

# Multistage Stochastic Programs

- Basic Formulations
- Multistage Stochastic *Linear* Program with Recourse: all functions are *linear in decision variables*
- Problem of Private Investor Revisited
- Horizon and Stages
- Flower-Girl Problem
- Scenario-Based Stochastic Linear Programs — Generation of Scenario Trees
- Financial Applications — General Scheme
- Illustration: Russell-Yasuda Kasai Model, ALM models

# Multistage stochastic programs I.

**$T$ -stage stochastic program** ... stochastic data process

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

realizations are (multidimensional) data trajectories;

vector decision process  $\mathbf{x} = \{\mathbf{x}_1, \dots, \mathbf{x}_T\}$ ,

$t \in \{1, \dots, T\}$  — stage index,  $T$  depth of the tree.

Components of  $\omega$  and decisions  $\mathbf{x}_2, \dots, \mathbf{x}_T$  are random vectors defined on probability space  $(\mathcal{Z}, \mathcal{F}, \mu)$ ,  $\mathbf{x}_1$  is nonrandom.

Decision process  $\mathbf{x}$  is NONANTICIPATIVE i.e. decisions taken at any stage of the process do not depend on *future realizations* of stochastic data nor on future decisions whereas past information and knowledge of probability distribution of the data process are exploited:

For  $t = 1, \dots, T$ , DENOTE  $\mathcal{F}_{t-1} \subset \mathcal{F}$ ,  $\sigma$ -field generated by the part  $\omega^{t-1, \bullet} := \{\omega_1, \dots, \omega_{t-1}\}$  of stochastic data process  $\omega$  that precedes stage  $t$ .

NONANTICIPATIVITY means that  $t$ -th stage decision  $\mathbf{x}_t$  is  $\mathcal{F}_{t-1}$ -measurable.

# Multistage stochastic programs II.

DENOTE  $\mathbf{x}^{t-1, \bullet} = (\mathbf{x}_1, \dots, \mathbf{x}_{t-1})$  sequence of decisions at stages  $t = 1, \dots, t - 1$ ,

$P$  distribution function of  $\omega$ ,

$P_t$  marginal probability distribution of  $\omega_t$ ,

$P_t(\cdot | \omega^{t-1, \bullet})$  its conditional probability distribution.

ASSUME that all infima are attained & all expectations exist.

1st-STAGE DECISIONS consist of all decisions that have to be selected BEFORE further information is revealed whereas the 2nd-stage decisions are allowed to adapt to this information, etc.

In each stage, decisions are limited by constraints that may depend on previous decisions and observations.

# Multistage stochastic programs III.

EXAMPLE: NESTED FORM of multistage stochastic LINEAR program resembles backward recursion of stochastic dynamic programming with additive overall cost function:

Minimize

$$\mathbf{c}_1^\top \mathbf{x}_1 + E_P \{ \varphi_1(\mathbf{x}_1, \omega) \}$$

on  $\mathcal{X}_1 := \{ \mathbf{x}_1 | \mathbf{A}_1 \mathbf{x}_1 = \mathbf{b}_1, \quad \mathbf{l}_1 \leq \mathbf{x}_1 \leq \mathbf{u}_1 \}$

$\varphi_{t-1}(\cdot, \cdot)$ ,  $t = 2, \dots, T$ , are defined recursively as

$$\varphi_{t-1}(\mathbf{x}^{t-1, \bullet}, \omega^{t-1, \bullet}) =$$

$$\min_{\mathbf{x}_t} \left[ \mathbf{c}_t(\omega^{t-1, \bullet})^\top \mathbf{x}_t + E_{P_t(\bullet | \omega^{t-1, \bullet})} \{ \varphi_t(\mathbf{x}^{t-1, \bullet}, \mathbf{x}_t, \omega^{t-1, \bullet}, \omega_t) \} \right]$$

subject to constraints

$$\mathbf{B}_t(\omega^{t-1, \bullet}) \mathbf{x}_{t-1} + \mathbf{A}_t(\omega^{t-1, \bullet}) \mathbf{x}_t = \mathbf{b}_t(\omega^{t-1, \bullet}) \quad \text{a.s.}$$

$$\mathbf{l}_t(\omega^{t-1, \bullet}) \leq \mathbf{x}_t \leq \mathbf{u}_t(\omega^{t-1, \bullet}), \quad \text{a.s.}$$

and  $\varphi_T \equiv f_T(\mathbf{x}_T)$  is explicitly given. For the 1st stage, all elements of  $\mathbf{b}_1, \mathbf{c}_1, \mathbf{A}_1$  are known and the **main decision variable** is  $\mathbf{x}_1$  that corresponds to 1st stage.

# Multistage stochastic programs IV.

FOR APPLICATIONS: Approximate  $P$  by a discrete distribution, carried by finite number of atoms (scenarios),  $\omega^k, k = 1, \dots, K \rightarrow$  supports of marginal and conditional probability distributions  $P_t, P_t(\cdot|\omega^{t-1,\bullet}) \forall t$  are finite sets.

For disjoint sets of indices  $\mathcal{K}_t, t = 2, \dots, T$ , we list as  $\tilde{\omega}_{k_t}, k_t \in \mathcal{K}_t$  all possible realizations of  $\omega^{t-1,\bullet}$  and denote by the same subscripts the corresponding values of  $t$ -th stage coefficients.

