# ON THE EFFICIENCY OF SOLUTIONS OF STOCHASTIC OPTIMAL CONTROL PROBLEM WITH DESCRETE TIME

Iouldouz S. Raguimov
Department of Mathematics and Statistics,
York University
4700 Keele Street,
Toronto, Ontario, Canada, M3J 1P3
Email: raguimov@mathstat.yorku.ca

The presentation will cover the results of a joint research with Dr. Igor I. Gasanov from the Department of Computational and Information Systems of the Computing Centre of the Russian Academy of Sciences.
E-mail: gasanov@ccas.ru

THE FIELDS INSTITUTE FOR RESEARCH IN MATHEMATICAL SCIENCE
222 College Street, Toronto, Ontario, Canada M5T 3J1

INDUSTRIAL-ACADEMIC WORKSHOP ON OPTIMIZATION IN FINANCE AND RISK MANAGEMENT
October 3-4, 2011

*"All the processes in the nature are influenced by surrounding stochastic factors with enormously large level of uncertainty. Same can be said about society!"*

*N.N.Moiseev*

n  Often a problem of control in economic and financial systems is formulated as a problem of decision-making under stochastics and/or uncertainty and formalization of such a problem usually requires a non-trivial non-classic approach

n  We will consider this as Operations Research problem where the results of an operation do depend not only on activities of the operating side – in an operation there exist some uncontrollable factors, that are having influence on the results of the operation, create so-called operational situation
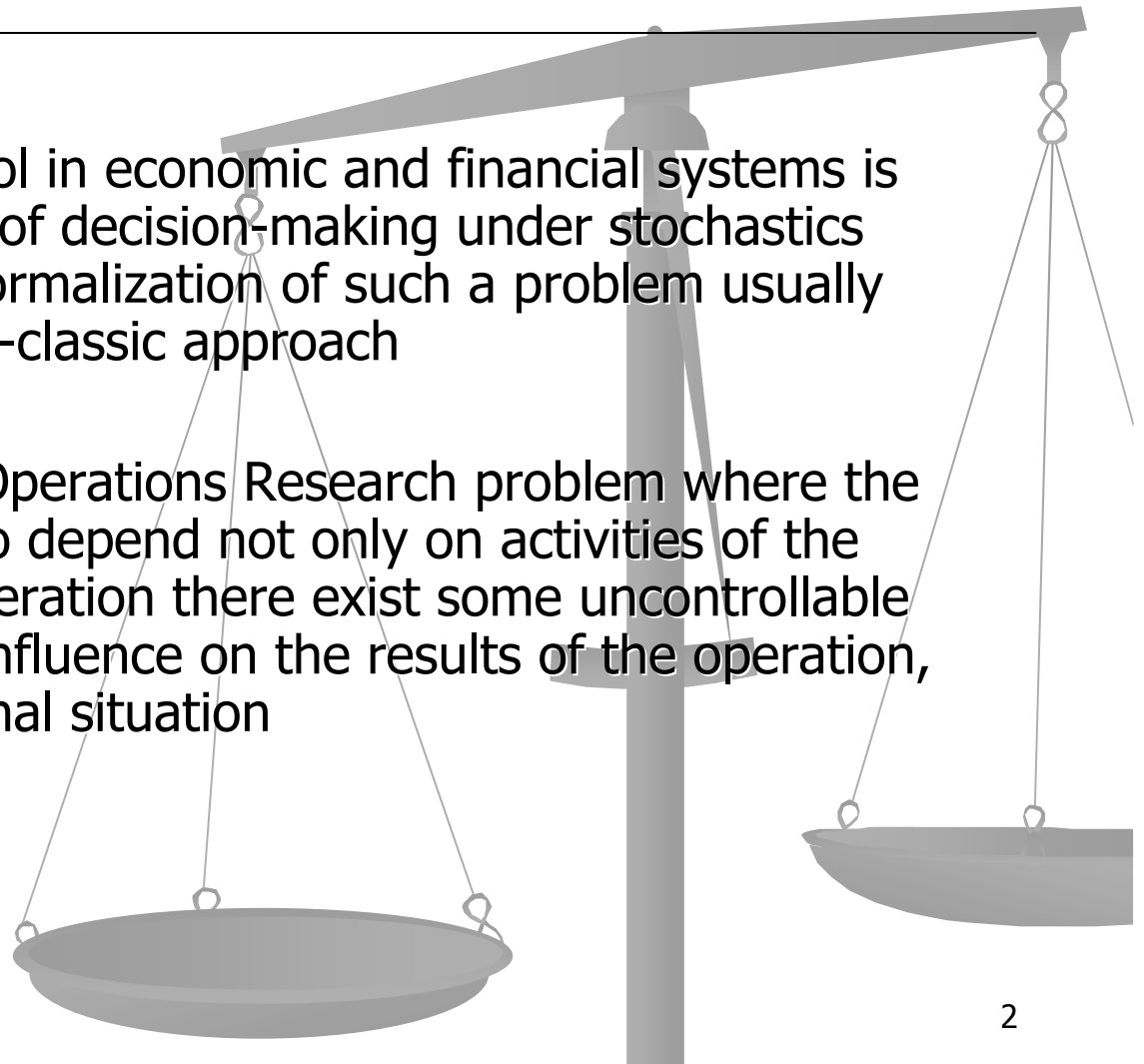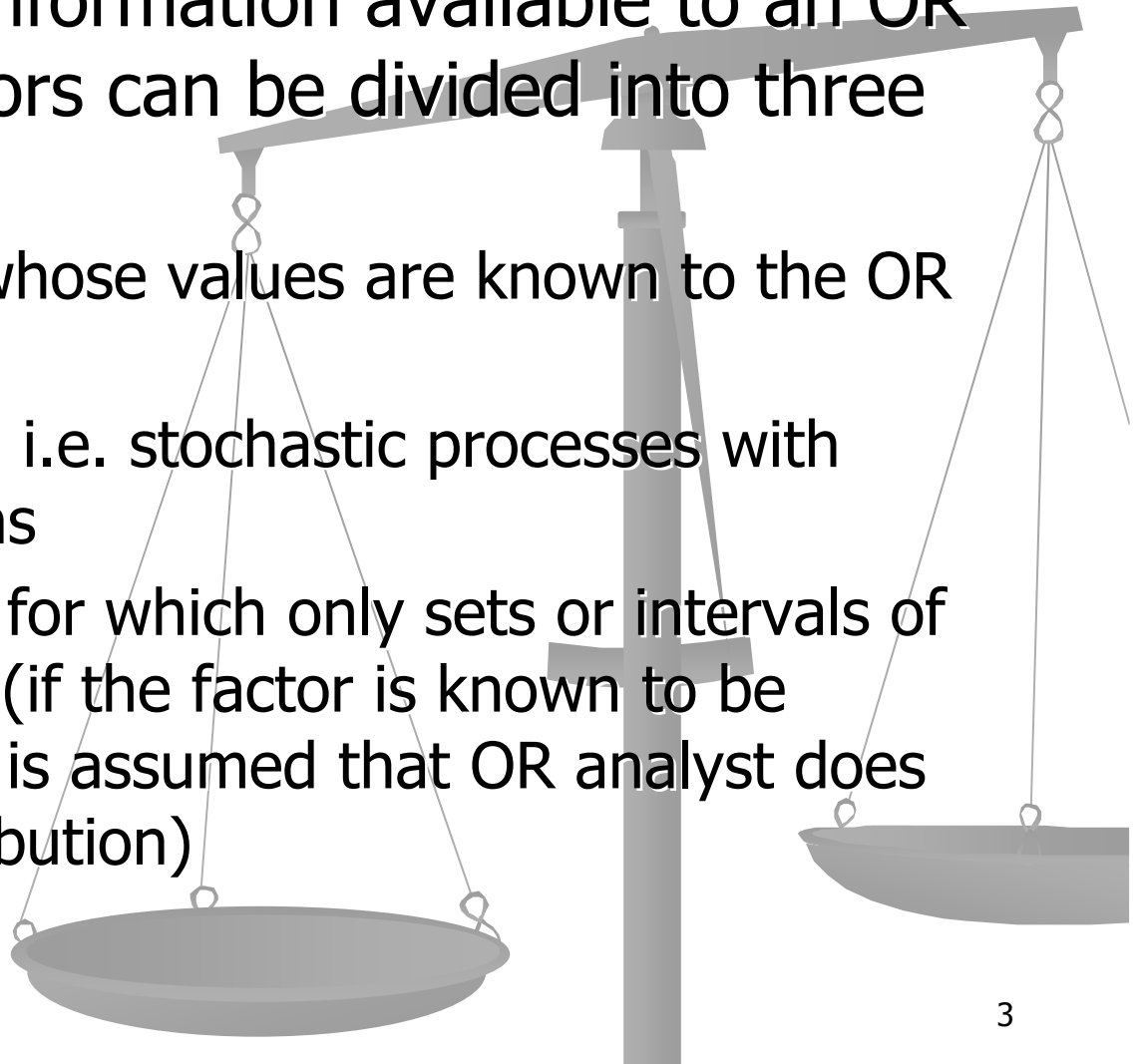
# Introduction

n According to the information available to an OR analyst these factors can be divided into three groups:

