it may require a "what-if" analysis to get information about robustness of the results.

*Example 1 – The Markowitz model.* In conclusions of his famous paper [49] on portfolio selection, Markowitz stated that "what is needed is essentially a 'probabilistic' reformulation of security analysis". He developed a model for portfolio optimization in an uncertain environment under various simplifications. It is a static, single period model which assumes a frictionless market. It applies to small rational investors whose investments cannot influence the market prices and who prefer higher yields to lower ones and smaller risks to larger ones. Let us recall the basic formulation: The composition of portfolio of $I$ assets is given by weights of the considered assets, $x_i$, $i = 1, \dots, I$, $\sum_i x_i = 1$. The unit investment in the $i$-th asset provides the random return $\rho_i$ over the considered fixed period. The assumed probability distribution of the vector $\boldsymbol{\rho}$ of returns of all assets is characterized by a known vector of expected returns $E\boldsymbol{\rho} = \boldsymbol{r}$ and by a covariance matrix $\boldsymbol{V} = [\mathrm{cov}(\rho_i, \rho_j), i, j = 1, \dots, I]$ whose main diagonal consists of variances of individual returns.

This allows to quantify the "yield from the investment" as the expectation $r(\boldsymbol{x}) = \sum_i x_i r_i = \boldsymbol{r}^\top \boldsymbol{x}$ of its total return and the "risk of the investment" as the variance of its total return, $\sigma^2(\boldsymbol{x}) = \sum_{i,j} \mathrm{cov}(\rho_i, \rho_j) x_i x_j = \boldsymbol{x}^\top \boldsymbol{V}\boldsymbol{x}$. According to the assumptions, the investors aim at maximal possible yields and, at the same time, at minimal possible risks – hence, a typical decision problem with two criteria, "max" $\{r(\boldsymbol{x}), -\sigma^2(\boldsymbol{x})\}$. The mean-variance efficiency introduced by Markowitz is fully in line with general concepts of multicriteria optimization. Hence, mean-variance efficient portfolios can be obtained by solving various optimization problems, e.g.,

$$\max_{\boldsymbol{x} \in \mathcal{X}} \{\lambda \boldsymbol{r}^\top \boldsymbol{x} - 1/2 \boldsymbol{x}^\top \boldsymbol{V} \boldsymbol{x}\} \tag{1}$$

where the value of parameter $\lambda \geq 0$ reflects investor's risk aversion. In classical theory, the set $\mathcal{X}$ is defined by $\sum_i x_i = 1$ without nonnegativity constraints, which means that short sales are permitted.

It was the introduction of risk into the investment decisions which was the exceptional feature of this model and a real breakthrough. The Markowitz model became a standard tool for portfolio optimization. It has been applied not only to portfolios of shares, but also to bonds [57], to international loans [68], etc., even to asset and liability management with portfolio returns replaced by the *surplus*, cf. [51]. However, there are many questions to be answered: Modeling the random returns to get their expectations, variances and covariances and analysis of sensitivity of the investment strategy on these estimated input values, the choice of the value of $\lambda$, etc. From the point of view of optimization, an inclusion of *linear* regulatory constraints does not cause any serious problems. This, however does not apply to minimal transaction unit constraints which introduce 0–1 variables; e.g., [71]. In the interpretation and application of the results one has to be aware of the model assumptions (not necessarily fulfilled in real-life), namely, that it is a one-period model based on the buy-and-hold strategy applied between the initial investment and the horizon of the problem so that decisions based on its repeated use over more that one period can be far from a good, suboptimal dynamic decision, cf. [9].

At the same time, Roy [67] proposed to use the *Safety-First* criterion

$$\max_{\boldsymbol{x} \in \mathcal{X}} P(\boldsymbol{\rho}^\top \boldsymbol{x} \geq r_p) \tag{2}$$

with a given level $r_p$ of the required return of the portfolio and other criteria were developed along the same way: The *chance-constrained* criterion by Telser [73] is

$$\max_{\boldsymbol{x}\in\mathscr{X}}\{r(\boldsymbol{x})\mid P(\boldsymbol{\rho}^\top\boldsymbol{x}\geq r_p)\geq 1-\alpha\} \tag{3}$$

with a prescribed level of the total return $r_p$ and with a given probability $\alpha\in(0,1)$. The *quantile* criterion introduced by Kataoka [38]

$$\text{maximize } r_p \text{ subject to } \boldsymbol{x}\in\mathscr{X} \text{ and } P(\boldsymbol{\rho}^\top\boldsymbol{x}\geq r_p)\geq 1-\alpha \tag{4}$$

for a prescribed probability $\alpha\in(0,1)$ can be evidently related with quantification of the risk of the investment by its *value at risk*, VaR, usually computed for $\alpha=0.01$ or 0.05. The generally accepted decision criterion of maximal *expected utility* of the total return, $\max_{\boldsymbol{x}\in\mathscr{X}} Eu(\sum_i \rho_i x_i)$, offers further possibilities and together with (2), (3), (4), it can be clearly classified as a *static, single period stochastic program.*

Numerous early studies dealt with the properties of the portfolio optimization problems based on various criteria mentioned above; we refer to the collection [79] of original papers with commentaries, to [42], to the recent survey [13] and to textbooks [30, 35]. For results on the equivalence of problems (2) and (4) under general assumptions on the probability distribution of returns see [37].

Objections against the symmetry of variance of returns as a measure of risk has lead to various asymmetric risk definitions, such as the quadratic semivariance $E\{([\sum_i r_i x_i - \sum_i \rho_i x_i]^+)^2\}$; the disadvantage is the decrease in numerical tractability of the resulting optimization problems to be solved. Sharpe [72] suggested to define "risk" as the *mean absolute deviation*

$$m(\boldsymbol{x}) := E\left|\sum_i \rho_i x_i - \sum_i r_i x_i\right|. \tag{5}$$

This idea was developed and applied, e.g., in [7, 44]. The values of (5) are constant multiples of the standard deviations $\sigma(\boldsymbol{x})$ provided that the probability distribution of the random returns $\boldsymbol{\rho}$ is normal $\mathscr{N}(\boldsymbol{r},\boldsymbol{V})$. The advantage of this approach is that the sample form of the optimization problem can be transformed to a linear program for which the coefficients result directly from sample observations and the required asymmetry of risk can be easily incorporated.

The numerical tractability of the piece-wise linear criteria extends to the piece-wise *linear-quadratic risk measures* [40] and the idea of decisions efficient with respect to multiple criteria appears also in various *tracking models*, e.g. [14, 75].

The crucial question is an extension of these *static* models to dynamic ones. For portfolio optimization problems, it is sufficient to deal with discrete time models as, contrary to the continuous *trading* problems, the investment decisions are done in specific time instants. The first step is an adequate treatment of transaction costs and tax [52, 60].

We shall focus on modeling via *multiperiod and multistage stochastic programming*. Contrary to the stochastic control models with discrete time, cf. [35], the main interest lies in the initial, first-stage decisions. The history of applications of multiperiod or multistage stochastic programs in portfolio allocation and management dates back to Bradley and Crane [5]. The important contributions in the eighties, e.g., [6, 7, 15, 45, 46, 56, 71], reviewed also in [18], together with the progress in numerical methods, software and computer technologies (e.g., [12, 53] and references therein) have opened possibilities of large scale real life applications reported, e.g., in [9, 10, 12, 54] and in the collection [78]. The strength of stochastic programming is in open possibilities to support the asset and liability management and the risk management decisions under various circumstances which reflect the goals and the restrictions of the users. Its concepts imply that the emphasis lies on strategic decisions.

However, some of the stumbling blocks mentioned in the context of the Markowitz model persist and they even grow into new dimensions: An adequate reflection of the common assumptions concerning the market and the investors; the recognition of goals and restrictions and the related problems of the choice of the objective function, of the time discretization and modeling of constraints; the input data generation; the interpretation of the obtained results and an analysis of their sensitivity to assumptions, data, etc. We shall deal with the last mentioned group of problems, offering selected methods suitable for analysis of relevance and robustness of the obtained results.

## 2 Multistage stochastic programs

In the general *T-stage stochastic program* we think of a stochastic data process

$$\omega = \{\omega_1, \ldots, \omega_{T-1}\} \quad \text{or} \quad \omega = \{\omega_1, \ldots, \omega_T\} \tag{6}$$

whose realizations are (multidimensional) data trajectories in $(\Omega, \mathscr{F}, P)$ and of a vector decision process $\boldsymbol{x} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T\}$, a measurable function of $\omega$. The whole sequence of decisions and observations is, e.g.,

$$\boldsymbol{x}_1, \omega_1, \boldsymbol{x}_2(\boldsymbol{x}_1, \omega_1), \omega_2, \ldots, \boldsymbol{x}_T(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{T-1}, \omega_1, \ldots, \omega_{T-1}) \tag{7}$$

in addition to $\omega_T$ that may also contribute to the overall observed costs. The decision process is *nonanticipative* in the sense that decisions taken at any stage of the process do not depend on future *realizations* of the random parameters or on future decisions, whereas it is the past information and the given *probabilistic specification* $(\Omega, \mathscr{F}, P)$ of the process $\omega$ which are exploited. The dependence of the decisions solely on the history and on the probabilistic specification can be expressed as follows: Denote by $\mathscr{F}_{t-1} \subseteq \mathscr{F}$ the $\sigma$-field generated by the observations $\omega^{t-1,\bullet} := \{\omega_1, \ldots, \omega_{t-1}\}$ of the part of the stochastic data process that precedes stage $t$. The dependence of the $t$-th stage decision $\boldsymbol{x}_t$ only on these past observations means that $\boldsymbol{x}_t$ is $\mathscr{F}_{t-1}$-adapted or, in other words, that $\boldsymbol{x}_t$ is measurable with respect to $\mathscr{F}_{t-1}$. In each of the stages, the decision is limited by constraints that may depend on the previous decisions and observations.

Two formulations of multistage stochastic programming problems can be used. For general results concerning their equivalence see, e.g., [66], for an introductory survey see [20].

Let $\mathscr{X}_t$ be given nonempty sets in $R^{n_t}, t = 1, \ldots, T$, and denote

$$\mathscr{X}^t(\omega) = \{\boldsymbol{x}^{t\bullet} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_t) \in \mathscr{X}_1 \times \cdots \times \mathscr{X}_t : \\ f_{ti}(\boldsymbol{x}^{t\bullet}, \omega^{t-1,\bullet}) \le 0, \; i = 1, \ldots, m_t\} \tag{8}$$

the set of the $t$-th stage constraints, $t = 2, \ldots, T$, and by $f_0(\boldsymbol{x}, \omega)$ the overall cost connected with the decision process (7).