Total number of scenarios  $K$  equals the number of elements of  $\mathcal{K}_T$ .

Each scenario  $\omega^k = \{\omega_1^k, \dots, \omega_{T-1}^k\}$  generates sequence of coefficients  $\{\mathbf{c}_{k_2}, \dots, \mathbf{c}_{k_T}\}, \{\mathbf{A}_{k_2}, \dots, \mathbf{A}_{k_T}\}, \{\mathbf{B}_{k_2}, \dots, \mathbf{B}_{k_T}\}, \{\mathbf{b}_{k_2}, \dots, \mathbf{b}_{k_T}\}, \{\mathbf{l}_{k_2}, \dots, \mathbf{l}_{k_T}\}, \{\mathbf{u}_{k_2}, \dots, \mathbf{u}_{k_T}\}.$

Data are organized in the form of SCENARIO TREE:

Nodes are determined by all considered realizations

$\tilde{\omega}_{k_t}, k_t \in \mathcal{K}_t, t = 2, \dots, T$ , and by the root indexed as  $k_1 = 1$ .

Each realization  $\tilde{\omega}_{k_{t+1}}$  of  $\omega^{t,\bullet}, t = 1, \dots, T$ , has a UNIQUE ANCESTOR  $\tilde{\omega}_{k_t}$  (realization of  $\omega^{t-1,\bullet}$ ), denoted by subscript  $a(k_{t+1})$ , and a finite number of DESCENDANTS — realizations of  $\omega^{t+1,\bullet}$ .

# Multistage stochastic programs V.

Given horizon & prescribed structure of stages  $\rightarrow$  “arborescent” form of  $T$ -stage scenario-based stochastic linear program with additive recourse

minimize

$$\mathbf{c}_1^\top \mathbf{x}_1 + \sum_{k_2 \in \mathcal{K}_2} p_{k_2} \mathbf{c}_{k_2}^\top \mathbf{x}_{k_2} + \sum_{k_3 \in \mathcal{K}_3} p_{k_3} \mathbf{c}_{k_3}^\top \mathbf{x}_{k_3} + \cdots + \sum_{k_T \in \mathcal{K}_T} p_{k_T} \mathbf{c}_{k_T}^\top \mathbf{x}_{k_T} \quad (1)$$

subject to

$$\begin{aligned} \mathbf{A}_1 \mathbf{x}_1 &= \mathbf{b}_1 \\ \mathbf{B}_{k_2} \mathbf{x}_1 + \mathbf{A}_{k_2} \mathbf{x}_{k_2} &= \mathbf{b}_{k_2}, \quad k_2 \in \mathcal{K}_2 \\ \mathbf{B}_{k_3} \mathbf{x}_{a(k_3)} + \mathbf{A}_{k_3} \mathbf{x}_{k_3} &= \mathbf{b}_{k_3}, \quad k_3 \in \mathcal{K}_3 \\ &\vdots \\ \mathbf{B}_{k_T} \mathbf{x}_{a(k_T)} + \mathbf{A}_{k_T} \mathbf{x}_{k_T} &= \mathbf{b}_{k_T}, \quad k_T \in \mathcal{K}_T \\ \mathbf{l}_1 \leq \mathbf{x}_1 \leq \mathbf{u}_1, \mathbf{l}_{k_t} \leq \mathbf{x}_{k_t} \leq \mathbf{u}_{k_t}, \quad k_t \in \mathcal{K}_t, \quad t = 2, \dots, T. \end{aligned} \quad (2)$$

Path probabilities  $p_{k_t} > 0 \forall k_t$ ,  $\sum_{k_t \in \mathcal{K}_t} p_{k_t} = 1$ ,  $t = 2, \dots, T$ , of partial sequences of coefficients are probabilities of realizations of  $\omega^{t-1, \bullet} \forall t$ . Probabilities  $p^k$  of individual scenarios  $\omega^k$  are equal to  $p_{k_T}$ .

Nonanticipativity constraints are included implicitly.

# Horizon and stages I.

NOTE: (1)–(2) may correspond also to  $T$ -period two-stage stochastic program based on the same scenarios: Except for the root, there is only one descendant  $d(k_t)$  of each of  $t$ -th stage nodes, that is, the transition probabilities  $\pi_{k_t, d(k_t)} = 1 \forall k_t \in \mathcal{K}_t, t = 2, \dots, T - 1$ . Scenarios are identified by sequences  $\{k_2, \dots, k_T\}$  such that  $k_t \in \mathcal{K}_t, k_{t+1} = d(k_t) \forall t \rightarrow$  objective function (1) simplifies to

$$\mathbf{c}_1^\top \mathbf{x}_1 + \sum_{k_T \in \mathcal{K}_T} p_{k_T} [\mathbf{c}_{k_2}^\top \mathbf{x}_{k_2} + \mathbf{c}_{k_3}^\top \mathbf{x}_{k_3} + \dots + \mathbf{c}_{k_T}^\top \mathbf{x}_{k_T}]. \quad (3)$$

Problem (3), (2) is called TWO-STAGE RELAXATION of MSLP (1)–(2).  
Insert FIGURE 22n

# Why to use multistage formulation?

## Difference between TWO-STAGE MULTIPERIOD AND MULTISTAGE STOCHASTIC PROGRAMS

- In multiperiod two-stage problems decisions at all time instances  $t = 1, \dots, T$  are made at once, no further information is expected;
- Hedging against all considered **unrelated** scenarios of possible developments is assumed;
- Except for the 1st stage no nonanticipativity constraints appear; The input — no tree structure, just *fan of scenarios*.
  - Robustness, stability of solutions: similar subscenarios result in similar decisions even for  $t > 1$ ;
  - Stochastic specification (interstage dependence) is reflected;
  - Number of nodes decreases.