Constant factors whose values are known to the OR analyst

Stochastic factors, i.e. stochastic processes with known distributions

Uncertain factors, for which only sets or intervals of values are known (if the factor is known to be stochastic, then it is assumed that OR analyst does not know its distribution)

n   In situations, when it is either impossible or unaffordable to determine an optimal solution to such a problem exactly, so-called method of *optimization on time series* is often used to determine an efficient solution to the problem:

Relying on information about realizations of uncontrollable uncertain parameter, we tend to determine a control that would be optimal once were to be used in the past

Here, we assume implicitly that since the uncertainty has a regular character, then if a control would have been optimal during some sufficiently continuous period of time in the past, it will be optimal in the future also.
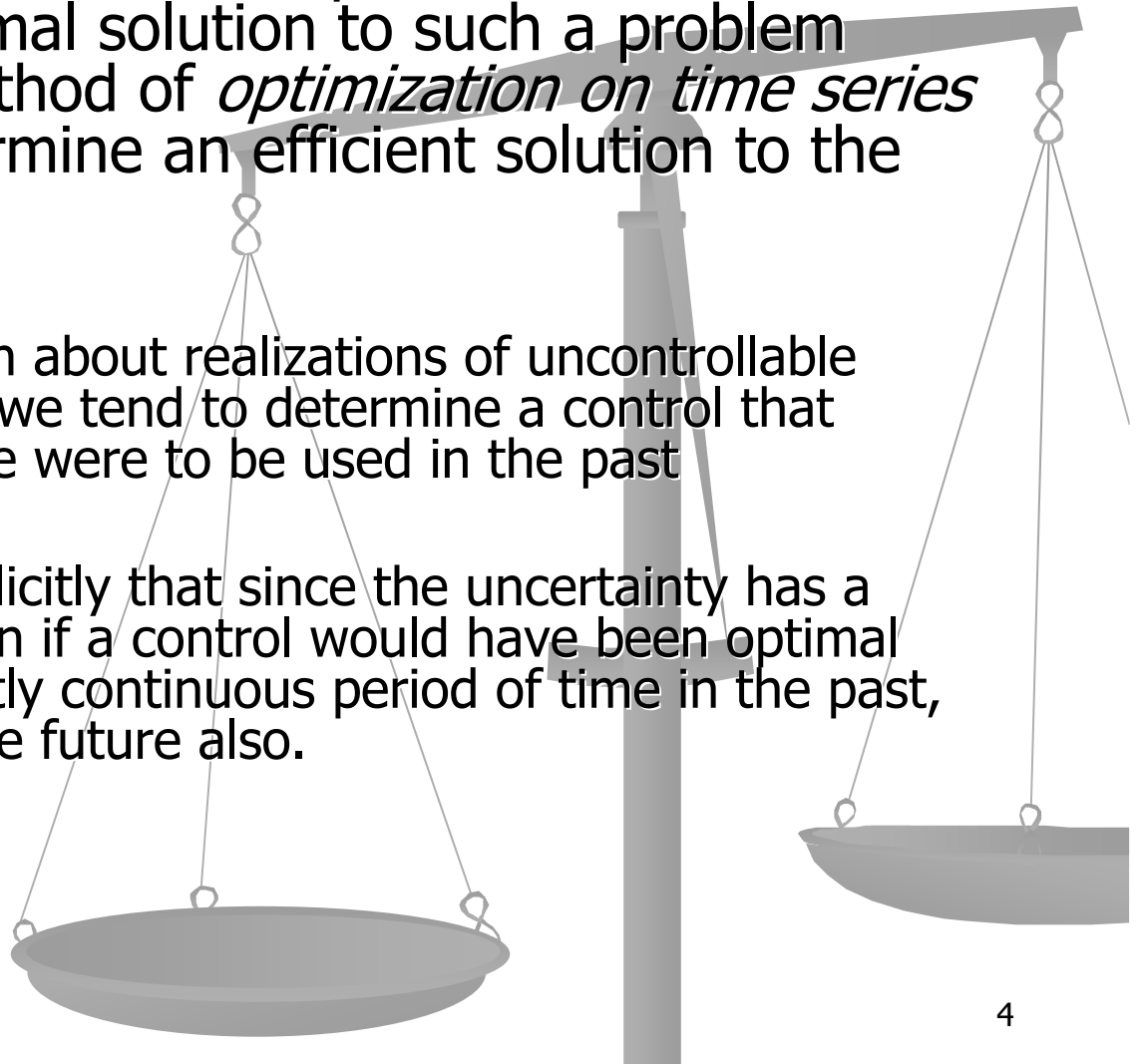
# Introduction (cont'd)

n The abovementioned idea appears to be rational, especially since the necessity of making decisions in such operations arises regularly, and there is no alternative approach to solving this type of decision-making problems.

n Questions and doubts:

Since optimal controls are determined and estimated on the same set of realizations of a stochastic parameter, by constructing an optimal control using time series, to what extend do we exploit systematic properties of the stochastic process, and to what – just make adjustments utilizing only non-significant for the future properties of the stochastic process?
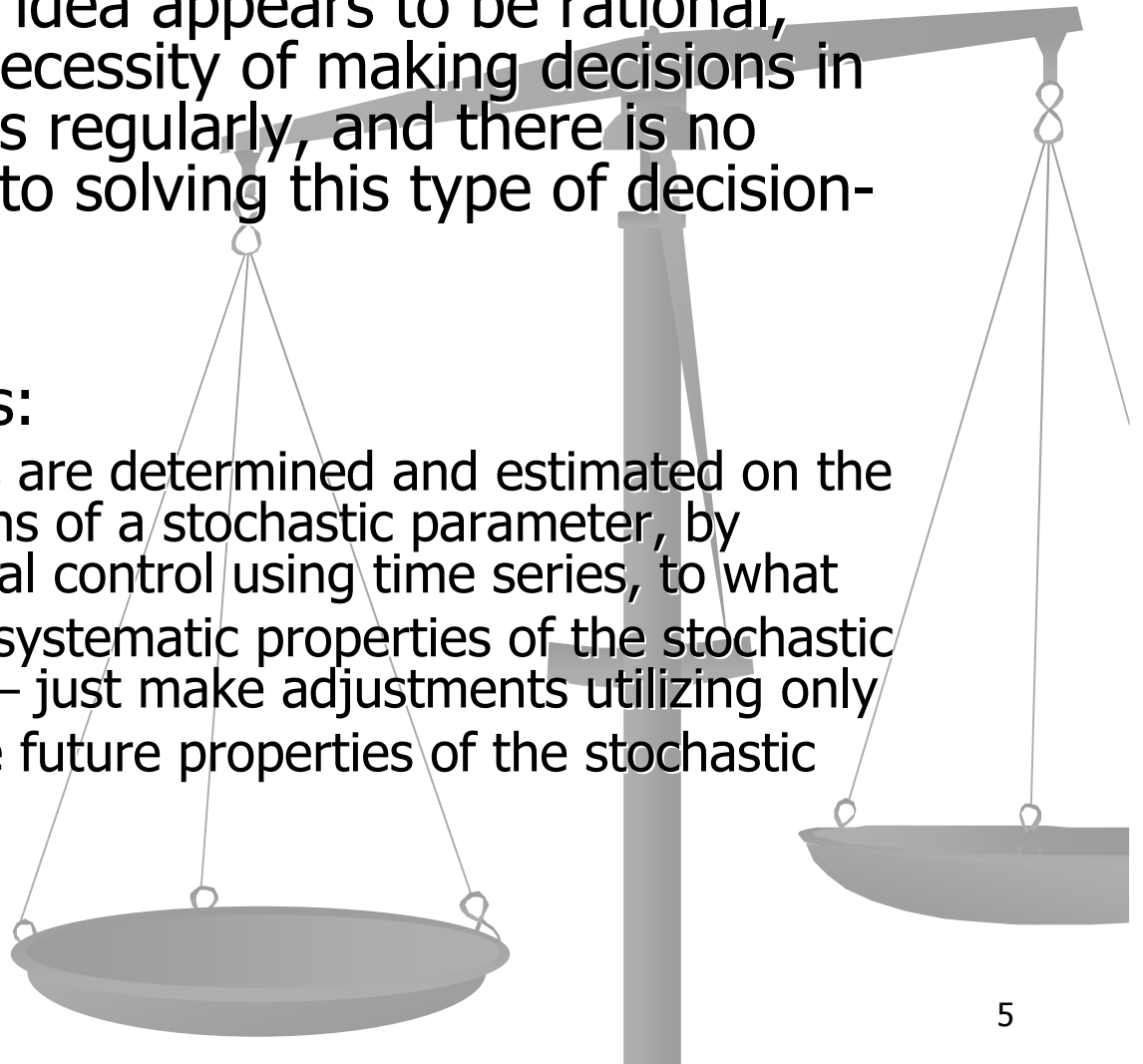
# Introduction (cont'd)

n  When an uncontrollable uncertain factor has a stochastic character, it usually represents a multidimensional stochastic process for which neither a structure nor parameters are known.