The *T-stage stochastic program* is to find

$$\boldsymbol{x}_1 \in \mathscr{X}_1 \text{ and } \boldsymbol{x}_t \quad \mathscr{F}_{t-1}\text{-measurable}, \; \boldsymbol{x}^{t\bullet} \in \mathscr{X}^t(\omega), \text{ a.s.}, \; t = 2, \ldots, T,$$

$$\text{that maximize } E\{f_0(\boldsymbol{x}_1, \boldsymbol{x}_2(\omega), \ldots, \boldsymbol{x}_T(\omega), \omega)\}. \tag{9}$$

The special choice of the function $f_0$ in (9) as an indicator function of a certain interval leads to the *probability objective function* of the form

$$P\{g_0(\boldsymbol{x}_1, \boldsymbol{x}_2(\omega), \ldots, \boldsymbol{x}_T(\omega), \omega) \in \mathscr{I}\}$$

where $\mathscr{I}$ is a given interval of desirable values of $g_0$; compare with (2). Similarly, the replacement of the constraints $\boldsymbol{x}^{t\bullet} \in \mathscr{X}_t(\omega)$, a.s., $t = 2, \ldots, T$, by the requirement that $\boldsymbol{x}^{t\bullet} \in \mathscr{X}_t(\omega), t = 2, \ldots, T$, holds true with a prescribed probability provides stochastic programs with *probabilistic* or *chance constraints*; compare with (4).

The *second formulation of the T-stage stochastic program* is based on a recursive evaluation of the overall objective function which allows us to write the multistage stochastic program as a sequence of nested two-stage programs. To this purpose, we denote by $P(\omega_t), t = 1, \ldots, T$, the marginal distributions and by $P(\omega_t|\omega^{t-1,\bullet})$ or $P(\omega_t|\omega_{t-1})$ for $t = 2, \ldots, T$, the conditional probability distributions of $\omega_t$ conditioned by the whole history $\omega^{t-1,\bullet}$ or by $\omega_{t-1}$, respectively. The $T$-stage stochastic program can be written as

$$\max E\{f_0(\boldsymbol{x}, \omega)\} := f_{10}(\boldsymbol{x}_1) + E_{P(\omega_1)}\{\varphi_{10}(\boldsymbol{x}_1, \omega_1)\} \tag{10}$$

subject to

$$\boldsymbol{x}_1 \in \mathscr{X}_1 \quad \text{and} \quad f_{1i}(\boldsymbol{x}_1) \le 0, \quad i = 1, \ldots, m_1,$$

where for $t = 2, \ldots, T, \varphi_{t-1,0}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{t-1}, \omega_1, \ldots, \omega_{t-1})$ is the optimal value of the stochastic program

$$\max f_{t0}(\boldsymbol{x}_t) + E_{P(\omega_t|\omega^{t-1,\bullet})}\{\varphi_{t0}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_t, \omega_1, \ldots, \omega_{t-1}, \omega_t)\} \tag{11}$$

with respect to $\boldsymbol{x}_t \in \mathscr{X}_t$ that fulfil

$$f_{ti}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{t-1}, \boldsymbol{x}_t, \omega_1, \ldots, \omega_{t-1}) \le 0, \quad i = 1, \ldots, m_t.$$

Here, $\varphi_{T,0} \equiv 0$ or is an explictly given function of $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_T, \omega_1, \ldots, \omega_T$. Again, all constraints involving random parameters hold almost surely. The main decision variable is $\boldsymbol{x}_1$ that corresponds to the first stage.

For *scenario-based* multistage stochastic programs one assumes that the probability distribution $P$ of $\omega$ is discrete one, concentrated in a finite number of points, say, $\omega^1, \ldots, \omega^S$. Accordingly, the supports $\mathcal{S}_t(\omega^{t-1,\bullet})$ of the conditional probability distributions of $P(\omega_t|\omega^{t-1,\bullet})$ are finite sets. The set of all considered scenarios is

$$\{\omega^1, \ldots, \omega^S\} = \{\omega | \omega_t \in \mathcal{S}_t(\omega^{t-1,\bullet}) \forall t > 1\}. \tag{12}$$

The associated conditional probabilities $P(\omega_t|\omega^{t-1,\bullet})$ on $\mathcal{S}_t(\omega^{t-1,\bullet})$ for $t > 1$ and the marginal probabilities $P(\omega_1)$ on $\mathcal{S}_1$ are called the *arc probabilities*. Their products

$$P(\omega^{t-1,\bullet}) = P(\omega_1) \prod_{\tau=2}^{t-1} P(\omega_\tau|\omega^{\tau-1,\bullet}) \tag{13}$$

are the *path probabilities* and the probability $p_s$ of scenario $\omega^s = \{\omega_1^s, \ldots, \omega_T^s\} \in \mathcal{S}^T$ is

$$p_s = P(\omega^s) = P(\omega_1^s) \prod_{\tau=2}^{T} P(\omega_\tau^s | \omega_1^s, \ldots, \omega_{\tau-1}^s).$$

The decisions at the beginning of the $t$-th period (at the stage $t = 1, \ldots, T$) depend on the sequence of observed realizations of the random variables in the preceding periods, e.g., on the section $(\omega_1^s, \ldots, \omega_{t-1}^s)$ in case of the $s$-th scenario. The first-stage decision variables are scenario independent and the last stage decisions at the beginning of the $T$-th period depend on the sections $\omega^{T-1,\bullet}$. These nonanticipativity conditions can be included into the problem formulation in various ways.

A specific problem related with multistage stochastic programs is the required special structure of the input in a form consistent with (9). We can view it as an oriented graph which starts from a root (the only node at level 0) and branches into nodes at level 1, each corresponding to one of the possible realizations of $\omega_1$, and the branching continues up to nodes at level $T$ assigned to the whole possible data paths $\omega^{T,\bullet}$. A common special arrangement is the *scenario tree* which is based on the additional assumption that there is a one-to-one correspondence between the sections $\omega^{t,\bullet}$ and the nodes of the tree at stage $t$ for $t = 1, \ldots, T$. This means that for any node at level $t$, each of the new observations $\omega_t$ must have only one immediate predecessor $\omega^{t-1,\bullet}$, i.e., a node at level $t-1$, and a (finite) number of descendants $\omega_{t+1}$ that result in nodes at level $t+1, t < T$.

An example is the scenario-based form of the $T$-stage stochastic *linear* program with recourse written in the arborescent form (compare with (9)):

$$\text{maximize } \boldsymbol{c}_1^\top \boldsymbol{x}_1 + \sum_{k_2=2}^{K_2} p_{k_2} \boldsymbol{c}_{k_2}^\top \boldsymbol{x}_{k_2} + \sum_{k_3=K_2+1}^{K_3} p_{k_3} \boldsymbol{c}_{k_3}^\top \boldsymbol{x}_{k_3} + \cdots + \sum_{k_T=K_{T-1}+1}^{K_T} p_{k_T} \boldsymbol{c}_{k_T}^\top \boldsymbol{x}_{k_T} \tag{14}$$

subject to constraints

$$
\begin{array}{llll}
A_1 x_1 & & = b_1, & \\
B_{k_2} x_1 + A_{k_2} x_{k_2} & & = b_{k_2}, & \\
& & & k_2 = 2, \ldots, K_2, \\
& B_{k_3} x_{a(k_3)} + A_{k_3} x_{k_3} & = b_{k_3}, & \\
& & & k_3 = K_2 + 1, \ldots, K_3, \\
& \ddots \qquad \ddots & \vdots & \\
& B_{k_T} x_{a(k_T)} + A_{k_T} x_{k_T} & = b_{k_T}, & \\
& & & k_T = K_{T-1} + 1, \ldots, K_T,
\end{array}
$$

$$
l_t \le x_{k_t} \le u_t, \qquad k_t = K_{t-1} + 1, \ldots, K_t, \quad t = 1, \ldots, T, \tag{15}
$$

with $K_1 = 1$. We denote here by $a(k_t)$ the immediate ancestor of $k_t$, so that $a(k_2) = 1, k_2 = 2, \ldots, K_2$, and $p_{k_t} > 0$ is the path probability assigned to the node $k_t$; compare with (13). Notice that the nonanticipativity constraints are spelled out implicitly.

Two special cases of the scenario tree are worth mentioning as they offer a relatively easy generalization of the existing theoretical and computational results valid for two-stage stochastic programs.

- For all stages $t = 2, \ldots, T$, the conditional probabilities $P(\omega_t|\omega^{t-1,\bullet})$ are independent of $\omega^{t-1,\bullet}$ and are equal to the marginal probabilities $P(\omega_t)$ – the *interstage independence*.
- For all stages $t = 2, \ldots, T$, the supports $\mathcal{S}_t(\omega^{t-1,\bullet})$ of conditional probability distributions of $\omega_t$ conditioned by realizations $\omega^{t-1,\bullet} = \{\omega_1, \ldots, \omega_{t-1}\}$ of sections $\omega^{t-1,\bullet}$ are singletons. This means that the scenario tree is nothing else but a "fan" of individual scenarios $\omega^s = \{\omega_1^s, \ldots, \omega_T^s\}$ which occur with probabilities $p_s = P(\omega_1^s)\forall s$ and, independently of the number of periods, the problem reduces to a *two-stage* multiperiod stochastic program.

Except for the two special cases mentioned above, to build a representative scenario tree seems to be presently the crucial problem for applications; we shall discuss it very briefly with the reference to [26] and [62].

In finance, there exist advanced continuous and discrete time stochastic models and historical time series that serve to calibrate these models. The examples are diffusion type models and multivatiate autoregression models used in [2, 12, 17, 31, 36, 54] or the Black-Derman-Toy model [4] for generating a binomial lattice for interest rates. These models employ a specified type of probability distributions, mainly the transformed (multi)normal ones. They have to be calibrated from the existing market data which involves suitable estimation and test procedures. Using the calibrated model, or its time discretization, one can simulate or select a large number of sample paths of $\omega$. Successful examples are the global scenario system developed by Towers Perrin [54] or the model FAM [16] for simulation of economic and capital market scenarios.

Nonparametric methods for scenario generation can be applied for very large families of probability distributions which cannot be indexed by a finite dimensional parameter in a natural way; another term used in this connection is *distribution-free* methods. The simplest idea is to use as scenarios the past observations obtained under comparable circumstances and assign them equal probabilities; see for instance [34, 56].

In case of truly multistage stochastic programs, these independently generated data trajectories are then used to build the scenario tree; the important initial decision is the number of stages (and the time discretization in general) and, depending on the algorithm, also the branching scheme. Possible updates of the structure of the scenario tree can be incorporated into the solution procedure, e.g. [12], or tested within the postoptimality analysis; see Section 4.