# Scenario splitted form

With EXPLICIT INCLUSION OF NONANTICIPATIVITY CONSTRAINTS, scenario-based multiperiod or multistage stochastic programs with linear constraints can be again written as a large-scale deterministic program:

Given scenario  $\omega^k$  denote by  $\mathbf{c}(\omega^k)$  vector composed of all corresponding coefficients, say,  $\mathbf{c}_1, \mathbf{c}_{k_t}, t = 2, \dots, T$ , in objective function, by  $\mathbf{A}(\omega^k)$  matrix of all coefficients of system of constraints (2) for scenario  $\omega^k$ , and by  $\mathbf{b}(\omega^k), \mathbf{l}(\omega^k), \mathbf{u}(\omega^k)$  vectors composed of right-hand sides in (2) and bounds of box constraints for scenario  $\omega^k$ .

SCENARIO-SPLITTED form of  $T$ -stage stochastic linear program is

$$\min_{\mathbf{x} \in \mathcal{C}} \left\{ \sum_{k=1}^K p^k \mathbf{c}(\omega^k)^\top \mathbf{x}(\omega^k) \right\} \quad (4)$$

subject to

$$\mathbf{A}(\omega^k) \mathbf{x}(\omega^k) = \mathbf{b}(\omega^k), \mathbf{l}(\omega^k) \leq \mathbf{x}(\omega^k) \leq \mathbf{u}(\omega^k) \forall k, \quad (5)$$

$\mathcal{X}$  is defined by deterministic constraints on  $\mathbf{x}_t(\omega^k) \forall t, k$ ,  $\mathcal{C}$  by nonanticipativity conditions, and  $\mathbf{x}(\omega^k)$  is the corresponding decision vector composed of stage related subvectors  $\mathbf{x}_t(\omega^k) \forall t$

# Problem of private investor revisited

Investor wishes to raise enough money for his child college education  $N$  years from now by investing  $w$  into some of  $I$  considered investments.

Tuition goal is  $g$ ,

exceeding  $g$  after  $N$  years  $\rightarrow$  additional income of  $q\%$  of the excess,  
not meeting the goal  $\rightarrow$  borrowing at the rate  $r > q$ .

Investor plans to *revise* his investment at certain time instances prior to  $N$  using additional information that will be available in future.

**Decision points** (stages) and the corresponding time periods are indexed by  $t = 1, \dots, T - 1$ , and horizon  $N$  corresponds to  $t = T$ .

**Main uncertainty:** returns  $\rho_i(t, \omega)$  on investments  $i$  in each period  $t$  depend on an underlying *random element*  $\omega$  and are observable at the end period  $t$ . Investment decisions  $x_i(t, \omega)$  made at the beginning of period  $t$  can be only based on the already observed part of the trajectory of  $\omega$ , i.e., they are *nonanticipative* of future outcomes.

$\implies$  at the beginning of 1st period, investment decisions  $x_i(1, \omega) = x_i(1)$  are fixed for all  $\omega$  belonging to a probability space  $(\Omega, \mathcal{F}, P)$ .

## Problem of private investor revisited cont.

Let  $T = 3$  — some of considered investments (term deposit or short term bond) mature in  $N_1 < N$  years  $\rightarrow$  portfolio has to be restructured at time  $N_1$ ; one more stage of decision process.

Put  $\omega = (\omega_1, \omega_2)$ , trajectory up to  $N_1$  and its continuation from  $N_1$  to  $N$ . Denote by  $\tilde{\omega}_{k_2}$ ,  $k_2 \in \mathcal{K}_2$  considered realizations of  $\omega_1$ ,  $p_{k_2}$  their probabilities and by  $\tilde{\omega}_{k_3}$ ,  $k_3 \in \mathcal{K}_3$  possible realizations of  $\omega^2$  grouped into sets  $\mathcal{D}(k_2)$ ,  $k_2 \in \mathcal{K}_2$  for which the conditional probabilities  $\pi_{k_2, k_3} \neq 0$ ; notice that  $\sum_{k_3 \in \mathcal{D}(k_2)} \pi_{k_2, k_3} = 1 \forall k_2$ . This information about the discrete probability distribution may be represented by a scenario tree, see FIGURE 21n