n  Moreover, to determine the structure and/or to calibrate parameters of a stochastic process an OR analyst is as a rule, experiencing a data insufficiency. Even when the model of a stochastic process has been formulated, obtained optimization problem is usually too complicated to be solved analytically.
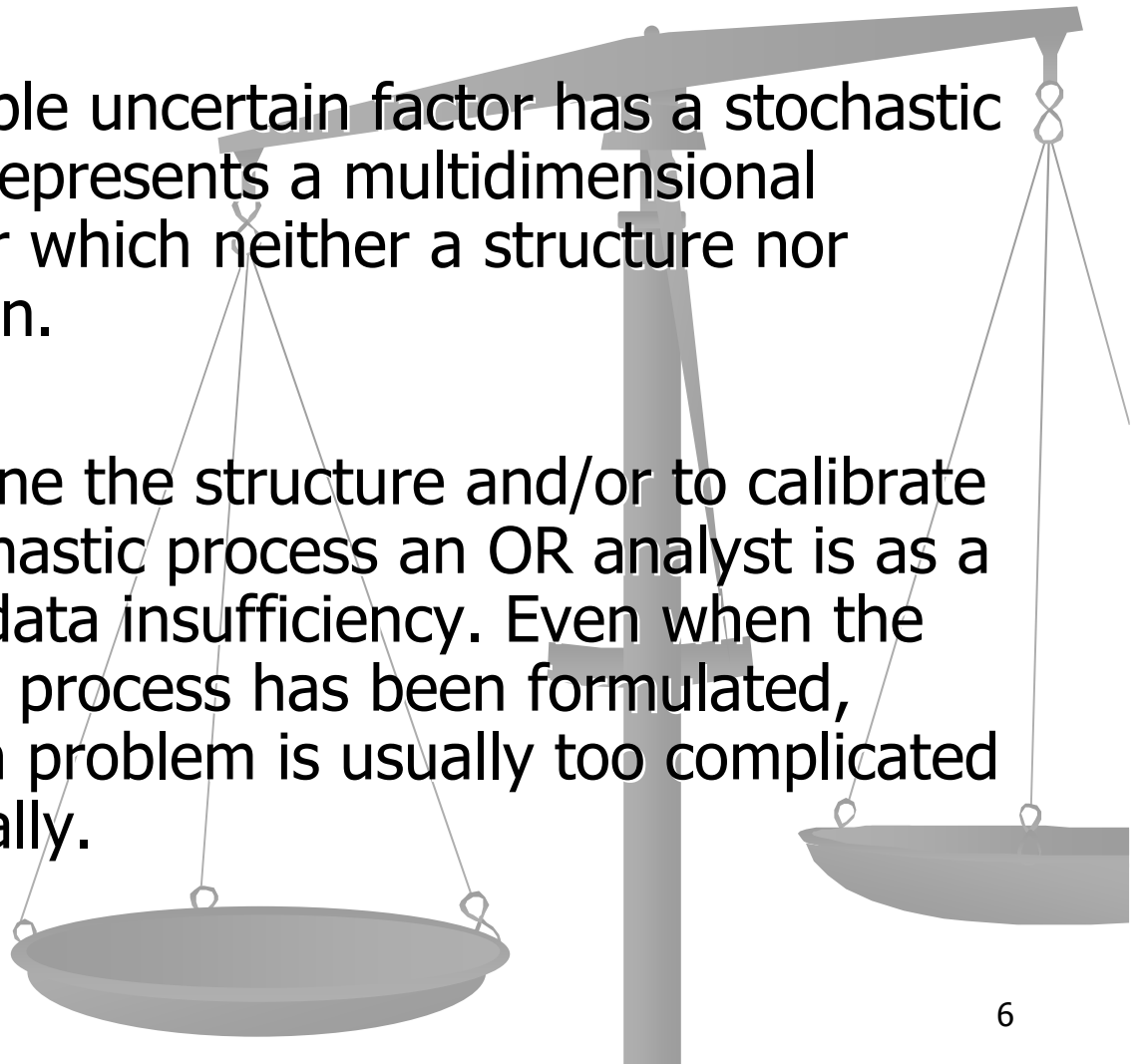
# Parametric Optimal Control Problem with Discrete Time

- n Suppose that at time (stage) $t$ = 1, 2, 3, … the state of a system is given by the characteristic vector $A_t \in \hat{A} \subset R^n$

- n Once a control $u_t$ at the stage (time period) $t$ is chosen and the value $\tilde{\xi}_t$ of the stochastic parameter $\xi$ is realized, the system moves on from the state $A_t$ to the state

$$A_{t+1} = \varphi(A_t, \tilde{\xi}_t, u_t)$$

- n where the parameter $\xi \in \Xi \subset R^m$ is a stationary Markov process with the transition probability function $\Phi(\xi_t \mid \xi_{t-1})$

- n Every ordered pair $S_t = (A_t, \xi_t)$ of arguments of the function $\varphi$ determines a state of the system at stage $t$.

- n The probability distribution $F^1(S_1)$ on the set of initial states $S_1$ of the system, i.e. the initial probability distribution is given.

7
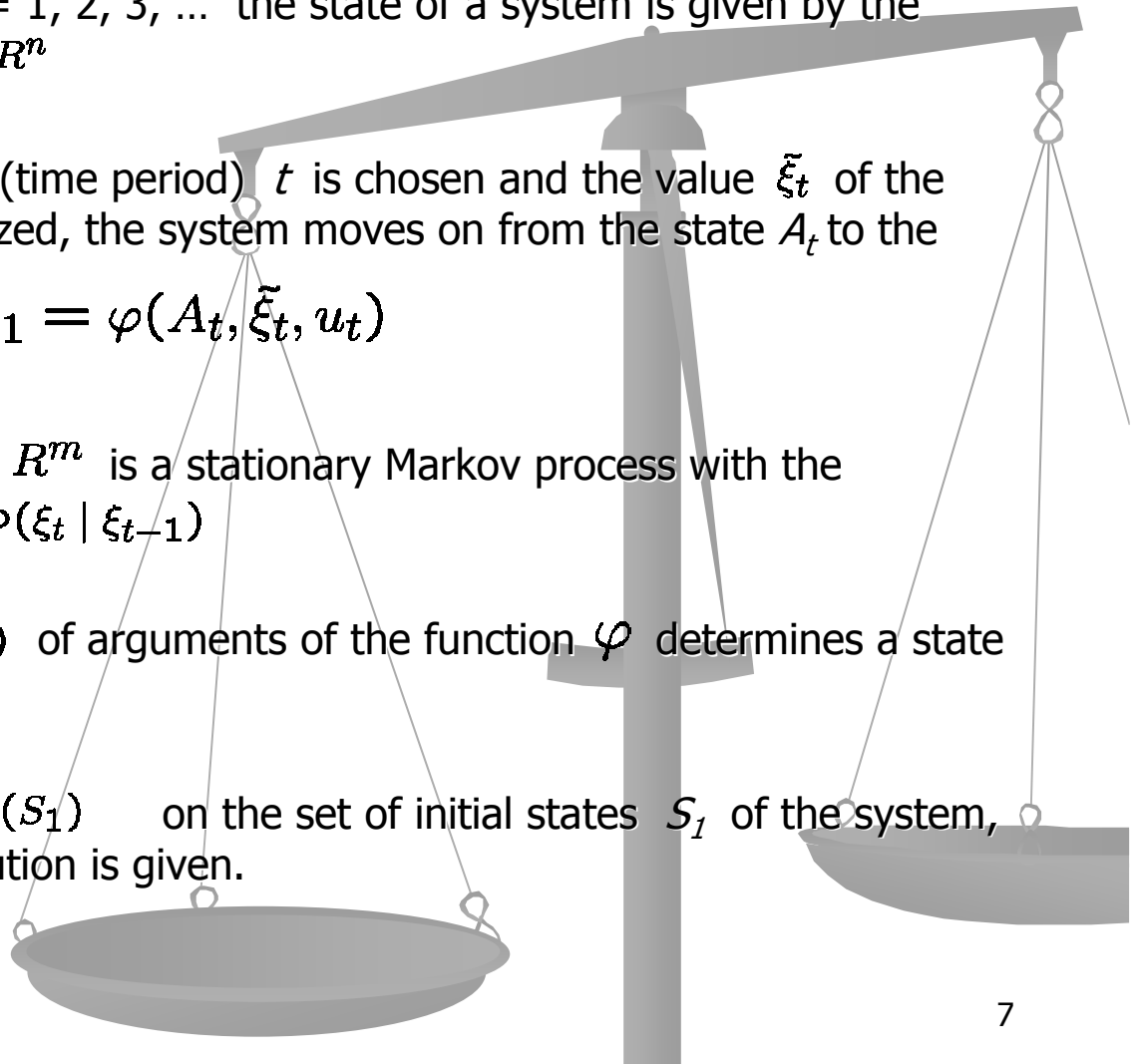
n  Every stage $t$ of the process is associated with a certain payoff function (expected reward)