## 3 Scenario based stochastic programs in portfolio optimization

We shall assume now that the uncertainty is described by a discrete probability distribution of random parameters carried by a finite number of scenarios with prescribed probabilities and that this discrete probability distribution is a reliable substitute of the true underlying probability distribution. To simplify the notation, we shall denote the coefficients and decision variables related with scenario $\omega^s$ simply by a superscript $s$.

A fundamental investment decision is the selection of asset *categories* and the wealth allocation over time. The level of aggregation depends on investor's circumstances. The planning horizon at which the outcome is evaluated is the endpoint $T_0$ of an interval $[0, T_0]$ which is further discretized, covered by nonoverlapping time intervals indexed by $t = 1, \dots, \tau$. An initial portfolio is constructed at time 0, i.e., at the beginning of the first period, and is subsequently rebalanced at the beginning of subsequent periods, i.e., for $t = 2, \dots, \tau$, to cover the target ratio or to contribute to maximization of the final performance at $T_0$. In our general setting of the $T$-stage stochastic programs, see (6)–(7), $\tau = T$. In some cases, additional time instants can be included at which some of economic variables are calculated; after $T_0$, no further active decisions are allowed.

It is important to realize that the stages do not necessarily correspond to time periods. The main interest lies in the first-stage decisions which consist of all decisions that have to be selected before the information is revealed, just on the basis of the already known probability distribution $P$, i.e., on the basis of the already designed scenario tree. The second-stage decisions are allowed to adapt to an additional information available at the end of the first-stage period, etc. For the sake of simplicity and accepting degenerated supports $\mathcal{S}(\omega^{t-1,\bullet})$ of the conditional probability distributions $P(\omega_t|\omega^{t-1,\bullet})$, we will not distinguish strictly between indices of stages and time periods in what follows.

The primary decision variable $h_i^s(t)$ represents the holding in asset category $i$ at the beginning of time period $t$ under scenario $s$ *after the rebalancing decisions took place*; the initial holding is $h_i(0)$. It can be included into the model as the amount of money invested in $i$ at the beginning of time period $t$, can be expressed in dollars of the initial purchase price, in face values, in number of securities or in lots, etc. Accordingly, the value of the holdings at the end of the period $t$ may be affected by the market returns; the *wealth accumulated* at the end of the $t$-th period before the next rebalancing takes place is then

$$w_i^s(t) := (1 + r_i^s(t))h_i^s(t) \forall i, t, s.$$

Purchases and sales of assets are represented by variables $b_i^s(t), s_i^s(t)$ with transaction costs defined via time-independent coefficients $\alpha_i$ and assuming mostly the symmetry in the transaction costs; it means that purchasing one unit of $i$ at the beginning of period $t$ requires $1 + \alpha_i$ units of cash and selling one unit of $i$ results in $1 - \alpha_i$ units of cash. The *flow balance constraint* for each asset category (except for cash, the asset indexed by $i = 0$), scenario and time period is

$$h_i^s(t) = (1 + r_i^s(t-1))h_i^s(t-1) + b_i^s(t) - s_i^s(t). \tag{16}$$

It restricts the cashflows at each period to be consistent.

The *flow balance equation for cash* for each time period and all scenarios is for instance

$$\begin{aligned} h_0^s(t) = h_0^s(t-1)(1 + r_0^s(t-1)) + c^s(t) + \sum_i s_i^s(t)(1-\alpha_i) - \sum_i b_i^s(t)(1+\alpha_i) \\ + \sum_i f_i^s(t)h_i^s(t) - y^{s-}(t-1)(1 + \delta^s(t-1)) - L^s(t) + y^{s-}(t) \end{aligned} \tag{17}$$

with $f_i^s(t)$ the cashflow generated by holding one unit of the asset $i$ during the period $t$ (coupons, dividends, etc.) under scenario $s$ and $L^s(t)$ the paydown of the commited liabilities in period $t$ and under scenario $s$. We denote $y^{s-}(t)$ borrowing in period $t$ under scenario $s$ at the borrowing rate $\delta^s(t)$ and $c^s(t) = c^{s+}(t) - c^{s-}(t)$ decision variables concerning the structure of external cash-flows in period $t$ under scenario $s$. (For simplicity we assume here that all borrowing is done on a single period basis; a generalization is possible.)

For holdings, purchases and sales expressed in numbers or in face values, the cash balance equation contains purchasing and selling prices, $\xi_i^s(t) > \zeta_i^s(t)$:

$$\begin{aligned} h_0^s(t) = h_0^s(t-1)(1 + r_0^s(t-1)) + c^s(t) + \sum_i \zeta_i^s(t)s_i^s(t) - \sum_i \xi_i^s(t)b_i^s(t) \\ + \sum_i f_i^s(t)h_i^s(t) - y^{s-}(t-1)(1 + \delta^s(t-1)) - L^s(t) + y^{s-}(t) \end{aligned} \tag{18}$$

and the flow balance constraints for assets assume a simpler form

$$h_i^s(t) = h_i^s(t-1) + b_i^s(t) - s_i^s(t) \tag{19}$$

as no wealth accumulation is considered.

The decision variables $h_i^s(t), b_i^s(t), s_i^s(t), y^{s-}(t)$ are nonnegative and it is easy to include further constraints which force a diversification, limit investments in risky or illiquid asset classes, limit borrowings, loan principal payments and turnovers, reflect legal and institutional constraints, etc. It is also possible to force a specific decision policy, e.g., the fixed-mix policy which can be expressed as

$$h_i^s(t) = \lambda_i \sum_{j=0}^{I} h_j^s(t) \forall i$$

where $\lambda_i$ is a fixed ratio of asset $i$ in the portfolio, cf. [53, 58].

Whereas the random liabilities $L^s(t)$ belong to the model input, various *decisions* concerning other liabilities can be included in the external cashflows and, similarly as in [45], one can separate decisions on accepting various types of deposits, on emission further debt instruments, decisions on specific goal payments, on long term debt retirement [15], etc. Naturally, the cash balance equation has to take into account the cost of the debt service.

The objective function is mostly related to the wealth at the end of the planning horizon $T_0$; this for each scenario consists of the amount of the total wealth $\sum_{i=0}^{I} w_i^s(T_0)$ reduced for the present value of liabilities and loans outstanding at the horizon. Risk can be reflected by the choice of a suitable utility function or incorporated into constraints. To include short term goals, cummulated penalties for shortfalls under scenario $s$ (e.g., the amounts $y^{s-}(t) > 0$) are subtracted from the final wealth computed for the same scenario [9]. A perspective alternative is to examine utility functions of *several* outcomes at specific time instants covered by the model. Also criteria *nonlinear* in the probability distributions can be applied; indeed, the Markowitz model (1) can be considered as an example, see also [61].

To initiate the model, one uses scenarios $r_i^s(t), \delta^s(t), f_i^s(t), L^s(t)$ of the returns, interest rates and liabilities for all $t$ and starts with the known, scenario independent initial holdings $h_i^s(0) \equiv h_i(0)$ of cash and all considered assets and with $y^{s-}(0) \equiv 0 \forall s$. If no ties in scenarios are considered they can be visualized as a *fan* of individual scenarios which start from the common known values $r_i(0), \delta(0), f_i(0), L(0)$ valid for $t = 0$. All decisions $h_i^s(t)$, $b_i^s(t)$, $s_i^s(t)$, $y^{s-}(t)$, $c^s(t) \forall i, s$ and $t \geq 1$ can be computed at once. In this case, only one additional requirement must be met: the initial decision $h_i^s(1)$, $b_i^s(1)$, $s_i^s(1)$, $y^{s-}(1)$, $c^s(1)$ must be *scenario independent*. This is a simple form of the nonanticipativity constraints and the resulting problem is a *multiperiod two-stage stochastic program*. An example is [24, 33, 45, 59].

For *multistage stochastic programs*, the input is mostly in the form of a scenario tree and the nonanticipativity constraints on decisions enter implicitly by using a decision tree which follows the structure of the already given scenario tree, e.g., (14)–(15), or in an explicit way by forcing the decisions based on the same history (i.e., on an identical part $\omega^{t,\bullet}$ of several scenarios) to be equal, as it was in the case of the two-stage stochastic program. With the explicit inclusion of the nonanticipativity constraints, the scenario-based multiperiod and multistage stochastic programs with linear constraints can be written in a form of a large-scale deterministic program

$$\max_{\mathscr{X} \cap \mathscr{N}} \left\{ \sum_s p_s u^s(\boldsymbol{x}^s) \mid \boldsymbol{A}^s \boldsymbol{x}^s = \boldsymbol{b}^s, s = 1, \ldots, S \right\}$$

where $\mathscr{X}$ is a set described by simple constraints, e.g., by nonnegativity conditions, $\mathscr{N}$ is defined by the nonanticipativity constraints and $u^s$ is the performance measure in case of scenario $s$.

A large class of solvers (CPLEX, MSLiP-OSL, OSL-SP, etc.) are currently available for the solution of multistage problems with linear constraints and nonlinear objectives. Nonlinear or integer constraints can be included but for the cost of an increased numerical complexity. On the other hand, if the resulting problem can be transformed into a large *linear* program, there are at disposal special decision support systems which are able to manage efficiently large scale scenario based stochastic linear programs for portfolio optimization including those with piece-wise linear concave utility functions.

## 4 How to draw inference about the true problem?

Already the early works on applications of stochastic programming in finance were aware of the fact that the obtained solution or policy can be influenced both by the choice and approximation of the presumably known probability distribution $P$. Up to now, the main tool has been sensitivity analysis via repeated runs of the optimization problem with a changed input; see for instance [7, 11, 33, 45, 59, 77]. We shall focus here on other approaches suitable for drawing conclusions about the optimal solutions and the optimal value of the true stochastic program when using results obtained by solving an approximate scenario-based program. Such possibilities depend essentially on the structure of the solved problem as well as on the origin of scenarios.

The main sources of errors come from simulation, sampling, estimation and also from incomplete or unprecise input information, the main tools are selected methods of probability theory, asymptotic and robust statistics, simulation methods and parametric optimization. To simplify the exposition, we shall concentrate on stochastic programs written in the form

$$\max_{\boldsymbol{x} \in \mathscr{X}} F(\boldsymbol{x}, P) \tag{20}$$

where the objective function is mostly an expectation of a rather complicated function $f_0(\boldsymbol{x}, \omega)$ and $\mathscr{X}$ is a nonempty closed set which does not depend on $P$.