1st stage decisions  $x_i(1)$  are scenario independent, returns  $\varrho_i(1, \omega)$  and decisions  $x_i(2, \omega)$  at 2nd stage of decision process depend only on the first part  $\omega_1$  of  $\omega$ ,

subsequent returns  $\varrho_i(2, \omega)$ , final decisions and compensations depend on whole history, i.e., on scenarios  $\omega^s$  which consist of  $\tilde{\omega}_{k_2}$  and of their “extension” to  $\tilde{\omega}_{k_3}$ ,  $k_3 \in \mathcal{D}(k_2)$ ,  $k_2 \in \mathcal{K}_2$ . Their probabilities  $p^s$  equal

$$p_{k_2} \pi_{k_2, k_3}.$$

## Problem of private investor revisited cont.

Following notation introduced above, we assign subscripts  $k_2, k_3$  to random coefficients  $\rho$  and to decision variables  $x_i, y^+, y^-$  which appear in 2nd and 3rd stages. The problem — a *three-stage stochastic linear program* — reads

$$\text{maximize} \quad \sum_{k_2 \in \mathcal{K}_2} p_{k_2} \sum_{k_3 \in \mathcal{D}(k_2)} \pi_{k_2, k_3} (qy_{k_3}^+ - ry_{k_3}^-)$$

subject to

$$\sum_i x_i(1) = w, \quad x_i(t) \geq 0 \quad \forall i, t$$

$$\sum_i \rho_{ik_2} x_i(1) - \sum_i x_{ik_2} = 0, \quad k_2 \in \mathcal{K}_2$$

$$\sum_i \rho_{ik_3} x_{ik_2} - y_{k_3}^+ + y_{k_3}^- = g, \quad k_3 \in \mathcal{D}(k_2), k_2 \in \mathcal{K}_2$$

and nonnegativity of all variables.

# Horizon and stages I.

Above formulations of SP model rely on *already fixed topology of stages*, possibly with long irregular time steps in comparison with time discretization of data process.

**Choice of stages, branching scheme, scenarios and their probabilities** influence the optimal 1st-stage decision and the overall optimal value. Stages do not necessarily refer to time periods, they correspond to steps in the decision process. To use multiperiod two-stage model or to assign one stage to each of possible discretization points are two extreme cases. Requirements of various applications may lead to different topologies of decision points.

In majority of cases, the horizon and the stages are declared as given. In practice, various situations can be distinguished:

- Both the horizon and stages are determined ad hoc, often for purposes of testing numerical approaches and/or software.
- Both the horizon and stages are determined, e.g., by the nature of the real-life technological process.

## Horizon and stages II.

- The horizon is tied to a fixed date, e.g., to the end of the fiscal year, to a date related with the annual Board of Directors' meeting, or to the end date of a screening study. Stages are sometimes dictated by the nature of the solved problem, e.g., by the dates of maturity of bonds, expiration dates of options or by periodic (quarterly, annual, etc.) management review meetings. In other cases, they are obtained by application of heuristic rules and/or experience, taking into account limitations due to numerical tractability. *Rolling forward* after the  $T$ -stage problem has been solved, the first-stage decision accepted and new information exploited means to solve a subsequent  $T - 1$ -stage stochastic program with a reduced number of stages or possibly another  $T$ -stage problem with a different topology of stages.
- The horizon is connected with a time interval of a fixed (possibly even infinite) length, given for instance by the periodicity of the underlying random process, and the number of stages is chosen in dependence on the available computing facilities. Rolling forward means here repeated solution of a  $T$ -stage problem of the same structure of stages with the initial state of the system determined by the applied first-stage decision and by observation of the value of

# Flower-Girl Problem I

Flower-girl sells roses at price  $c$  and has to buy them at cost  $p$  before she starts selling.

Flowers left over at the end of the day can be stored and sold the next day, when she starts selling the old roses.

The roses cannot be carried over more than one additional day at the end of which they are thrown away.

The demand is random,  $\omega_t$  denotes demand on the  $t$ -th day. The flower-girl wants to maximize her total expected profit.

**Horizon** is related to the number of days for which the flower-girl continues selling roses without break (and also to the fact that our formulation treats only one-period carry-over).

Assume, that flower-girl sells roses only during weekend, orders amount  $x_1$  on Friday evening, registers demand  $\omega_1$  on Saturday, stores unsold roses (without any additional cost) and, possibly, buys  $x_2(\omega_1)$  new roses.

Denote  $s_2(\omega_1)$  stock left for 2nd day,  $z(\omega_1, \omega_2)$  amount of unsold roses at the end of 2nd day which is also affected by the demand  $\omega_2$  on Sunday.

# Flower-Girl Problem II

All variables are nonnegative. Constraints:

$$x_1 - s_2(\omega_1) \leq \omega_1, \quad x_2(\omega_1) + s_2(\omega_1) - z(\omega_1, \omega_2) \leq \omega_2;$$

total profit is

$$(c - p)(x_1 + x_2(\omega_1)) - cz(\omega_1, \omega_2).$$

If the demand  $\omega_1, \omega_2$  is known in advance, then one of the optimal solutions is to buy  $x_1 = \omega_1$  and  $x_2(\omega_1) = \omega_2$  roses which gives the maximal profit of  $(c - p)(\omega_1 + \omega_2)$ .

Scenario-based version of this 3-stage problem. Scenario tree consists of  $S$  scenarios corresponding to the considered realizations  $\omega^s = (\omega_1^s, \omega_2^s)$ ,  $s = 1, \dots, S$ , of the demand on 1st and 2nd day, their probabilities are  $p^s$ ,  $s = 1, \dots, S$ . We denote  $\tilde{\omega}_{k_2} = b_{k_2}$ ,  $k_2 \in \mathcal{K}_2$  possible realizations of  $\omega_1$ , by  $p_{k_2}$  their probabilities, by  $\omega_{k_3}$  the possible realizations of  $(\omega_1, \omega_2)$  conditional on  $\omega_{k_2}$  and by  $\pi_{k_2, k_3}$  their (conditional) probabilities.