$$h_t = h_t(S_t, u_t)$$

n  Suppose that the decision maker is interested in the maximization of the average reward earned per period, i.e. s/he is solving the following maximization problem

$$Q = \lim_{n \to \infty} \frac{1}{n} E(\Sigma_{i=1}^{n} h_t(S_t, u_t)) \implies \max_{u} \quad (1)$$

n  Here, generally speaking, the decision maker tends to maximize the function (1) with respect to $u_t = u(S_t)$, where $u$ is a mapping

$$\hat{A} \times \Xi \to \hat{U}$$

n  It is clear that for the existence of the expected value in (1) the functions involved in the model should satisfy certain constraints.

n  The application of the method of optimization on time series to parametric optimization problem where a certain parametric class of control functions is considered and a problem is formulated as a problem of choosing optimal values for parameters of the functions from the considered class.

8

- Let's consider the problem of maximization (1) on the set of control functions $\widehat{U}_\alpha$ with $\alpha \in A$, where $\widehat{U}_\alpha$ is a class of the functions $u(S; \alpha)$, such that there exists a one-to-one correspondence between $\widehat{U}_\alpha$ and $A$

- Consequently, the original problem is reduced to a problem of finding a value of $\alpha$, such that

$$Q = \lim_{n \to \infty} \frac{1}{n} E(\Sigma_{t=1}^n h_t(S_t, u(S_t; \alpha))) \implies \max_\alpha \qquad (2)$$

- Denote the formulated problem as <u>Problem 1</u> and compare it with its discrete analogue <u>Problem 1D</u> of maximization of function (2) on $\widehat{U}_\alpha^D$, the parametric class of discrete functions.

- It is assumed that there exists a solution to Problem 1 and we suppose that these two problems are such that when an appropriate (small) mesh for the grid is chosen, solutions to Problem 1D closely approximate solutions to Problem 1. Therefore, determining an approximate solution to Problem 1 is the same as finding an approximate solution to Problem 1D.

# Problem of Parametric Optimization on Time Series

- n Suppose, we are given the sequence of realizations $\{\tilde{\xi}_1, \tilde{\xi}_2, \cdots, \tilde{\xi}_T\} \subset \Xi^D$ of the stochastic parameter $\xi$ , and the initial state $\tilde{A}_1 \in A^D$ of the system.

- n Problem 1R: Maximize the objective function

$$\tilde{Q}^T = \frac{1}{T}\Sigma_{t=1}^T h_t(A_t, \tilde{\xi}_t, u(A_t, \tilde{\xi}; \alpha)) \qquad (3)$$

on the set of control functions $u_t = u(A_t, \tilde{\xi}_t),$ subject to the constraint $A_1 = \tilde{A}_1$

- n Control function $u$ is said to be *everywhere optimal*, if it is optimal for every initial distribution $F^1(S_1)$ .

- n **Theorem 1.** If there exists an everywhere optimal control function for Problem 1D, then the optimal objective value of Problem 1R almost surely converges to the optimal objective value of Problem 1D, provided that the size of the sample increases.

# Optimization of Parameters of Stochastic Process on Time Series

n   In case, when the given data does not allow to construct a reliable model of a stochastic process or Problem 1 is too complicated to be solved in the original form, the stochastic process is replaced with a simple one, which according to the opinion of OR analyst reflects the essential characteristics of the original process.

n   The parameters $\beta \in R^n$ of the auxiliary process are calibrated on the given series of realizations $\bar{\xi}_1, \bar{\xi}_2, \cdots, \bar{\xi}_T$ and a problem with the correspondingly modified stochastic process is solved ( let's call this problem as <u>Problem 1M</u>).

n   Denote the objective value of Problem 1 corresponding to the control function $u$ by $Q(u)$ and consider the parametric class of problems of the type 1M, where as parameters are considered calibrated coefficients of the modified stochastic process.

n   Denote a solution to Problem 1M corresponding to fixed values of the coefficients $\beta$ by $u^\beta$ and the corresponding objective value by $Q(u^\beta)$

- Consider the problem of maximization of $Q(u^\beta)$ on the set of calibrated parameters $\beta$ and denote it as <u>Problem 1A</u>. Let's estimate the parameters of the stochastic process using one of the commonly used statistical methods, namely using Monte-Carlo method, and consider Problem 1M corresponding to the obtained parameters. Denote problem as <u>Problem 1S</u>.

- As before, formulate discrete analogues of the problems 1M, 1A and 1S. Denote these problems as <u>Problem 1MD</u>, <u>Problem 1AD</u> and <u>Problem 1SD</u>, correspondingly.

- The efficiency of a control function obtained by solving Problem 1MD can be estimated on the sample $\bar{\xi}_1, \bar{\xi}_2, \cdots, \bar{\xi}_T$ by calculating the value of objective function (3).

- Now, formulate <u>Problem 1MR</u> as a problem of finding the values of the parameters of the stochastic model that maximize the value of objective function (3).

- Let $u^{1SD}(\bar{\xi}_1, \bar{\xi}_2, \cdots, \bar{\xi}_T)$ and $u^{1MR}(\bar{\xi}_1, \bar{\xi}_2, \cdots, \bar{\xi}_T)$ be optimal control functions for Problem 1SD and Problem 1MR, correspondingly. Using Theorem 1, the following theorem can be proved.

- **Theorem 2.** If there exists $u^{1AD}$ an everywhere optimal control function for Problem 1AD, then the optimal objective value
$$Q(u^{1MR}(\bar{\xi}_1, \bar{\xi}_2, \cdots, \bar{\xi}_T))$$
almost surely converges to the optimal objective value $Q(u^{1AD})$ as $T$ tends to infinity. At the same time, $Q(u^{1SD}(\bar{\xi}_1, \bar{\xi}_2, \cdots, \bar{\xi}_T)) \leq Q(u^{1AD})$ $(\forall(\bar{\xi}_1, \bar{\xi}_2, \cdots, \bar{\xi}_T))$

# Design of Computational Experiments

- The asymptotic preference of one method to another does not provide, however, formal grounds to consider the first method as more efficient in solving practical problems where data samples are always limited and mostly not large enough.

- Mathematical models of controlled systems containing models of a stochastic process of the form of Markov process are constructed where the Markov process is modeling a financial market.

- Using the Markov process, data imitating series of observations is generated, then from the point of view of the operation analyst who does not know the structure of the stochastic process and knows only the series of observations, various problems of optimization on time series are solved.