To get (20), it is enough to separate the first-stage decisions $\boldsymbol{x}_1$ and $\omega_1$ in the nested form (10)–(11) and to consider the first-stage probem (10). Notice, however, that the function $\varphi_{10}(\boldsymbol{x}_1, \omega_1)$ in (10) depends also on the conditional probability distribution $P(\omega|\omega_1)$. For scenario-based stochastic *linear* programs, e.g., (14)–(15), the functions $\varphi_{10}(\boldsymbol{x}_1, \omega_1)$ can be computed as optimal values of $(T-1)$-stage stochastic linear programs written in the arborescent form, each corresponding to one considered realization of $\omega_1$. Dependence on the conditional discrete distributions becomes obvious. The large deterministic linear program (14)–(15), and also those for evaluation of $\varphi_{10}(\boldsymbol{x}_1, \omega_1)$, can be rearranged so that it assumes the form of a large *two-stage* stochastic linear program with *random* recourse; the fixed recourse is obtained only when the random, scenario dependent, coefficients are on the right-hand sides and in the objective function. This observation suggests a way how to extend the results known for two-stage programs to the multistage case, but, at the same time, it points out at the problems to be expected in postoptimality, stability and worst-case analysis for multistage stochastic programs. On the other hand, for a *discrete* probability distribution $P$, the first-stage problem

$$\max_{\boldsymbol{x}_1 \in \mathscr{X}_1} \boldsymbol{c}_1^\top \boldsymbol{x}_1 + E_{P(\omega_1)}\{\varphi_{10}(\boldsymbol{x}_1, \omega_1)\} \tag{21}$$

is a *convex* program with a concave piece-wise linear objective function. This property is also valid whenever all objective functions $f_{t0}$ in (11) are concave piece-wise linear.

*Example 2 – A multiperiod bond portfolio management problem.* Consider $S$ scenarios of interest rates $\boldsymbol{r}^s \in R^T, s = 1, \ldots, S$ which occur with probabilities $p_s > 0 \forall s, \sum_s p_s = 1$. These scenarios enter coefficients of a multiperiod two-stage problem for bond portfolio management, see [24, 33], with scenario independent liabilities. The problem can be rewritten into the form (20) as

$$\max_{\boldsymbol{x} \in \mathscr{X}} \sum_{s=1}^{S} p_s W(\boldsymbol{x}, \boldsymbol{r}^s); \tag{22}$$

the first-stage decision vector $\boldsymbol{x}$ consists of components of the scenario independent first-stage decision variables $h_i(1), b_i(1), s_i(1) \forall i$ in the notation introduced in Section 3, $W(\boldsymbol{x}, \boldsymbol{r}^s)$ denotes the *maximal* market value of the portfolio which can be attained at time $T_0$ for a given scenario $\boldsymbol{r}^s$ and an already fixed first-stage decision $\boldsymbol{x}$ and the set $\mathscr{X}$ is described by deterministic linear constraints on holdings and cashflow (compare with (18)–(19))

$$h_i(1) = h_i(0) + b_i(0) - s_i(0), i = 1, \ldots, I, \tag{23}$$

$$\begin{aligned} h_0(1) = h_0(0)(1 + r_0(0)) &+ \sum_{i=1}^{I} \zeta_i(1) s_i(1) \\ &- \sum_{i=1}^{I} \xi_i(1) b_i(1) + \sum_{i=1}^{I} f_i(1) h_i(1) - L(1) + y^-(1) \end{aligned} \tag{24}$$

and by nonnegativity of all variables. The set $\mathscr{X}$ of feasible first-stage decisions is nonempty, convex and compact provided that the initial wealth is positive and borrowing in the first stage is restricted.

Thanks to unlimited possibilities of borrowing in the subsequent periods (see (18)), the second-stage linear programs which determine $W(\boldsymbol{x}, \boldsymbol{r}^s)$ are feasible for each $\boldsymbol{x}, \boldsymbol{r}^s$ and the first-stage decision variables enter only the right-hand sides of the system of their (linear) constraints. With prices $\xi_i^s(t) > \zeta_i^s(t)$ and with $\delta^s(t) > 0 \forall i, t, s$, the dual programs are also feasible, hence, for each $\boldsymbol{r}^s$, $W(\bullet, \boldsymbol{r}^s)$ is a finite, piece-wise linear concave function on $\mathscr{X}$ and these properties extend evidently also to the objective function in (22).

In general, we assume that the set $\mathscr{X}$ is nonempty, convex and closed and the objective function in (20) is concave in $\boldsymbol{x}$ and convex in $P$; the last assumption covers also the mean-variance model or models of robust optimization [55]. The probability distribution function plays a role of an abstract parameter which is estimated or approximated by another probability distribution $\hat{P}$ obtained by parametric or nonparametric methods and by sampling, discretization and simulation techniques. The optimization problem is solved with $\hat{P}$ and the question is how the obtained optimal value $\varphi(\hat{P})$ and the set of optimal solutions $\Psi(\hat{P})$ relate to the optimal value $\varphi(P)$ and to the set of optimal solutions $\Psi(P)$ for the true problem. Depending on the nature of the approximation, the problem can be for instance treated within asymptotic statistics or in the frame of quantitative stability for parametric programs. In general, it is much easier to get an estimate of precision of the optimal value than that of the optimal solutions.

For the *expected value objective* $F(\boldsymbol{x}, P) := E_{P(\omega)} f_0(\boldsymbol{x}, \omega)$, a natural idea is to use the empirical probability distribution $P^\nu$ based on $\nu$ i.i.d. sample values of $\omega$ and to study the asymptotic properties of the optimal value $\varphi(P^\nu)$ and of the set of optimal solutions $\Psi(P^\nu)$ of the sample problem

$$\max_{\boldsymbol{x} \in \mathcal{X}} F(\boldsymbol{x}, P^\nu) = \max_{\boldsymbol{x} \in \mathcal{X}} \frac{1}{\nu} \sum_{j=1}^{\nu} f_0(\boldsymbol{x}, \omega^j).$$

The notion of the *weak convergence* [69] is a legitimate starting point. It implies convergence of the expectations $F(\boldsymbol{x}, P^\nu) = \int_\Omega f_0(\boldsymbol{x}, \omega) P^\nu(d\omega)$ for a fixed $\boldsymbol{x}$ if $f_0(\boldsymbol{x}, \bullet)$ is a bounded continuous function of $\omega$ or if the set of probability distributions is restricted to a subset with respect to which the function $f_0(\boldsymbol{x}, \bullet)$ is uniformly integrable.

*Classical consistency results.* Under assumption that $P_\nu \to P$ weakly and that $f_0(\boldsymbol{x}, \omega)$ is a continuous bounded function of $\omega$ for every $\boldsymbol{x} \in \mathcal{X}$, the point-wise convergence of the expected value objectives $F(\boldsymbol{x}, P^\nu) \to F(\boldsymbol{x}, P)\ \forall \boldsymbol{x} \in \mathcal{X}$ follows directly from the definition of weak convergence. If $\mathcal{X}$ is compact and the convergence is uniform on $\mathcal{X}$ we get immediately convergence of the optimal values

$$\varphi(P^\nu) \to \varphi(P).$$

If, moreover, $\mathcal{X}$ is convex and $f_0(\bullet, \omega)$ is strictly convex on $\mathcal{X}$ it is easy to get in addition the convergence of the (unique) optimal solutions $\boldsymbol{x}(P^\nu)$ of $\max_{\boldsymbol{x} \in \mathcal{X}} F(\boldsymbol{x}, P^\nu)$ to the unique optimal solution $\boldsymbol{x}(P)$ of the initial problem (20).

*Bias of the empirical optimal value.* It is possible to prove that for empirical probability distributions, $\varphi(P^\nu)$ has a *one-directional bias* in the sense that

$$E\varphi(P^\nu) \geq \varphi(P).$$

Indeed, for i.i.d. values $\omega^i$ and for any fixed $\boldsymbol{x} \in \mathcal{X}$, the function values $f_0(\boldsymbol{x}, \omega^j)$ are i.i.d. and

$$\begin{aligned} E\varphi(P^\nu) &= E \max_{\boldsymbol{x}} \frac{1}{\nu} \sum_{j=1}^{\nu} f_0(\boldsymbol{x}, \omega^j) \\ &\geq \max_{\boldsymbol{x}} \frac{1}{\nu} E \sum_{i=1}^{\nu} f_0(\boldsymbol{x}, \omega^j) = \max_{\boldsymbol{x}} E f_0(\boldsymbol{x}, \omega) = \varphi(P). \end{aligned}$$

An empirical point estimate of $E\varphi(P^\nu)$ follows from the Law of Large Numbers and an asymptotic confidence interval for this upper bound on the true optimal value $\varphi(P)$ can be obtained from the Central Limit Theorem. This idea was elaborated in [48] and applied in [3] to a bond portfolio management problem akin to that of Example 2.

To get consistency results under less stringent assumptions concerning the true and approximating probability measures, functions $f_0$ and for multiple optimal solutions requires a different methodology; see Theorem 3.9 of [29] which will be applied below.

*General consistency results for a discrete true distribution P.* Let $\omega^1, \ldots, \omega^N$ be the atoms of $P$ and $\pi_j, j = 1, \ldots, N$, their probabilities, let $\mathscr{X}$ be a nonempty convex polyhedron and $f_0(\boldsymbol{x}, \omega)$ a piece-wise linear concave function of $\boldsymbol{x}$. This means that $F(\boldsymbol{x}, P) := \sum_{j=1}^N \pi_j f_0(\boldsymbol{x}, \omega^j)$ is also a piece-wise linear concave function, hence, there exists a finite number of nonoverlapping nonempty convex polyhedra $\mathscr{X}^k, k = 1, \ldots, K$, such that $\mathscr{X} = \bigcup_{k=1}^K \mathscr{X}^k$ and $F(\boldsymbol{x}, P)$ is linear on each of $\mathscr{X}^k$. Then the set of optimal solutions $\Psi(P)$ evidently intersects the set $\bar{\mathscr{X}}(P)$ of all extremal points of $\mathscr{X}^k, k = 1, \ldots, K$. The true distribution $P$ is estimated by empirical distributions $P^\nu$ based on finite samples of sizes $\nu$ from $P$, hence, carried by subsets of $\{\omega^1, \ldots, \omega^N\}$. The empirical objective functions $F(\boldsymbol{x}, P^\nu)$ are also concave, piece-wise linear and the sets of the related extremal points $\bar{\mathscr{X}}(P^\nu) \subset \bar{\mathscr{X}}(P)$. This means that the assumptions of Theorem 3.9 in [29] are fulfilled with the compact set $\mathscr{D} = \bar{\mathscr{X}}(P)$. Consequently, with probability one, any cluster point of any sequence of points $x(P^\nu) \in \Psi(P^\nu) \cap \mathscr{D}$ is an optimal solution of the true problem.