Corresponding realizations of demand on the second day will be denoted  $b_{k_3}$ .



# Flower-Girl Problem – Formulation

In the introduced notation the problem reads:

$$\text{maximize } (c - p)x_1 + \sum_{k_2 \in \mathcal{K}_E} p_{k_2} [(c - p)x_{2k_2} - c \sum_{k_3 \in \mathcal{D}(k_2)} \pi_{k_2, k_3} z_{k_3}]$$

subject to

$$x_1 - s_{2k_2} \leq b_{k_2}, \quad k_2 \in \mathcal{K}_2$$

$$x_{2k_2} + s_{2k_2} - z_{k_3} \leq b_{k_3}, \quad k_3 \in \mathcal{D}(k_2), \quad k_2 \in \mathcal{K}_2$$

and nonnegativity constraints.

# Flower-Girl Problem – T stages

Generalization to  $T$ -stage problem ( $T > 3$ ) is obvious.

We index by  $t$  all variables related with the stage  $t$ , i.e., amount of roses ordered ( $x$ ), stored ( $s$ ) and thrown away ( $z$ ) at the end of  $(t - 1)$ st day; notice that  $z_T$  plays the role of the only decision variable at the last stage. We obtain:

$$\text{maximize } (c - p)x_1 + E\left\{(c - p) \sum_{t=2}^{T-1} x_t(\omega^{t-1, \bullet}) - c \sum_{t=2}^T z_t(\omega^{t-1, \bullet})\right\}$$

subject to

$$x_1 + s_1 - s_2(\omega_1) - z_2(\omega_1) \leq \omega_1$$

$$x_t(\omega^{t-1, \bullet}) + s_t(\omega^{t-1, \bullet}) - s_{t+1}(\omega^{t, \bullet}) - z_{t+1}(\omega^{t, \bullet}) \leq \omega_t, \quad t = 2, \dots, T - 1$$

$$s_t(\omega^{t-1, \bullet}) - z_{t+1}(\omega^{t, \bullet}) \leq \omega_t, \quad t = 1, \dots, T - 1$$

with  $s_T(\omega) \equiv 0$  and nonnegativity of all variables. In case that the initial supply  $s_1 = 0$ , one gets  $z_2(\omega_1) \equiv 0$ .

*Number of stages equals one plus the number of days for which the flower-girl sells roses without break, for  $T = 3$ , the last inequalities are redundant.*

# Flower-Girl Problem – Discussion

Scenario-based formulation of the  $T$ -stage problem can be written in arborescent form or in split variable form with explicit nonanticipativity constraints.

Notice that flower-girl problem should be more realistically formulated as an *integer* stochastic program.

Imagine now that the flower-girl wants to earn as much as possible during the two months of her high-school vacations; such a 63 stage problem may be solvable thanks to its simple form. Still some other possibilities should be examined.

Program may be rolled forward in time with an essentially shorter horizon, say, for  $T = 8$  which covers a whole week. This means that the flower-girl decides as if she plans to maximize her profit over each one-week period and solves the problem every day with a known non-zero initial supply of roses and with a new scenario tree spanning over the horizon of the next 8 days.

Another possibility is aggregation of stages above a tractable horizon.

# Generation of Scenario Trees I.

To generate **SCENARIOS FOR MULTISTAGE SP** means to replace the initial probability distribution  $P$  of  $\omega = (\omega_1, \omega_2, \dots, \omega_{T-1})$  by a discrete distribution carried by a finite number of *atoms*

$$\omega^k = (\omega_1^k, \omega_2^k, \dots, \omega_{T-1}^k), k = 1, \dots, K$$

hence, to replace the conditional probability distributions  $P_t(\cdot|\omega^{t-1}, \bullet)$  and the marginal distributions  $P_t$  of  $\omega_t \forall t$  by discrete distributions whose supports are **finite**.

**Arc or transition probabilities** are  $P(\omega_1)$  and  $P(\omega_t|\omega^{t-1}, \bullet) \forall t > 1$

**Path probabilities**

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

and probability of scenario  $\omega^k = (\omega_1^k, \dots, \omega_{T-1}^k)$  is

$$p_k = P(\omega^k) = P(\omega_1^k) \prod_{t=2}^{T-1} P(\omega_t^k|\omega_1^k, \dots, \omega_{t-1}^k)$$

**This information is organized in SCENARIO TREE:**

THERE IS EXACTLY ONE ANCESTOR of  $\omega_t^k \forall k, t$ , but multiple descendants are allowed.

# Two Special Types of Scenario Tree

- INTERSTAGE INDEPENDENCE – For all stages conditional distribution of  $\omega_\tau$

$$P_\tau(\cdot|\omega^{\tau-1,\bullet}) = P_\tau,$$

the marginal distribution.