- Certainly, such experiments cannot be considered as a formal proof of the efficiency of the presented method, nevertheless from the point of view of its further utilization the results of the experiments seem to be essential.
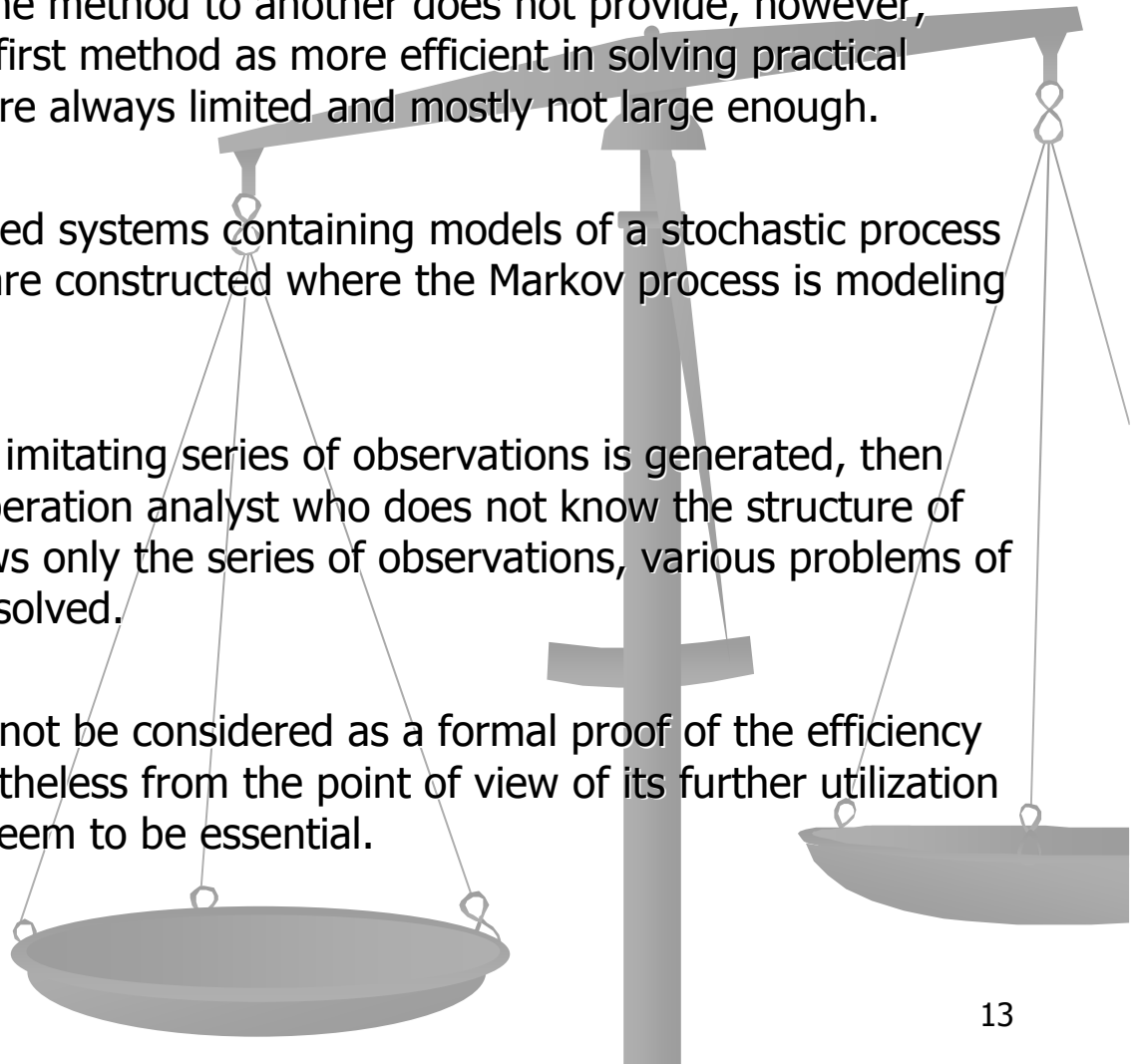
- n    Consider 20 types of securities with different redemption dates but the same redemption value of $100.

- n    Assume that the $k$-th security is redeemed in $T_k = \lfloor \frac{365}{20} \rfloor k$ days, $k = 1, 2, \cdots, 20$; and let $C_{kt}$ be the price of the $k$-th security at the end of the period $t$ (each of these 18-day periods is called a session).

- n    So, $C_{k,18k} = 100$, $k = 1, 2, \cdots, 20$. Hence, the nominal yield rate, $i_{kt}$ of the $k$-th security purchased at $C_{kt}$ at the end of the session $t$ and redeemed in $T_k = 18k$ days is determined as

$$i_{kt} = 100(\frac{100}{C_{kt}} - 1)\frac{20}{k}\%$$

or numerically

$$i_{kt} = (\frac{100}{C_{kt}} - 1)\frac{20}{k},$$

- n    Solving the latter for $C_{kt}$ (the purchase price of the $k$-th security at the end of the session $t$)

$$C_{kt} = \frac{2000}{ki_{kt} + 20}$$

14

- The prices $C_{kt}$ are assumed to be random variables determined according to the following principle: for every session $t$ there exists a basis yield rate $\alpha_t$, a random variable forming a stationary Markov process of the deepness three, i.e. for some constants $b_1^0$, $b_2^0$ and $h^0$

$$\alpha_t = \alpha_{t-1}(1 + h^0\xi_t) + b_1^0(\alpha_{t-1} - \alpha_{t-2}) + b_2^0(\alpha_{t-2} - \alpha_{t-3}) + 0.01(1 - \alpha_{t-1})^3 \quad (6)$$

where $\xi_t$ is a random variable with standard normal distribution.

- The yield rate of the $k$-th security is supposed to be a certain deviation of the current basis yield rate $\alpha_t$ and defined by

$$i_{kt} = \alpha_t(1 + d_{kt}) \quad (7)$$

where $d_{kt}$ is a random variable defined by

$$d_{kt} = b'd_{k,t-1} + h'\xi_{kt} \quad (8)$$

- Given that the $k$-th security is purchased for the price of $C_{kt}$ at the end of the session $t$ and sells at the price of $C_{k,t+1}$ at the end of the session $t+1$, investor's revenue $B_{k,t+1}$ for each unit of investment is determined as

$$B_{k,t+1} = \frac{C_{k,t+1}}{C_{kt}} - 1. \quad (9)$$

- It is supposed that an algorithm of control, that is an algorithm to select the type of a security to be purchased at each session depends on operator's information about the process of forming the yield rates for the securities.

15

n    We consider several scenarios each of which corresponds to the different amount of information being on the possession of the operator.

n    It is assumed that an operation analyst implements and optimizes algorithms of control using series of observations about the process, that is using information about the closing purchase prices of the securities on the market place during a certain time period corresponding to the sessions $t = -2,-1,0, 1, ... , N+1,$ as well as about the values of $\alpha$, whenever it is possible.

n    The sample data is used in such a fashion as the initial investment to be made at the end of session one, and the final operation will take place at the end of session $N$.

n    The result of a portfolio control is estimated at the end of session $N+1$. Computational experiments are used to compare the efficiency of various algorithms of control corresponding to the different information assumed to be in the possession of the operation analyst.

n    The efficiency of a control is estimated by the average relative increment of the cost of investor's portfolio for a session, that is by the average value of the ratio $\frac{D_{t+1}}{D_t}$ where $D_t$ is the cost of the portfolio at session $t$.

16
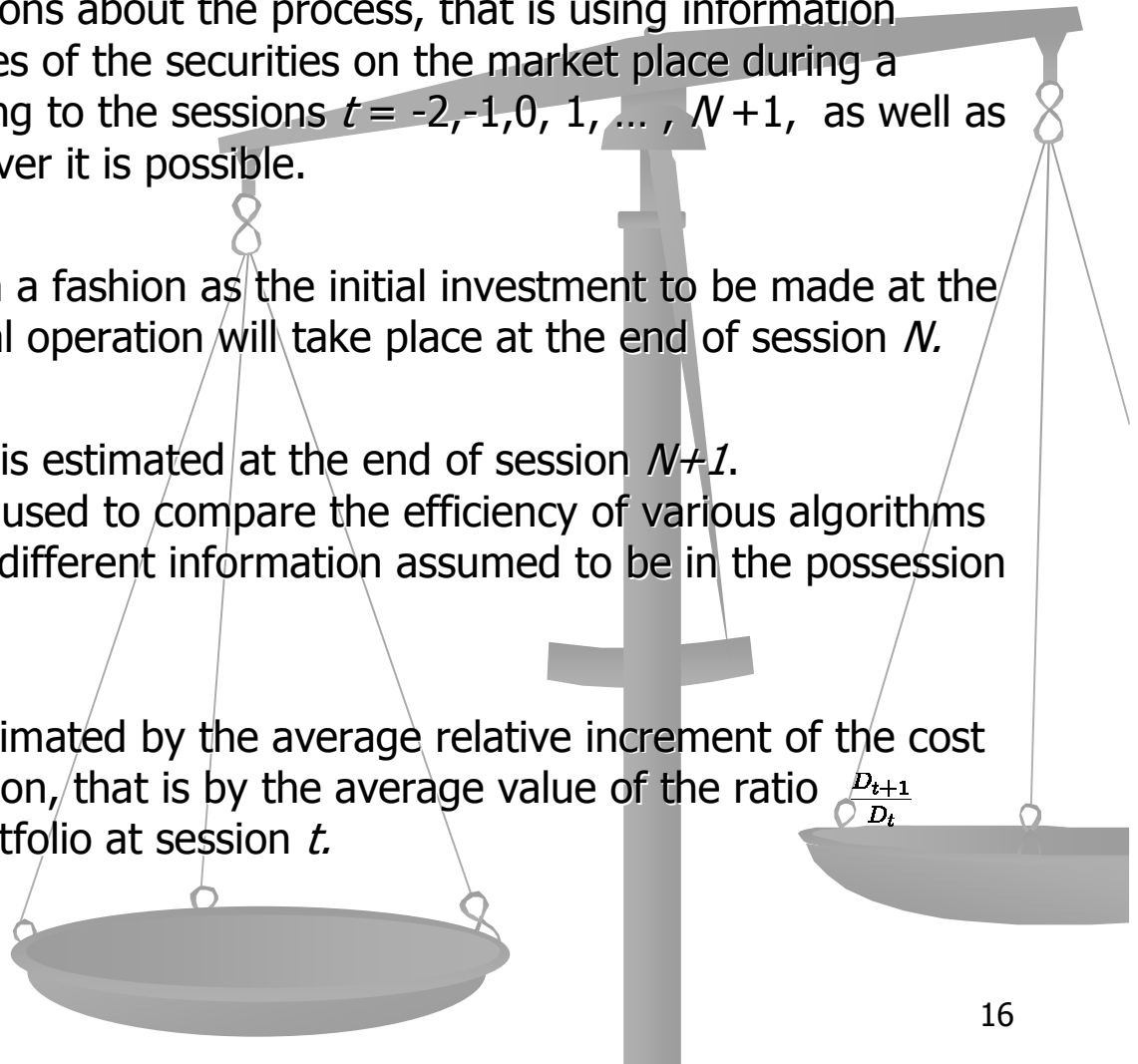
n   The mathematical expectation of the ratio is evaluated using Monte-Carlo method.