Assume in addition that there is a unique optimal solution $\boldsymbol{x}(P)$ of the true problem (20). In this case there is a measurable selection $\boldsymbol{x}^*(P^\nu)$ from $\Psi(P^\nu) \cap \mathscr{D}$ such that with probability 1, $\lim_{\nu\to\infty} \boldsymbol{x}^*(P^\nu) = \boldsymbol{x}(P)$. Due to the special form of the objective functions and of the sets $\Psi(P^\nu) \cap \mathscr{D}$, this is equivalent to

$$\boldsymbol{x}^*(P^\nu) \equiv \boldsymbol{x}(P) \quad \text{a.s. for } \nu \text{ large enough.} \tag{25}$$

*Example 2 – continuation.* Assume that the scenarios of interest rates in the bond portfolio management problem have been generated according to the Black-Derman-Toy [4] model. This provides a large number, say $N$, of equiprobable scenarios. As the number of scenarios can be very large, e.g., $2^{360}$ if bonds with maturity 360 months are considered, one solves a problem based on a relatively small sample of scenarios. For i.i.d. sample of interest rate vectors form the full binomial lattice and for incresing sample sizes, the above consistency result applies: There is an integer $\nu_0$ such that for increasing sample sizes $\nu > \nu_0$ a (sub)sequence of the respective first-stage optimal solutions $\boldsymbol{x}(P^\nu)$ becomes stationary, $\boldsymbol{x}(P^\nu) \equiv \boldsymbol{x}(P)$ a.s. with $\boldsymbol{x}(P) \in \Psi(P)$, an optimal solution of the full problem. This is an encouraging result; however, an important question is the rate of the convergence or the related problem of the required minimal sample size $\nu_0$.

*Asymptotic normality* of the optimal value $\varphi(P^\nu)$ can be proved under relatively weak assumptions, e.g., for a compact set $\mathscr{X} \neq \emptyset$, unique true optimal solution $\boldsymbol{x}(P)$ and continuous $f_0(\bullet, \omega)\forall\omega$ with finite expectation $E\{f_0(\boldsymbol{x}(P), \omega)\}^2$, see Theorem 3.3 of [70]. This allows to construct approximate confidence intervals for the true optimal value. On the other hand, a similar result for optimal solutions $\boldsymbol{x}(P^\nu)$ cannot be expected even when all solution sets $\Psi(P), \Psi(P^\nu)\forall\nu$ are singletons. Moreover, for typical stochastic programs, the higher order differentiability assumptions concerning $f_0(\boldsymbol{x}(P), \omega)$ or $F(\boldsymbol{x}, P)$ needed for the second order analysis of the optimal value function are not fulfilled. To an extent, smoothness can be substituted by convexity properties. An example is the asymptotic distribution result in [41] applied to the static portfolio optimization problem based on a piece-wise linear-quadratic convex tracking function.

Theoretically, the above asymptotic results apply also to the multistage stochastic programs. However, the assumption of an infinitely increasing sample size means that at every node of the scenario tree, the number of branches grows to infinity and the sample based problems become very quickly intractable.

*Asymptotic results for a parametric family.* A simplification is possible whenever the general stability properties with respect to the probability distribution can be reduced to a finite dimensional parameter case. An example are probability distributions of a given parametric form and the desired results concern differences between the optimal values $\varphi(\boldsymbol{\theta}_0)$ and $\varphi(\boldsymbol{\theta}_\nu)$ and between the solution sets obtained for the true parameter value $\boldsymbol{\theta}_0$ and for its estimate $\boldsymbol{\theta}_\nu$, respectively. For sufficiently smooth optimal value function $\varphi$ and for unique optimal solutions, the statistical properties of $\varphi(\boldsymbol{\theta}_\nu)$ and of $\boldsymbol{x}(\boldsymbol{\theta}_\nu)$ follow from the statistical properties of the estimates $\boldsymbol{\theta}_\nu$ by application of results concerning transformed random sequences [69].

*Example 1 – continuation.* Simulation experiments [11] indicate that the results of the Markowitz model are rather sensitive on the assumed values of expected returns. Let us assume that the covariance matrix $\boldsymbol{V}$ in (1) is a known positive definite matrix, the set $\mathscr{X}$ a nonempty convex polyhedron with nondegenerated vertices, $\lambda > 0$ a chosen parameter value and that the true expected return $\boldsymbol{r}_0$ was estimated by an asymptotically normal estimate $\boldsymbol{r}_\nu$ based on a sample size $\nu$: $\sqrt{\nu}(\boldsymbol{r}_\nu - \boldsymbol{r}_0) \sim \mathscr{N}(\boldsymbol{0}, \boldsymbol{\Sigma})$. Under these assumptions about (1), for each $\boldsymbol{r}$, there is a unique optimal solution $\boldsymbol{x}(\boldsymbol{r})$ and the optimal value function $\varphi(\boldsymbol{r}) := \max_{\boldsymbol{x}\in\mathscr{X}}[\lambda\boldsymbol{r}^\top\boldsymbol{x} - \frac{1}{2}\boldsymbol{x}^\top\boldsymbol{V}\boldsymbol{x}]$ is a piece-wise linear-quadratic differentiable convex function with the gradient $\lambda\boldsymbol{x}(\boldsymbol{r})$. Hence, asymptotic normality of the sample optimal value function follows:

$$\sqrt{\nu}(\varphi(\boldsymbol{r}_\nu) - \varphi(\boldsymbol{r}_0)) \sim \mathscr{N}(\boldsymbol{0}, \nabla\varphi(\boldsymbol{r})^\top\boldsymbol{\Sigma}\varphi(\boldsymbol{r})). \tag{26}$$

The asymptotic result in (26) holds true also with the variance replaced by its sample counterpart, i.e., by $\lambda^2\boldsymbol{x}(\boldsymbol{r}_\nu)^\top\boldsymbol{\Sigma}_\nu\boldsymbol{x}(\boldsymbol{r}_\nu)$ and asymptotic confidence intervals for the optimal value can be constructed.

The optimal solutions $\boldsymbol{x}(\boldsymbol{r})$ of (1) are directionally differentiable, piece-wise linear on certain nonoverlapping convex stability sets which means that the asymptotic normality can hold true only if the true expected return $\boldsymbol{r}_0$ lies in the interior of one of these stability sets. Continuity of $\boldsymbol{x}(\boldsymbol{r})$ is sufficient for consistency of the optimal solutions based on a consistent estimate of the true expected return; we refer to [22] for details.

For various reasons, empirical estimates of the probability distribution $P$ or of the true parameter $\boldsymbol{\theta}_0$ are not always available and, moreover, they need not provide the best approximation technique: They focus solely on the probability distribution, which is not the only ingredient of the stochastic portfolio optimization models, they do not take into account any expert knowledge or foresight and for technical reasons, they cannot be based on very large samples. Moreover, the goal is to get a sensible approximation of the optimal solution and of the optimal value, not an approximation of the probability distribution. Stability results with respect to *probabilities* of the selected scenarios can be found in [65]. We shall look into stability analysis of (20) with respect to the abstract parameter $P$.

*Quantitative stability results.* The success and applicability of the quantitative stability results depend essentially on an appropriate choice of the *probability metric* $d$ used to measure the perturbations in the probability distribution $P$. The probability metrics should be closely tailored to the structure of the considered stochastic program and/or to the particular type of approximation of probability distribution $P$. The desired results are, e.g., a Lipschitz (or Hölder) property of the optimal value

$$d(P, P') < \eta \quad \Rightarrow \quad |\varphi(P) - \varphi(P')| < K\eta$$

and possibly also a Lipschitz (or Hölder) property of the Hausdorff distance of the corresponding solution sets with respect to perturbations of $P$ measured by $d$; naturally, the Lipschitz (or Hölder) constants depend on the chosen metric $d$. Again, special assumptions are needed for to get such results for optimal solutions whereas for the *sets of ε-optimal solutions*

$$\Psi_\varepsilon(P) = \varepsilon - \arg\min_{\boldsymbol{x} \in \mathscr{X}} F(\boldsymbol{x}, P) = \{\boldsymbol{x} \in \mathscr{X} \,|\, F(\boldsymbol{x}, P) \le \varphi(P) + \varepsilon\}$$

quantitative stability results hold true under more general circumstances [1].

An important class of probability metrics in our context, is based on the Kantorovich-Rubinstein functional with a continuous distance function $c : R^s \times R^s \to R^1_+$, see Chapter 5 in [63]:

$$d_c(P, Q) = \inf\left\{\int_{R^s \times R^s} c(\boldsymbol{u}, \boldsymbol{w})\eta(d\boldsymbol{u}, d\boldsymbol{w})\right\} \tag{27}$$

over all finite Borel measures $\eta$ on $\Omega \times \Omega$ such that $\eta(B \times \Omega) - \eta(\Omega \times B) = P(B) - Q(B) \forall B \in \mathscr{F}$. An example is the *Fortet-Mourier metric* obtained for

$$c(\boldsymbol{u}, \boldsymbol{w}) = \|\boldsymbol{u} - \boldsymbol{w}\| \max\{1, \|\boldsymbol{w}\|^{p-1}, \|\boldsymbol{u}\|^{p-1}\}. \tag{28}$$

The distance function $c$ has to be chosen so that the integrands $f_0(\boldsymbol{x}, \bullet)$ exhibit a generalized Lipschitz property with respect to $\omega$:

$$|f_0(\boldsymbol{x}, \omega) - f_0(\boldsymbol{x}, \tilde{\omega})| \le c(\omega, \tilde{\omega}) \forall \omega, \tilde{\omega} \in \Omega. \tag{29}$$

Then the general results of parametric programming imply that small changes in the probability distribution measured by $d_c$ result in small changes of the optimal values and of the sets of $\varepsilon$-optimal solutions, cf. [27] for theoretical results and their application to the bond portfolio management problem. These quantitative stability results can help in designing a discrete approximation $P'$ of $P$ which is representative enough and such that the obtained solution enjoys plausible robustness properties. For *scenario-based* programs they can be used to quantify the desirable robustness properties also in rather complicated instances of stochastic programs with random recourse. Moreover, for two discrete probability distributions, say $P = \sum p_i \delta_{\omega^i}, Q = \sum_j q_j \delta_{\tilde{\omega}^j}$, $d_c(P, Q)$ is the optimal value of a finite-dimensional *transportation problem.*

*An application – Deleting scenarios.* Assume again that the discrete probability distribution $P$ is carried by a finite number of *scenarios* $\omega^1, \ldots, \omega^N$, which occur with positive probabilities $p_1, \ldots, p_N, \sum_i p_i = 1$; the number $N$ can be very large. Such distribution can be the true underlying discrete distribution, an example is the distribution of interest rates obtained according to the Black-Derman-Toy model [4], or an already accepted good discrete approximation of a true distribution that can be obtained by sampling, discretization techniques, etc.