- For all stages, supports of conditional distributions  $P_t(\cdot|\omega^{t-1,\bullet})$  are SINGLETONS

→

the tree collapses into FAN of individual scenarios

$$\omega^k = (\omega_1^k, \dots, \omega_{T-1}^k)$$

which occur with probabilities

$$p_k = P(\omega_1^k), k = 1, \dots, K$$

→ MULTIPERIOD TWO-STAGE PROBLEM

# From a Fan of Scenarios to Scenario Tree

ASSUME:

- ∃ a given structure of the scenario tree, i.e.
  - horizon
  - time discretization
  - stages
  - branching scheme
- ∃ sufficiently many scenarios

## Various ways to create scenario tree

- Ad hoc / expert cutting and pasting
- Conditional / importance sampling
- Clustering
- Moments fitting
- Techniques for scenario tree construction by minimization of distances of probability distributions
- Discretization schemes used instead of simulation or sampling

# Problem Oriented Requirements

One should respect **problem specific requirements** and to avoid as much as possible distortions of available input information. Motivation comes from various problem areas.

- Goal of this procedure does not reduce to approximation of probability distribution  $P$  but to creating input which provides applicable solutions of real-life problem.
- Scenarios based solely on past observations may ignore possible time trends or exogenous knowledge or expectations of the user.
- In financial applications, one prefers that scenario-based estimates of future asset prices in portfolio optimization model do not allow *arbitrage opportunities*; this may put additional requirements on scenario selection.

Explicitly formulated additional requirements concerning properties of probability distribution may help. They can be made concrete through a suitable massaging of data to obtain **prescribed moments values** (given a fixed tree structure. )

SUGGESTION: build scenario tree so that some of statistical properties, e.g. some moments, of data process are retained

# Why Matching Moments?

∃ THEORY: representation of probability distributions by (infinite) sequences of moments and approximating them using only a few moments. Moreover, given  $m$  admissible values of moments, there exists *discrete* probability distribution with these moments and its support has at most  $m + 2$  points.

For our purposes it means that for given values of certain moments or expectations of continuous functions, say,

$$\mu_k = \int_{\Omega} g_k(\omega) P(d\omega), k = 1, \dots, m,$$

∃ modest number of scenarios  $\omega^s, s = 1, \dots, S$ , and their probabilities  $p^s, s = 1, \dots, S, \sum_s p^s = 1$  so that the moment values are retained, i.e.,

$$\sum_s p^s g_k(\omega^s) = \mu_k, k = 1, \dots, m. \quad (6)$$

To get scenarios and their probabilities means to find a solution  $\omega^s$  and  $p^s \geq 0, s = 1, \dots, S$ , of system (6) extended for additional constraint  $\sum_s p^s = 1$ . This is a highly nonlinear numerical problem.

System of equations (6) can be further extended for other constraints on selection of scenarios to represent certain strata, to cover extremal cases



# Fitting Moment Values

For simplicity assume that  $\omega = (\omega_1, \omega_2)$  is two-dimensional random vector with prescribed first three moments  $\mu_k(1), \mu_k(2), k = 1, 2, 3$  of marginal probability distributions and with covariance  $\rho$  of their joint probability distribution. To cover an important extremal case, we require that for at least one scenario,  $\omega_1^s \geq l_1, \omega_2^s \geq l_2$  holds true.

We want to get discrete two-dimensional probability distribution carried by  $S$  atoms  $\omega^s = (\omega_1^s, \omega_2^s), s = 1, \dots, S$ , which matches the true one. Hence, we search values of pairs  $(\omega_1^s, \omega_2^s)$  and scalars  $p^s$  such that

$$\sum_{s=1}^S p^s (\omega_1^s)^k = \mu_k(1) \text{ for } k = 1, 2, 3$$

$$\sum_{s=1}^S p^s (\omega_2^s)^k = \mu_k(2) \text{ for } k = 1, 2, 3$$

$$\sum_{s=1}^S p^s (\omega_1^s - \mu_1(1))(\omega_2^s - \mu_1(2)) = \rho$$

$$\omega_1^s \geq l_1, \omega_2^s \geq l_2, p^s \geq 0, s = 1, \dots, S, \sum_{s=1}^S p^s = 1.$$

# Goal Programming Technique

For small number of scenarios or for inconsistent moment values existence of solution is not guaranteed. Almost feasible solution can be found by *goal programming technique*:

scenarios  $\omega^s$  and probabilities  $p^s$  can be obtained for instance by solving weighted least squares minimization problem

minimize

$$\sum_{k=1}^3 \alpha_k \left( \sum_{s=1}^S p^s (\omega_1^s)^k - \mu_k(1) \right)^2 + \sum_{k=1}^3 \beta_k \left( \sum_{s=1}^S p^s (\omega_2^s)^k - \mu_k(2) \right)^2 + \gamma \left( \sum_{s=1}^S p^s (\omega_1^s - \mu_1(1)) (\omega_2^s - \mu_1(2)) - \rho \right)^2$$

subject to

$$\omega_1^1 \geq h_1, \omega_2^1 \geq h_2$$
$$p^s \geq 0, s = 1, \dots, S, \sum_s p^s = 1.$$

From the optimization point of view, it is non-convex problem and may have many local minima.

☐ SOFTWARE

Advantage of this formulation is that the optimal value is zero if the data is consistent and  $S$  is large enough, but that its optimal solution is also a good representation of data in the case of inconsistency. Parameters  $\alpha, \beta$  and  $\gamma$  can be used to reflect importance and quality of data.