n   "Running" the considered control through sufficiently large series of generated observations, that is through a matrix $P$ of the size $[(\overline{N}+4) \times 21]$ , where the columns correspond to the realizations of the values of  $\alpha$  (column one) and

purchase prices  $C_k$ of the securities for sessions (columns 2 to 21)

the number $\overline{N}$  is determined by the power of a used computational tool (it is preferred to have $\overline{N} \geq 10\,000$ ).

n   The expected value of the relative increment in investor's portfolio cost for a session is computed. It is necessary to keep the matrix $P$ the same for all computational experiments ( $P$ is called *proving*  or *testing* matrix ).

n   Existing series of observations is imitated by a generated matrix $X$ of the size $[(N+4) \times 21]$, where the entries have the same meaning as above for $P$.

n   Both matrices $P$ and $X$ are obtained using the same random number generator imitating the random variable $\xi$  with standard normal distribution, and calculating the corresponding entries by formulae (6), (7),(8) and (5).

17
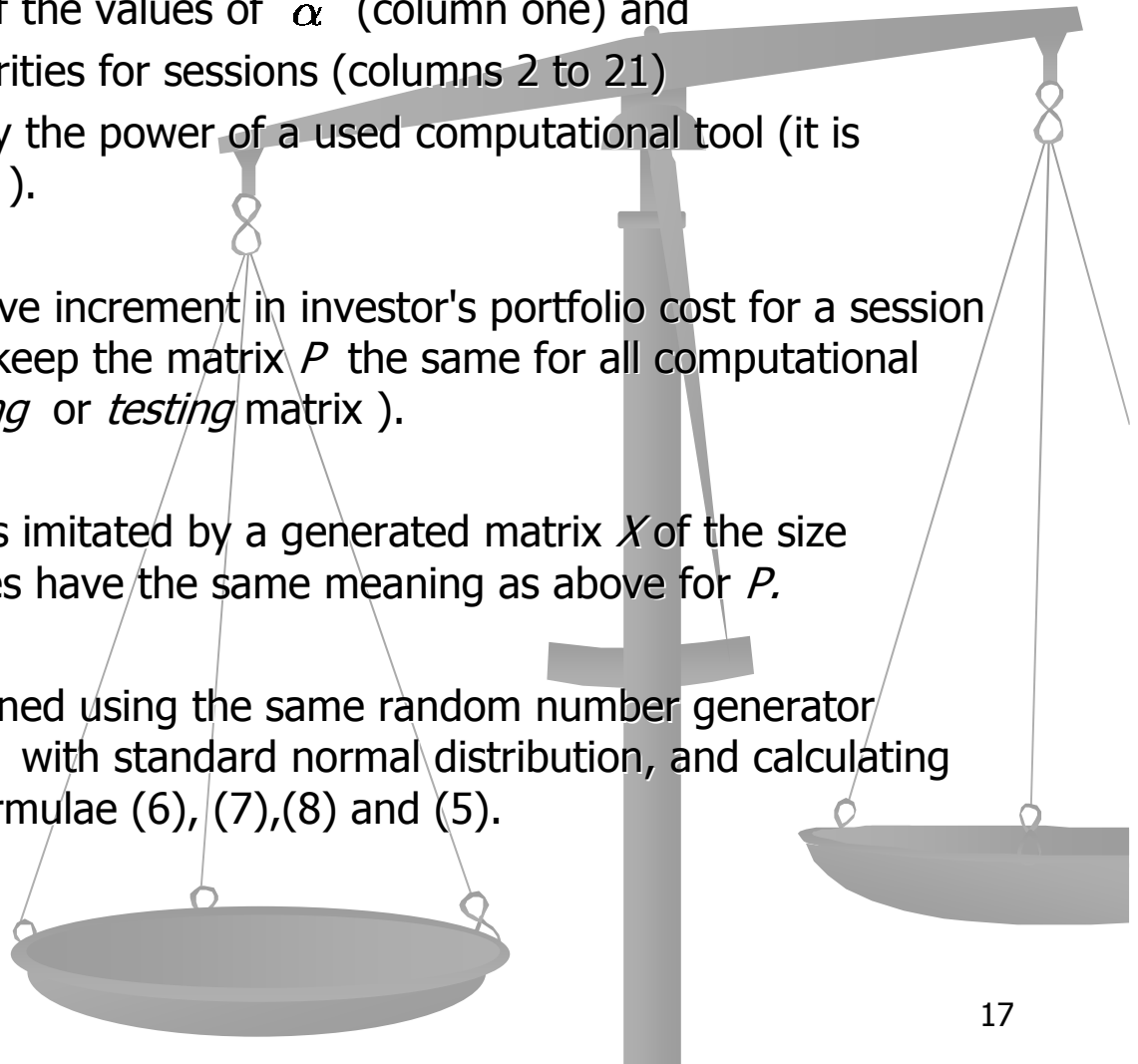
n  If at session one $1 is invested, then the efficiency of a control on series $R$ is calculated as

$$O_R = (D_{N_R+1})^{\frac{1}{N_R}}$$

where $N_R+1$ is the index of the final session on the considered series of observations, and $D_{N_R+1}$ is the cost of the portfolio on that session.

n  The monthly efficiency of the control then calculated as $O_R^M = O_R^{22}$, where the number 22 corresponds to the approximate number of the daily trade sessions taking place within a month.

n  Depending on operation analyst's information about an operational situation the following scenarios are to be considered and the corresponding optimization problems on time series are solved:

(A)  The operation analyst does know the original pricing model including all its parameters and the current value of $\alpha$ at each session exactly. The obtained result will determine the upper bound for an objective value that may be obtained within the considered model.

(B)  The operation analyst knows the structure of the original pricing model, but has no information about the coefficients $b$. The current value of $\alpha$ at each session is known exactly. Based on the generated series of the realizations the operation analyst determines values of the coefficients $b$.