A possible reduction of the number of atoms of the discrete probability distribution $P$ can be based on the distance $d_c$ of probability distributions defined in (27) with a distance function $c$ for which (29) is satisfied and the triangle inequality holds true. We focus on deleting scenarios so that the distance of the reduced probability distribution from the initial one is small.

The resulting *optimal deletion rule* [28] reads: Delete the scenario $\omega^k$ for which

$$p_k \min_{i \neq k} c(\omega^i, \omega^k) = \min_{l=1,\ldots,N} p_l \min_{i \neq l} c(\omega^i, \omega^l).$$

The optimal reduced discrete probability distribution is $Q = \sum_{i=1, i \neq k}^{N} q_i \delta_{\omega^i}$ with

$$q_i = p_i, i \neq i_k, q_{i_k} = p_{i_k} + p_k \quad \text{for } c(\omega^{i_k}, \omega^k) = \min_{i \neq k} c(\omega^i, \omega^k).$$

The result is rather canonical because it says that deletion should be done where the scenarios are close together in the sense of the distance $c$ or where probabilities are small. The results and the rule can be extended to deletion of a fixed set of scenarios and to prescribed redistribution rules. Such redistribution rule follows for instance from the requirement that the remaining scenarios are equiprobable again.

*Example 2 – continuation.* For the bond portfolio management problem, $f_0(\boldsymbol{x}, \omega) = W(\boldsymbol{x}, \boldsymbol{r})$ – the optimal wealth for the initial decision $\boldsymbol{x}$ and for a given trajectory of interest rates $\boldsymbol{r}$. It is possible to prove, cf. [23], that the corresponding second-stage linear program is regular in the sense of [64]. This means, inter alia, that $W(\boldsymbol{x}, \boldsymbol{r})$ is *jointly* continuous in $\boldsymbol{x}, \boldsymbol{r}$. As the Black-Derman-Toy lattice provides a *compact* support of distribution $P$, an appropriate choice of $c$ is (28) with $p = 1$. To delete one of selected scenarios $\boldsymbol{r}^s, s = 1, \ldots, S$, according to the introduced deletion rule means to find the minimal element $\|\boldsymbol{r}^{i_0} - \boldsymbol{r}^{j_0}\|^2$ of the matrix $\|\boldsymbol{r}^i - \boldsymbol{r}^j\|^2, 1 \le i,j \le S$, to delete scenario $\boldsymbol{r}^{i_0}$ and to allocate its probability to $\boldsymbol{r}^{j_0}$ or vice versa. Hence, the scenarios will no more be equiprobable. An alternative redistribution rule is to redistribute the probability $p_k (= \frac{1}{S})$ uniformly among the $S - 1$ remaining scenarios. It changes the above deletion rule to deletion of scenario $\boldsymbol{r}^k$ for which the average distance from other scenarios is minimal:

$$\sum_s \|\boldsymbol{r}^s - \boldsymbol{r}^k\| = \min_l \sum_s \|\boldsymbol{r}^s - \boldsymbol{r}^l\|.$$

Again, it is expedient to model the perturbations of the probability distribution $P$ using a finite dimensional parameter. This can be done by the contamination method.

*The contamination method* does not require any specific properties of the probability distribution $P$. In the framework of scenario-based stochastic programs, it can be used to study the influence of the assigned values of probabilities $p_s$ and of the whole probability structure, including additional stages and additional scenarios or branches of the scenario tree on the optimal value; see [19, 21].

Contamination means to model the perturbed probability distribution as

$$P_\lambda = (1 - \lambda)P + \lambda Q, \quad 0 \le \lambda \le 1, \tag{30}$$

the *probability distribution P contaminated by the probability distribution Q*. The contamination neighborhood $\mathcal{O}_\varepsilon(P) := \{\hat{P} | \hat{P} = (1 - \varepsilon)P + \varepsilon Q, \ \forall$ probability distributions $Q\}$ is not a neighborhood in the topological sense, but for $\varepsilon$ small enough, the contaminated distributions fall into a neighborhood of $P$. For fixed probability distributions $P, Q$, the objective function $F(\boldsymbol{x}, P_\lambda)$ in (20) computed for the contaminated distribution $P_\lambda$ is *linear or convex* in the parameter $\lambda$ and under modest assumptions, its optimal value

$$\varphi(\lambda) := \max_{\boldsymbol{x} \in \mathcal{X}} F(\boldsymbol{x}, P_\lambda)$$

is a finite convex function on $[0, 1]$ with a derivative $\varphi'(0^+)$ at $\lambda = 0^+$.

Bounds on the optimal value $\varphi(\lambda)$ for an arbitrary $\lambda \in [0, 1]$ follow by properties of convex functions:

$$(1 - \lambda)\varphi(0) + \lambda\varphi(1) \ge \varphi(\lambda) \ge \varphi(0) + \lambda\varphi'(0^+) \quad \forall \lambda \in [0, 1]. \tag{31}$$

For $F(\boldsymbol{x}, P)$ *linear* in $P$, the lower bound for the derivative $\varphi'(0^+)$ equals $F(\boldsymbol{x}(P), Q) - \varphi(0)$ where $\boldsymbol{x}(P)$ is an arbitrary optimal solution of the initial problem (20) obtained for the probability distribution $P$; if the optimal solution is unique, this lower bound is attained. Hence, evaluation of bounds in (31) requires the solution of another stochastic program of the type (20) for the new distribution $Q$ to get $\varphi(1)$ and evaluation of the expectation $F(\boldsymbol{x}(P), Q)$ at an already known optimal solution $\boldsymbol{x}(P)$ of the initial problem (20) but for the contaminating distribution $Q$.

For stability studies with respect to small changes in the underlying probability distribution, small values of the contamination parameter $\lambda$ are typical. The choice of $\lambda$ may reflect the degree of confidence in expert opinions represented as the contaminating distribution $Q$, and so on. Using a contaminating distribution $Q$ carried by additional scenarios or branches of the scenario tree, one can study the influence of including these additional "out-of-sample" scenarios; cf. [25] for application in portfolio optimization. By a suitable choice of $Q$, criteria on a right number of stages (cf. [43]) can be tested, the response on an increasing importance of a scenario can be quantified, etc.

*Example 3*. Consider the problem of investment decisions in the international debt and equity markets. Assume that historical data allows us to construct many scenarios of returns of investments in the considered assets categories. We denote these (in principle equiprobable) scenarios by $\omega^s, s = 1, \dots, S$, and let $P$ be the corresponding discrete probability distribution. Assume that for each of these scenarios, an outcome of a feasible investment strategy, say, $\boldsymbol{x} \in \mathcal{X}$ can be evaluated as $f_0(\boldsymbol{x}, \omega^s)$. Maximization of the expected outcome

$$F(\boldsymbol{x}, P) = \frac{1}{S}\sum_{s=1}^{S} f_0(\boldsymbol{x}, \omega^s) \quad \text{with respect to } \boldsymbol{x} \in \mathcal{X}$$

provides the optimal value $\varphi(P)$ and an optimal investment strategy $\boldsymbol{x}(P)$.

The historical data definitely do not cover all possible extremal situations on the market. Assume that experts suggest an additional scenario $\omega^*$. This is the only atom of the degenerated probability distribution $Q$, for which the best investment strategy is $\boldsymbol{x}(Q)$ – an optimal solution of $\max_{\boldsymbol{x} \in \mathcal{X}} f_0(\boldsymbol{x}, \omega^*)$. The contamination method explained above is based on the probability distribution $P_\lambda$, carried by the scenarios $\omega^s, s = 1, \dots, S$, and on the experts scenario $\omega^*$ with probabilities $\dfrac{1-\lambda}{S}$ for $s = 1, \dots, S$, and $p_* = \lambda$. The probability $\lambda$ assigns a weight to the view of the expert and the bounds (31) are valid for all $0 \le \lambda \le 1$. They clearly indicate how much the weight $\lambda$, interpreted as the degree of confidence to the investor's view, affects the outcome of the portfolio allocation.

The impact of a modification of every single scenario according to the investor's views on the performance of each asset class can be studied in a similar way. We use the initial probability distribution $P$ contaminated by $Q$, which is now carried by equiprobable scenarios $\hat{\omega}^s = \omega^s + \delta^s, s = 1, \dots, S$. The contamination parameter $\lambda$ relates again to the degree of confidence to the expert's view.

More specifically, assume now that the investment problem has been formulated and solved as a two-stage multiperiod stochastic program with the second-stage constraints of the type (16)–(17) or (18)–(19) for $t = 2, \dots, T_0$. The members of the investment committee expect that a special event at time $t^*$ will result in new developments which might require some changes of the investment policy. For simplicity assume that $t^*$ is one of time points of the discretization $t = 2, \dots, T_0$. Under these circumstances the initial decision problem should be formulated as a three-stage stochastic program with branching at time $t^*$. A possibility is to use first the contamination technique to analyze the influence of the new developments on the maximal expected outcome for various values of $\lambda$ corresponding to the degree of confidence in the new developments. The contaminating probability distribution $Q$ is carried again by $S$ equiprobable scenarios, say, $\tilde{\omega}^s, s = 1, \dots, S$, whose components $\tilde{\omega}^s_t = \omega^s_t, t < t^*$, whereas the remaining components reflect the foreseen new developments and are designed by the experts. Bounds (31) on the optimal value for the contaminated probability distribution $P_\lambda$ quantify the impact of the additional stage. This approach can be modified to cover more than one date of the foreseen changes.

*The worst case analysis and the minimax decision rule* are mostly used in cases of an *incomplete information about the probability distribution P* which is known to belong into a family $\mathscr{P}$ of probability distributions identified, e.g., by known values of some moments, by a given support, by qualitative properties, such as unimodality, or by unprecise values of probabilities of expert scenarios. We refer to [8, 50] for early examples of applications of the minimax decision rule in portfolio optimization.