Inconsistency can appear if the information about moments comes from different sources, if implicit specifications are inconsistent with explicit ones, etc. Consider for instance a problem which covers two periods. Let us specify the variance of  $\omega_1$  and the variance of the sum  $\omega_1 + \omega_2$ . This is reasonable as many users have difficulties to provide conditional statements about second period variances unless these are equal for all periods. But specifying these two variances, we have said something about the correlation over time. If we now explicitly specify correlations over time, we are likely up with two inconsistent specifications of the same entity.

There is a numerical evidence in favor of performance of stochastic programs based on scenario trees with moment values fitted at each node over those based only on a few randomly sampled realizations.

# Structure of Portfolio Optimization Models I.

Assume now that uncertainty is described by discrete probability distribution of random parameters carried by finite number of scenarios with prescribed probabilities and that this discrete probability distribution is an acceptable substitute of the true underlying probability distribution.

Denote coefficients and decision variables related with scenario  $\omega^s$  simply by superscript  $s$ .

**Fundamental investment decision:** selection of asset *categories* and wealth allocation over time. Level of aggregation depends on investor's circumstances.

**Planning horizon** at which the outcome gets evaluated is endpoint  $T_0$  of interval  $[0, T_0]$  which is further discretized, covered by nonoverlapping time intervals indexed by  $t = 1, \dots, \tau$ .

**Initial portfolio** is constructed at time 0, i.e., at the beginning of the 1st 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  $T_0$ .

# Structure of Portfolio Optimization Models II.

In our general setting of the  $T$ -stage stochastic programs,  $\tau = T$ . 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.

## **Stages do not necessarily correspond to time periods.**

Main interest lies in 1st-stage decisions — all decisions that have to be selected *before* 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.

2nd-stage decisions are allowed to adapt to additional information available at the end of 1st-stage period, etc.

# Decision Variables and Constraints I.

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*; initial holding is  $h_i(0)$ . It can be included into model as

- amount of money invested in  $i$  at the beginning of time period  $t$  or expressed in dollars of the initial purchase price,
- in face values, in number of securities or in lots, etc.

Accordingly, in the first case, the value of holdings at the end of period  $t$  may be affected by 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 symmetry in 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.

## Decision Variables and Constraints II.

The *flow balance constraint* for each asset category (except for cash, the asset indexed by  $i = 0$ ), scenario and time period is

$$h_i^s(t) = (1 + r_i^s(t-1))h_i^s(t-1) + b_i^s(t) - s_i^s(t). \quad (7)$$

It restricts cash flows to be consistent.

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

$$\begin{aligned} h_0^s(t) = & h_0^s(t-1)(1 + r_0^s(t-1)) + c^s(t) + \sum_i s_i^s(t)(1 - \alpha_i) \\ & - \sum_i b_i^s(t)(1 + \alpha_i) + \sum_i f_i^s(t)h_i^s(t) \\ & - y^{s-}(t-1)(1 + \delta^s(t-1)) - L^s(t) + y^{s-}(t) \end{aligned} \quad (8)$$

with  $f_i^s(t)$  — cash flow generated by holding one unit of asset  $i$  during period  $t$  (coupons, dividends, etc.) under scenario  $s$  and  $L^s(t)$  paydown of committed liabilities in period  $t$  under scenario  $s$ . We denote  $y^{s-}(t)$  borrowing in period  $t$  under scenario  $s$  at 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  $s$ .

## Decision Variables and Constraints III.

For holdings, purchases and sales expressed in *numbers or in face values*, cash balance equation contains purchasing and selling prices,

$\xi_i^s(t) > \zeta_i^s(t)$ :

$$\begin{aligned} h_0^s(t) = & h_0^s(t-1)(1+r_0^s(t-1)) + c^s(t) + \sum_i \zeta_i^s(t)s_i^s(t) \\ & - \sum_i \xi_i^s(t)b_i^s(t) + \sum_i f_i^s(t)h_i^s(t) \\ & - y^{s-}(t-1)(1+\delta^s(t-1)) - L^s(t) + y^{s-}(t) \end{aligned} \quad (9)$$

and flow balance constraints for assets assume a simpler form

$$h_i^s(t) = h_i^s(t-1) + b_i^s(t) - s_i^s(t) \quad (10)$$

as *no wealth accumulation is considered*.

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 diversification, limit investments in risky or illiquid asset classes, limit borrowings, loan principal payments and turnovers, reflect legal and institutional constraints, or fixed-mix policy  $h_i^s(t) = q_i \sum_{j=0}^I h_j^s(t) \forall i$ ,  $q_i$  fixed.



# Objective Function

**Random liabilities**  $L^s(t)$  belong to model input, various *decisions* concerning other liabilities can be included in the external cash flows and 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, etc. Cash balance equation has to take into account cost of debt service.

**Objective function** is mostly related to the *wealth at the end of planning horizon*  $T_0$ ;

This for each scenario consists of amount of 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 choice of suitable utility function or risk measure and incorporated into objective function or into constraints. Perspective alternative is to examine utility functions of *several* outcomes at specific time instants covered by the model. Also criteria *nonlinear* in probability distributions can be applied; the Markowitz model is an example.

# Solving Portfolio Optimization Problems I.

To **initiate the model**, one uses scenarios  $r_i^s(t), \delta^s(t), f_i^s(t), L^s(t)$  of returns, interest rates and liabilities for all  $t$  and starts with 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*.