18
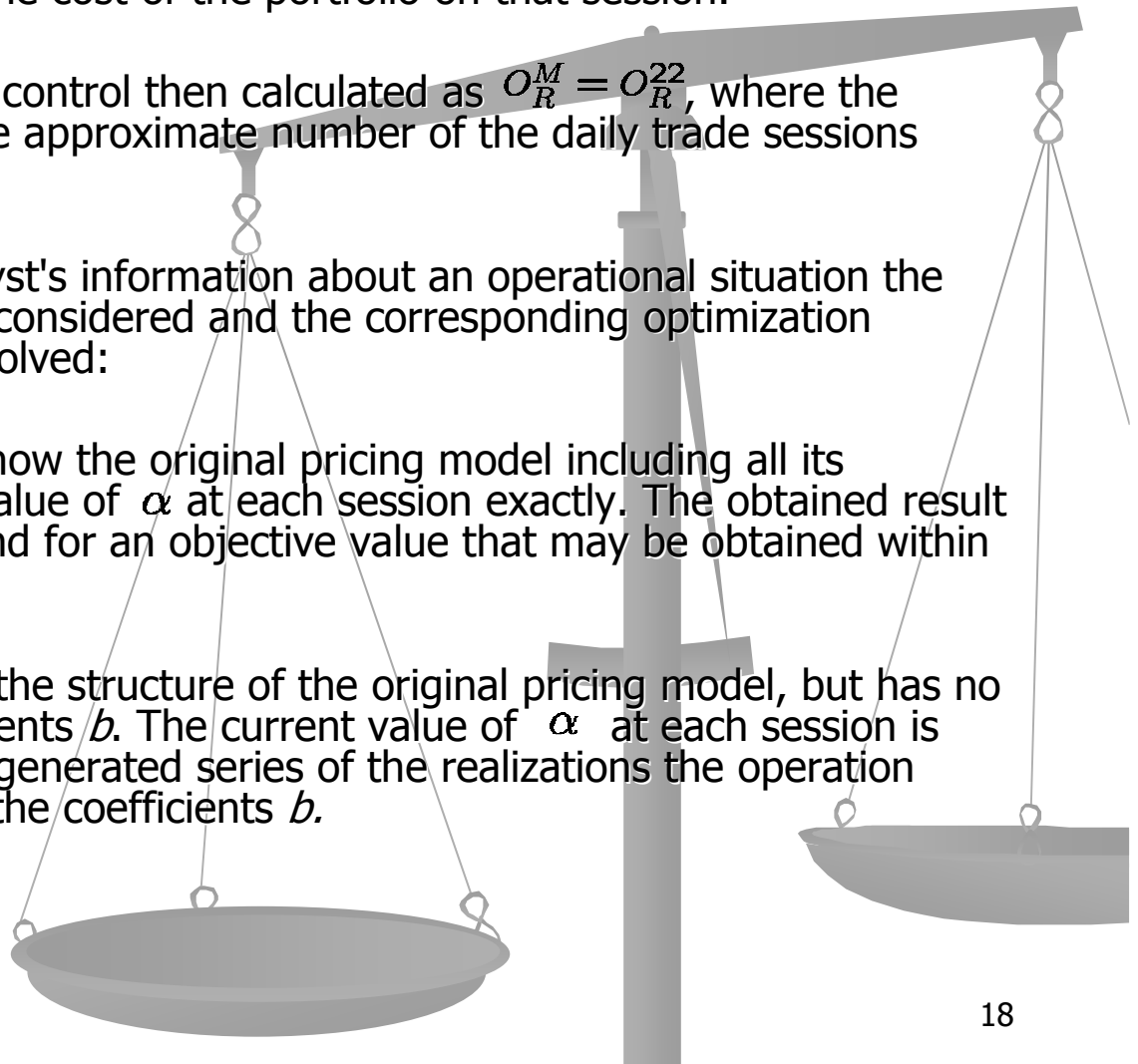
(C) (Independent yield rates) The operation analyst has no information about the structure of the original pricing model; and using a linear regression model, s/he forecasts a closing price of each security separately for the next session assuming that there is no influence between the yield rates of different securities.

(D) (Correlated yield rates) The operation analyst does not know the price forming mechanism for the securities, but s/he makes assumptions about possible influences of the yield rates of different securities to each other.

- Using the graphs of closing prices of the securities from session to session, the operation analyst assumes that at each period of time the points whose $x$-coordinates are the lengths of time to maturity, and $y$-coordinates -- the prices, are grouped about a certain theoretical curve, namely a parabola. Based on data modeling the series of realizations, the operation analyst calculates the coefficients of the parabolas for every past session using the method of least squares.

- Then, the coefficients of the parabolas for the next session are forecasted using the corresponding coefficients for the current and several previous sessions, i.e. based on the obtained so-called "restored" parabolas. At the same time, using the deviations of the prices from the theoretical curve on the current session, the corresponding deviations for the next session is forecasted. For forecasting, as before, a linear regression model is used.
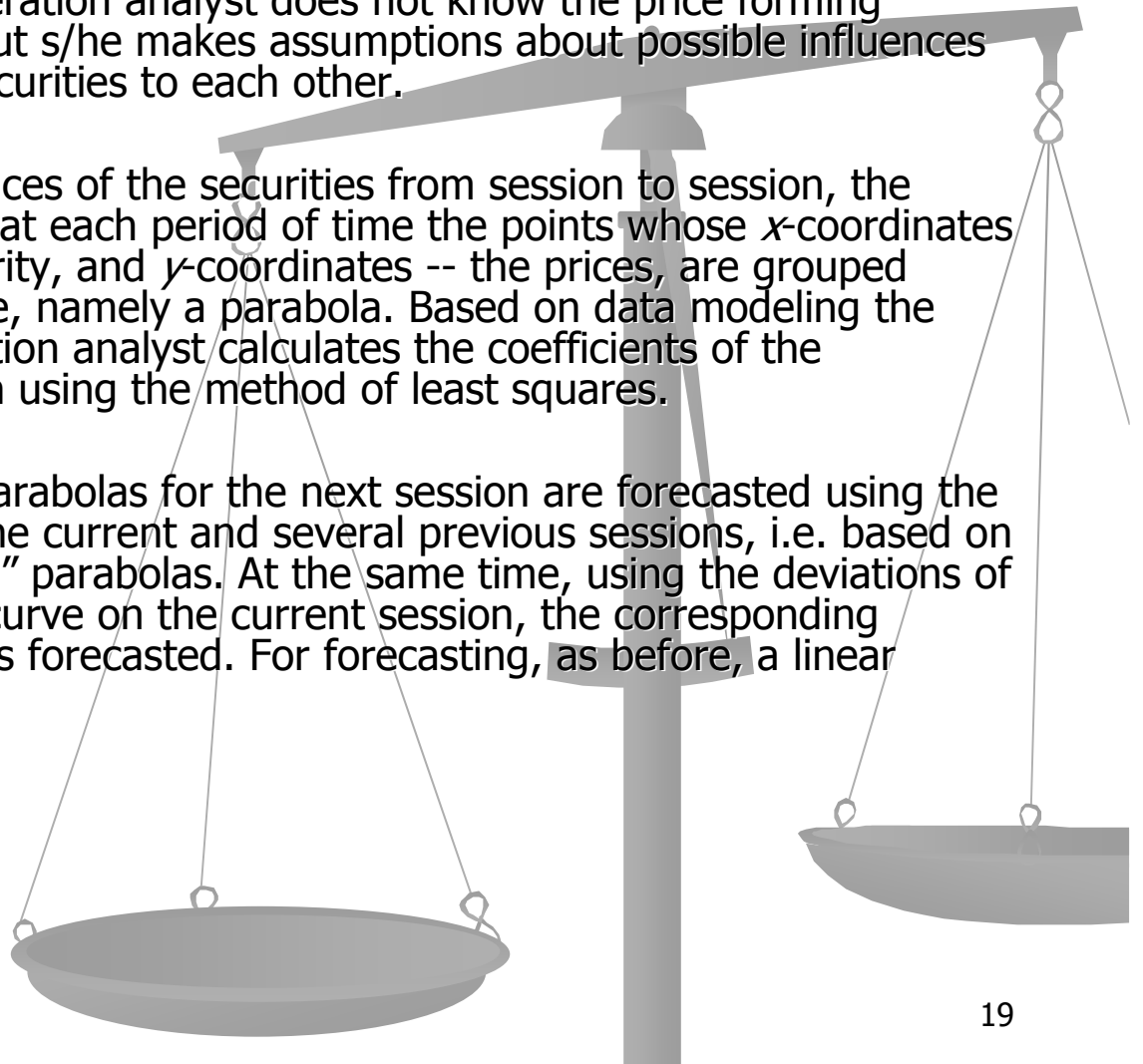
# Conclusions

n    The stochastic optimal control problem with discrete time is considered. The efficiency of its solutions corresponding to the parameters of a stochastic process determined by the method of optimization on time series is analyzed in comparison to that of solutions corresponding to the parameters obtained using a common statistical method of estimation.

n    Parametric optimization problems for continuous and discrete stochastic optimization problems have been introduced and the corresponding problems of optimization on time series formulated. The asymptotic properties of solutions for the formulated problems as a sample size increases analyzed.

n    Applications of the obtained results to a financial portfolio management problem considered.

n    A mathematical model of the financial portfolio management problem, containing a stochastic factor in the form of Markov process is formulated, with the Markov process modeling a financial market. Using the Markov process, data for imitation of the series of observations were generated, and depending on the operational situation and information provided OR analyst, various scenarios implemented.