When applied to (20) with $F(\boldsymbol{x}, P) := E_{P(\omega)} f_0(\boldsymbol{x}, \omega)$, the minimax decision rule means to select the decision $\boldsymbol{x}^*$ which maximizes the smallest possible expectation:

$$\boldsymbol{x}^* \in \arg\max_{\boldsymbol{x} \in \mathscr{X}} \left[ \min_{P(\omega) \in \mathscr{P}} E_{P(\omega)} f_0(\boldsymbol{x}, \omega) \right]. \tag{32}$$

The objective function of the inner minimization problem in (32) is linear in $P$, which means that for convex, compact set $\mathscr{P}$, the optimal *worst-case* probability distribution $P^* \in \mathscr{P}$ is one of extremal points of $\mathscr{P}$. In the framework of the moment problem, see [39], these extremal points are well described for $\mathscr{P}$ defined by a given support and by known values of certain generalized moments: For admissible moment values, the extremal distributions are discrete ones, concentrated in a modest number of points; hence, (32) is a *scenario-based program*. The problem is to get one of these discrete distributions, the worst-case distribution $P^*$ in its *dependence on the first-stage decision* $\boldsymbol{x}$. To this purpose, convexity, concavity or saddle properties of the integrand $f_0(\boldsymbol{x}, \bullet)$ are essential; recall the Jensen inequality valid for all probability distributions $P$ with a prescribed expected value $E_{P(\omega)}\omega = \bar{\omega}\ \forall P \in \mathscr{P}$ and for convex $f_0(\boldsymbol{x}, \bullet)$, so that the worst case distribution, independently on $\boldsymbol{x}$, is degenerated, concentrated in $\bar{\omega}$. On the other hand, for concave $f_0(\boldsymbol{x}, \bullet)$ and compact convex polyhedral support $\Omega$, the worst case distributions are concentrated in the extremal points of $\Omega$. This is an old result of Edmundson and Madansky [47] which provides the worst-case distribution independent of $\boldsymbol{x}$ under special assumptions, e.g., when $\Omega$ is a simplex or a Cartesian product of simplices.

This gives some possibilities in modeling and software development also for applications of multiperiod and multistage stochastic programs in portfolio optimization, cf. [31]. However, the required convexity or saddle property of $f_0(\boldsymbol{x}, \bullet)$ restricts the models to fixed recourse stochastic programs with random right-hand sides and/or coefficients in the objective functions that are supposed to depend on the random parameters in a *linear* way. Moreover for multistage stochastic programs, convexity of $f_0(\boldsymbol{x}, \bullet)$ in (32) depends upon special assumptions about the probability distribution of the stochastic data process (6).

*Example 4 – Incomplete knowledge of liabilities.* Let us modify the bond portfolio management problem treated in Example 2: Assume that the interest rate scenarios and their probabilities have been already fixed so that except for liabilities, the scenario dependent coefficients in the second-stage constraints (18) are known. Let us turn our attention to liabilities. Their full knowledge is not realistic for instance in management of newly established pension funds, portfolios of insurance companies, etc. We shall assume now that liabilities are random, $\boldsymbol{L}(\eta)$ with $\eta$ *independent of random interest rates.* The objective function, now the expected utility of the final wealth, assumes the form

$$E_{P(\eta)}\left\{\sum_s p_s U(W(\boldsymbol{x}, \boldsymbol{r}^s; \boldsymbol{L}(\eta)))\right\} \tag{33}$$

where $U(W(\boldsymbol{x}, \boldsymbol{r}^s; \boldsymbol{L}(\eta)))$ denotes the maximal contribution of portfolio management for a feasible first-stage decision $\boldsymbol{x}$, a given scenario $\boldsymbol{r}^s$ of interest rates and a realization $\boldsymbol{L}(\eta)$ of liabilities. However, the probability distribution of $\boldsymbol{L}(\eta)$ is not known completely. Using the available information, we want to get bounds on the maximal value of (33) subject to (23)–(24) and nonnegativity constraints.

As the first step, it is easy to realize that for a concave, nondecreasing utility function $U$, the individual terms $U(W(\boldsymbol{x}, \boldsymbol{r}^s; \boldsymbol{L}(\eta)))$ in the objective function (33) are concave in the right-hand sides $\boldsymbol{L}(\eta)$ taken as a parameter in evaluating the maximal attainable final wealth under scenario $\boldsymbol{r}^s$ for a fixed feasible first-stage decision $\boldsymbol{x}$. It means that Jensen's inequality provides an upper bound for the objective function (33):

$$E_{P(\eta)}\left\{\sum_s p_s U(W(\boldsymbol{x}, \boldsymbol{r}^s; \boldsymbol{L}(\eta)))\right\} \leq \sum_s p_s U(W(\boldsymbol{x}, \boldsymbol{r}^s; E_{P(\eta)}\boldsymbol{L}(\eta))).$$

The corresponding *upper bound for the optimal value* of (33) subject to constraints on the first-stage variables equals

$$\max_{\boldsymbol{x}} \sum_s p_s U(W(\boldsymbol{x}, \boldsymbol{r}^s; E_{P(\eta)}\boldsymbol{L}(\eta))).$$

Hence, *replacing the random liabilities by their expectations in the bond portfolio management problem leads to overestimating the maximal expected gain.*

The lower bound can be based on the Edmundson-Madansky inequality if, in addition, the components $L_t(\eta) := L(t;\eta)$ of $\boldsymbol{L}(\eta)$ are known to belong to finite intervals, say $[L_t', L_t'']$ for each $t$. The general bound, however, is computationally expensive unless the objective function is separable in individual liabilities, which is not our case. A trivial lower bound can be obtained by replacing all liabilities by their upper bounds $L_t''$; this bound will be rather loose. Another possibility is to assume a special structure of liabilities (their independence, a Markov property, etc.) in which case the lower bound can be simplified provided that the objective function remains concave with respect to the random variables used to model the liabilities. Assume for instance that

$$L(\eta) = Ga(\eta)$$

with a given matrix $G$ of the size $T_0 \times J$ and $a_j(\eta), j = 1, \dots, J$, mutually independent random variables with known expectations $E_{P(\eta)} a_j$ and known supports $[a'_j, a''_j]\ \forall j$. Accordingly, the individual objective functions

$$U(W(x, r^s; L(\eta))) = U(W(x, r^s; Ga(\eta))) := \tilde{U}^s(x; a(\eta)))$$

are concave in $a(\eta)$.

For small $J$, the following string of inequalities valid for each of scenarios $r^s$ and for all feasible first-stage solutions can be useful:

$$\begin{aligned} E_{P(\eta)} \tilde{U}^s(x; a(\eta)) &\geq \lambda_1 E_{P(\eta)} \tilde{U}^s(x; a'_1, a_2(\eta), \dots, a_J(\eta)) \\ &\quad + (1 - \lambda_1) E_{P(\eta)} \tilde{U}^s(x; a''_1, a_2(\eta), \dots, a_J(\eta)) \\ &\geq \sum_{\mathscr{J} \subset \{1,\dots,J\}} \prod_{j \in \mathscr{J}} \lambda_j \prod_{j \notin \mathscr{J}} (1 - \lambda_j) \tilde{U}^s(x; a_{\mathscr{J}}) \end{aligned} \tag{34}$$

where the components of $a_{\mathscr{J}}$ equal $a'_j$ for $j \in \mathscr{J}$ and $a''_j$ for $j \notin \mathscr{J}$ and

$$\lambda_j = \frac{a''_j - E_{P(\eta)} a_j}{a''_j - a'_j}, \quad j = 1, \dots, J.$$

Inequalities (34) imply that the lower bound for the maximal value of the objective function (33) can be obtained by solving the corresponding stochastic program based on $2^J S$ scenarios.

For instance for *pension funds* it is natural to assume that $G$ is a lower triangular matrix: The liabilities $L_1$ to be paid at the beginning of the first period are known with certainty and their portion, say, $\gamma L_1$ corresponds to unrepeated payments (e.g., final settlements or premiums) whereas the remaining main part of $L_1$ will be paid also in the subsequent period (continuing pensions). The liabilities $L_2(\eta)$ to be paid at the beginning of the period 2 can be modeled as

$$L_2(\eta) = (1 - \gamma) L_1 + a_2(\eta),$$

etc. Moreover, it is possible to assume that $a_j(\eta)$ are mutually independent so that (34) is a valid and tight lower bound that applies whenever the intervals $[a'_j, a''_j]\ \forall j$ and the expectations $E_{P(\eta)} a(\eta)$ are known.

If such *maximin and maximax bounds*

$$\max_{x \in \mathscr{X}} \inf_{P \in \mathscr{P}} F(x, P) \leq \varphi(P) \leq \max_{x \in \mathscr{X}} \sup_{P \in \mathscr{P}} F(x, P) \forall P \in \mathscr{P} \tag{35}$$

on the optimal value of the true program are available, they provide an important information on robustness of the optimal value within the considered family of probability distributions. Special algorithms [32] designed for this purpose, have not yet been applied in the context of portfolio management. A related, though less ambitious problem is to get bounds on the performance of an optimal solution $\boldsymbol{x}(P)$ obtained for a probability distribution $P \in \mathscr{P}$ using the corresponding worst-case and best-case probability distributions from $\mathscr{P}$.

## 5 Conclusions

We have presented several techniques suitable for analysis of results obtained by solution of stochastic programs designed to support decisions of portfolio managers. Such stochastic programs are just an approximation of reality and the goal is to get sensible and robust decisions for the underlying real-life problem. Depending on its nature, on the sources of errors due to approximation and simplification, the introduced methods indicate how to bridge the gap between the results obtained for the approximate problem and those valid for the true one, using the available information; the first numerical experience is reported in [3, 24, 25].

*Acknowledgement:* This work was partly supported by research project "Mathematical Methods in Stochastics" – CEZ:J13/98:113200008 and by grants No. 201/99/0264 and 402/99/1136 of the Grant Agency of the Czech Republic.