# Solving Portfolio Optimization Problems II.

For *multistage stochastic programs*, input is mostly in the form of **scenario tree** and nonanticipativity constraints on decisions enter implicitly by using a decision tree which follows the structure of the already designed scenario tree, or in explicit way by forcing decisions based on the same history (i.e., on identical part  $\omega^{t,\bullet}$  of several scenarios) to be equal.

With explicit inclusion of nonanticipativity constraints, scenario-based multiperiod and multistage stochastic programs with linear constraints can be written in a form of **large-scale deterministic program**

$$\min_{\mathcal{X} \cap \mathcal{C}} \left\{ \sum_s p^s u^s(\mathbf{x}^s) \mid \mathbf{A}^s \mathbf{x}^s = \mathbf{b}^s, s = 1, \dots, S \right\}$$

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

## Solving Portfolio Optimization Problems III.

A large class of solvers (CPLEX, MSLiP-OSL, OSL-SP, etc.) are currently available for solution of multistage problems with linear constraints and convex 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, e.g., SLP-IOR, which are able to manage efficiently large scale scenario-based stochastic linear programs for portfolio optimization including those maximizing expectations of piece-wise linear concave utility functions.

# Yasuda Kasai's Problem

1990 – Reason for developing new method according to Kunihiro Sasamoto, director and deputy president of Yasuda Kasai

*The liability structure of the property and casualty insurance business has become very complex, and the insurance industry has various restrictions in terms of asset management. We concluded that existing models, such as Markowitz mean-variance, would not function well and that we need to develop a new asset/liability management model.*

## Requirements of RYK model:

- Adequately represent the book and market value goals
- Incorporate regulations
- Reflect multiple, conflicting goals (maximize long-run value of company and high-quality service to customers)
- Capture multiperiod nature of goals and constraints
- Reflect uncertainty of investment process, financial markets, liabilities (claims)
- Solve quickly using given computer technology
- Be believable and understandable by managers

## ASSET-LIABILITY MANAGEMENT (ALM) PROBLEM with

- New types of liabilities (savings-oriented policies)
- Complex constraints mostly by Ministry of Finance
- Multiperiod risks and objectives – maximize both the long-run total return and current yields

## PREVIOUS APPROACH

Mean-Variance Markowitz model applied repeatedly in time, expectations and variance matrix computed from scenarios

Cannot capture many NEW FEATURES, such as

- General asset classes and restrictions
- Inclusion of liability balances and cashflows

# Main Formulation Blocks

- Horizon  $T$  and decision stages, period length need not be uniform
- Liabilities
- Asset classes, including illiquid loan assets
- Asset accumulation
- Shortfalls — Income yield earned in a year should exceed interest credited, otherwise recourse action (penalty)
- Objective function — maximize expected return at the horizon minus expected penalty for shortfalls at each of considered time point
- Scenarios

*Trade off between detail and complexity* is needed, e.g. choice of

- manageable No.of asset classes, possibly different asset classes in different periods

- number and allocation of stages within the chosen horizon

- number of scenarios to represent uncertainty in interest rates, returns, prices, etc.

INSERT Scheme of stages and scenario tree

# Main Formulation Blocks

- **Liabilities** Liability growth flows

$$L_{t+1} = (1 + g_{t+1})L_t + F_{t+1} - P_{t+1} - I_{t+1}$$

$L_t$  liability balance

$F_{t+1}$  deposit inflow

$g_{t+1}$  credited growth

$P_{t+1}, I_{t+1}$  principal resp. interest payout

Generated off-line by aggregating over all policies

- **Asset accumulation**

Denote

$X_{jt}$  market value in asset  $j$  at  $t$

$A_t$  total fund market value at  $t$ , i.e.

$$A_t - \sum_j X_{jt} = 0 \quad (11)$$

$\varrho_{jt+1}$  asset  $j$  return from end of  $t$  to end of  $t + 1$

(its part  $\varrho_{jt+1}^I$  denotes the income return)

$$A_{t+1} = \sum_j (1 + \varrho_{jt+1})X_{jt} + F_{t+1} - P_{t+1} - I_{t+1} \quad (12)$$



# Model Skeleton

$$\text{Maximize } E[A_T - \sum_{t=1}^T c_t w_t]$$

subject to budget and asset accumulation constraints (11), (12) with income shortfall definitions

$$\sum_j \varrho_{jt+1}' X_{jt} + w_{t+1} \geq g_{t+1} L_t$$

and nonnegativity of all variables  $X_{jt}, w_t$   
for all  $t = 0, \dots, T - 1$

The income shortfall in period  $t$  corresponds to positive value  $w_t$  and is penalized by (possibly discounted cost of borrowing)  $c_t$  in the objective function.

Other constraints, other types of shortfalls can be included

# Requirements on Implementation

- Software was to be delivered to the client
- Client staff would create inputs and run the model repeatedly each quarter
- Specified hardware platform

Applied since 1991 — New potfolio produced higher income with no loss in total return

Compared with previous technique — Markowitz model with comparable input data

Tested over single period and over multiple periods

Poor results for multiple periods

## **IMPULS TO SP APPLICATIONS IN VARIOUS ALM PROBLEMS**

THE MAIN AREA IS ALM FOR PENSION FUNDS

The common feature: Models are problem and country specific, take into account details and requirements of users.

Reflection of RISK is one of the major issues in modeling ALM.