## References

[1] Attouch H, Wets RJ-B (1993) Quantitative stability of variational systems: III. $\varepsilon$-approximate solutions. Math. Progr. 61:197–214
[2] Beltratti A, Consiglio A, Zenios SA (1999) Scenario modeling for the management of international bond portfolios. In: [74], pp. 259–280
[3] Bertocchi M, Dupačová J, Moriggia V (1999) Sensitivity of bond portfolio's behavior with respect to random movements in yield curve: A simulation study. To appear in Annals of Oper. Res. (APMOD'98)
[4] Black F, Derman E, Toy W (1990) A one-factor model of interest rates and its application to treasury bond options. Financial Analysts Journal: 33–39
[5] Bradley SP, Crane DB (1972) A dynamic model for bond portfolio management. Manag. Sci. 19:139–151
[6] Bradley SP, Crane DB (1980) Managing a bank bond portfolio over time. In: Dempster MAH (ed.) Stochastic programming. Academic Press, London, pp. 449–471
[7] Brodt AI (1984) Intertemporal bank asset and liability management. J. Bank. Res. 15:82–94
[8] Bühler W (1981) Flexible Investitions- und Finanzplannung bei unvollkommen bekannten Übergangswahrscheinlichkeiten. OR Spektrum 2:207–221
[9] Cariño DR, Myers DH, Ziemba WT (1998) Concepts, technical issues and uses of the Russell-Yasuda Kasai financial planning model. Oper. Res. 46:450–462
[10] Cariño DR, Ziemba WT (1998) Formulation of the Russel-Yasuda Kasai financial planning model. Oper. Res. 46:433–449
[11] Chopra WK, Ziemba WT (1993) The effect of errors in means, variances and covariances on optimal portfolio choice. J. of Portfolio Mgt. 19:6–11
[12] Consigli G, Dempster MAH (1998) Dynamic stochastic programming for asset-liability management. Annals of Oper. Res. 81:131–161
[13] Constantinides GM, Malliaris AG (1995) Portfolio theory. In: Jarrow RA, Maksimovic V, Ziemba WT (eds.) Finance. Handbooks in OR & MS, Vol. 9, Elsevier, pp. 1–30
[14] Dembo RS (1993) Scenario immunization. In: [76], pp. 290–308

[15] Dempster MAH, Ireland AM (1988) A financial expert decision support system. In: Mitra G (ed.) Mathematical models for decision support. NATO ASI Series Vol. F48, Springer, Berlin, pp. 415–440
[16] Dempster MAH, Thorlacius AE (1998) Stochastic simulation of international economic variables and asset returns: The Falcon asset model. Preprint
[17] Dert C (1995) Asset liability management for pension funds. A multistage chance-constrained programming approach. Ph.D. thesis, Erasmus University
[18] Dupačová J (1991) Stochastic programming models in banking. Tutorial paper, IIASA Laxenburg. See also Ekonomicko – Matematický Obzor 27:201–234 (in Czech)
[19] Dupačová J (1995) Postoptimality for multistage stochastic linear programs. Annals of Oper. Res. 56:65–78
[20] Dupačová J (1995) Multistage stochastic programs: The state-of-the art and selected bibliography. Kybernetika 31:151–174
[21] Dupačová J (1996) Scenario based stochastic programs: Resistance with respect to sample. Annals of Oper. Res. 64:21–38
[22] Dupačová J (1996) Uncertainty about input data in portfolio management. In: Bertocchi M et al. (eds.) Modelling techniques for financial markets and bank management. Physica Verlag, pp. 17–33
[23] Dupačová J (1999) Stability properties of a bond portfolio management problem. To appear in Annals of Oper. Res. (APMOD'98)
[24] Dupačová J, Bertocchi M (1999) From data to model and back to data: A bond portfolio management problem. To appear in EJOR
[25] Dupačová J, Bertocchi M, Moriggia V (1998) Postoptimality for scenario based financial planning models with an application to bond portfolio management. In: [78], pp. 263–285
[26] Dupačová J, Consigli G, Wallace SW (1998) Scenarios for multistage stochastic programs. Submitted to Annals of Oper. Res
[27] Dupačová J, Römisch W (1998) Quantitative stability for scenario-based stochastic programs. In: Hušková M et al. (eds.) Prague Stochastics'98. JČMF, Prague, pp. 119–124
[28] Dupačová J, Römisch W (1999) Scenario-based stochastic programs: Changes of the probability distribution. Extended abstract for IFIP TC7 Conference, Cambridge
[29] Dupačová J, Wets RJ-B (1988) Asymptotic behavior of statistical estimators and of optimal solutions of stochastic optimization problems. Ann. Statist. 16:1517–1549
[30] Elton EJ, Gruber MJ (1987) Modern portfolio theory and investment analysis. Wiley, New York (3rd edition)
[31] Frauendorfer K, Schürle M (1998) Barycentric approximation of stochastic interest rate processes. In: [78], pp. 231–262
[32] Gaivoronski AA (1991) A numerical method for solving stochastic programming problems with moment constraints on a distribution function. Annals of Oper. Res. 31:347–369
[33] Golub B et al. (1995) Stochastic programming models for portfolio optimization with mortgage-backed securities. EJOR 82:282–296
[34] Grauer RR, Hakansson NH (1985) Returns on levered actively managed long-run portfolios of stocks, bonds and bills. Financial Analyst Journal: 24–43
[35] Ingersol JEJr. (1987) Theory of financial decision making. Rowman and Littlefield, Totowa
[36] Jamshidian F, Zhu Yu (1997) Scenario simulation: Theory and methodology. Finance and Stochastics 1:43–67
[37] Kan YuS, Mistryukov AA (1998) On the equivalence in stochastic programming with probability and quantile objectives. In: Marti K, Kall P (eds.) Stochastic programming methods and technical applications. LN in Economics and Math. Systems 458, Springer, Berlin, pp. 145–153
[38] Kataoka S (1963) A stochastic programming model. Econometrica 31:181–196
[39] Kemperman JMB (1968) The general moment problem, a geometric approach. Ann. Math. Statist. 39:93–122
[40] King AJ (1993) Asymmetric risk measures and tracking models for portfolio optimization under uncertainty. Annals of Oper. Res. 45:165–178
[41] King AJ, Rockafellar RT (1993) Asymptotic theory for solutions in statistical estimation and stochastic programming. Math. of Oper. Res. 18:148–162
[42] Kira D, Ziemba WT (1977) Equivalence among alternative portfolio selection criteria. In: Levy H, Sarnat M (eds.) Financial decision making under uncertainty. Academic Press, New York, pp. 151–161

[43] Klaassen P (1998) Financial asset-pricing theory and stochastic programming models for asset/liability management: A synthesis. Manag. Sci. 44:31–48
[44] Konno H, Yamazaki H (1991) Mean-absolute deviation portfolio optimization model and its applications to Tokyo stock market. Manag. Sci. 37:519–531
[45] Kusy MI, Ziemba WT (1986) A bank asset and liability management model. Oper. Res. 34:356–376
[46] Lane M, Hutchinson P (1980) A model for managing a certificate of deposit portfolio under uncertainty. In: Dempster MAH (ed.) Stochastic programming. Academic Press, London, pp. 473–495
[47] Madansky A (1959) Bounds on the expectation of a convex function of a multivariate random variable. Ann. Math. Statist. 30:743–746
[48] Mak Wai-Kei, Morton DP, Wood RK (1999) Monte Carlo bounding techniques for determining solution quality in stochastic programs. OR Letters 24:47–56
[49] Markowitz HM (1952) Portfolio selection. J. of Finance 6:77–91
[50] Moèseke P van (1980) Stochastic portfolio programming: The game solution. In: Dempster MAH (ed.) Stochastic programming. Academic Press, London, pp. 497–505
[51] Mulvey JM (1989) A surplus optimization perspective. Investment Management Review 3:31–39
[52] Mulvey JM (1993) Incorporating transaction costs in models for asset allocation. In: [76], pp. 243–259
[53] Mulvey JM, Rosenbaum DP, Shetty B (1997) Strategic financial risk management and operations research. EJOR 97:1–16
[54] Mulvey JM, Thorlacius AE (1998) The Towers Perrin global capital market scenario generation system. In: [78], pp. 286–312
[55] Mulvey JM, Vanderbei RJ, Zenios SA (1995) Robust optimization of large scale systems. Oper. Res. 43:264–281
[56] Mulvey JM, Vladimirou H (1989) Stochastic network optimization models for investment planning. Annals of Oper. Res. 20:187–217
[57] Mulvey JM, Zenios SA (1994) Capturing the correlations of fixed-income instruments. Manag. Sci 40:1329–1342
[58] Mulvey JM, Ziemba WT (1995) Asset and liability allocation in a global environment. In: Jarrow RA, Maksimovic V, Ziemba WT (eds.) Finance. Handbooks in OR & MS, Vol. 9, Elsevier, pp. 475–463
[59] Nielsen SS, Zenios SA (1996) A stochastic programming model for funding single premium deferred annuities. Math. Progr. 75:177–200
[60] Perold AF (1986) Large-scale portfolio optimization. Manag. Sci. 30:1143–1160
[61] Pflug GCh (1997) Risk-reshaping contracts and stochastic optimization. CEJORE 5:205–229
[62] Pflug GCh, Swietanowski A (1998) Optimal scenario tree generation for multiperiod financial optimization. Technical Report AURORA TR1998-22, University of Vienna
[63] Rachev ST (1991) Probability metrics and the stability of stochastic models. Wiley, New York
[64] Robinson SM (1977) A characterization of stability in linear programming. Oper. Res. 25:435–447
[65] Robinson SM, Wets RJ-B (1987) Stability in two-stage stochastic programming. SIAM J. Control and Optimization 25:1409–1416
[66] Rockafellar RT, Wets RJ-B (1976) Nonanticipativity and $\mathscr{L}^1$-martingales in stochastic optimization problems. Math. Progr. Study 6:170–187
[67] Roy AD (1952) Safety-first and the holding of assets. Econometrica 20:431–449
[68] Seppälä J (1994) The diversification of currency loans: A comparison between safety-first and mean-variance criteria. EJOR 74:325–343
[69] Serfling RJ (1980) Approximation theorems in mathematical statistics. Wiley, New York
[70] Shapiro A (1991) Asymptotic analysis of stochastic programs. Annals of Oper. Res. 30:169–186
[71] Shapiro JF (1988) Stochastic programming models for dedicated portfolio selection. In: Mitra B (ed.) Mathematical models for decision support. NATO ASI Series, Vol. F48, Springer, Berlin, pp. 587–611
[72] Sharpe WF (1971) Mean-absolute-deviation characteristic lines for securities and portfolios. Manag. Sci. 18:B1–B13
[73] Telser LG (1955–56) Safety first and hedging. Review of Economic Studies 23:1–16

[74] Wets RJ-B, Ziemba WT (eds.) (1999) Stochastic programming: The state of the art 1998. Annals of Oper. Res. 89
[75] Worzel KJ, Vassiadou-Zeniou Ch, Zenios SA (1994) Integrated simulation and optimization models for tracking indices of fixed-income securities. Oper. Res. 42:223–233
[76] Zenios SA (ed.) (1993) Financial optimization. Cambridge University Press
[77] Zenios SA et al. (1998) Dynamic models for fixed-income portfolio management under uncertainty. J. of Economic Dynamics and Control 22:1517–1541
[78] Ziemba WT, Mulvey JM (eds.) (1998) Worldwide asset and liability modeling. Cambridge University Press
[79] Ziemba WT, Vickson RG (eds.) (1975) Stochastic optimization models in finance. Academic Press, London