n    The general plan of computational experiments for different sizes of data imitating the series of realizations and different behaviors of the operating side implemented.
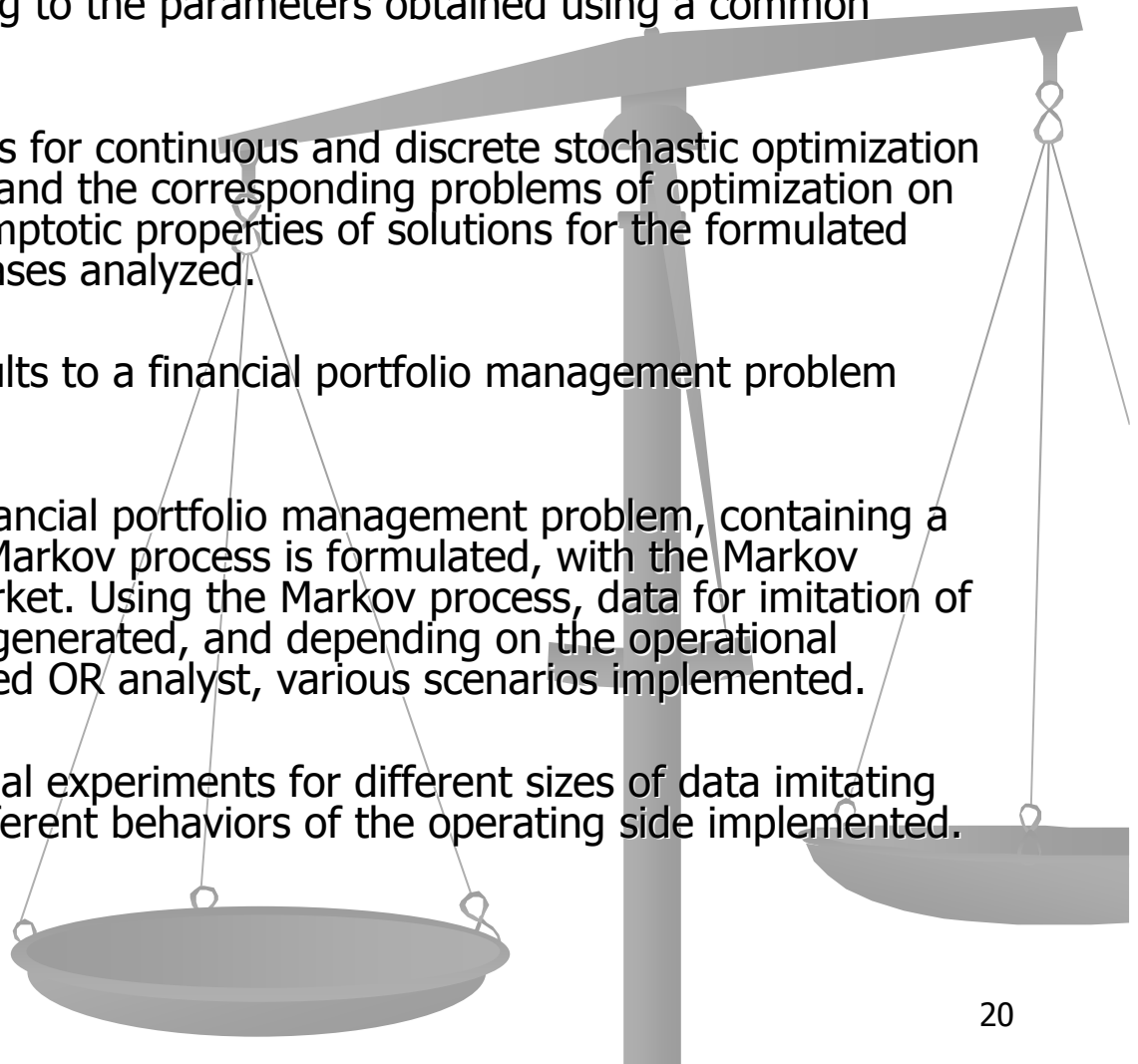
# Future Research

- It is planned to conduct a large number of computational experiments for each scenarios using the both methods of calibrations of the parameters of the stochastic process. Obtained controls and their estimates on the given series of realizations will be compared with their "real" efficiency on the original Markov process.

- Of course, such experiments cannot be considered as a formal proof of the efficiency of the presented method, nevertheless from the point of view of its further utilization the results of the experiments seem to be essential.
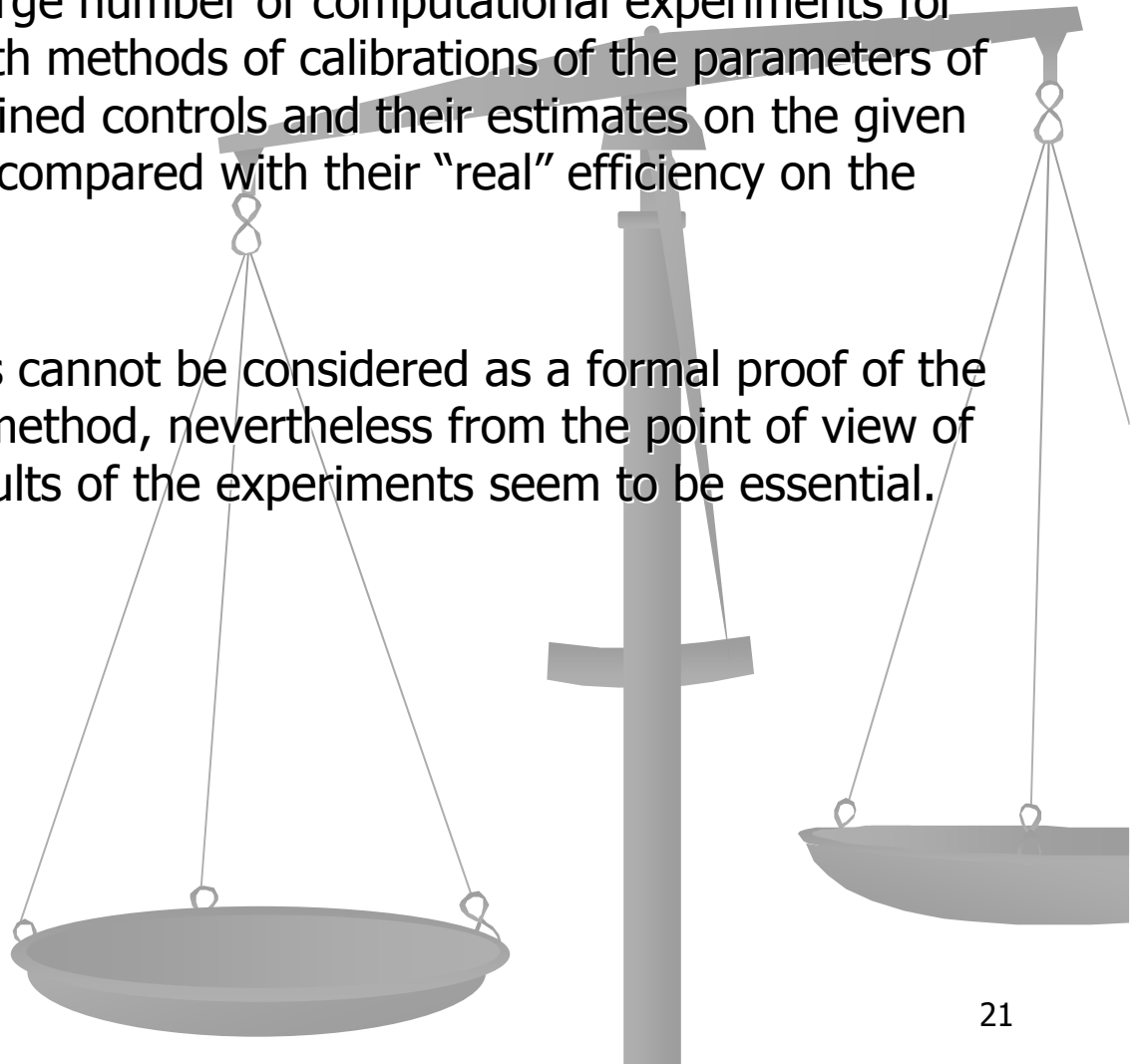
# References

n    Winston, W. L., *Operations Research: Applications and Algorithms*, Belmont, California: Duxbury Press, 1994

n    Dynkin, E. B., and A. A. Yushkevich, *Controlled Markov Processes*, New York: Springer-Verlag, 1979

n    Bertsekas, D. P., and S. E. Shreve, *Stochastic Optimal Control. The Discrete Time Case*, New York: Academic Press, 1978

n    Gasanov I. I., Raguimov, I. S., "On algorithm of financial portfolio control by the method of optimization on time series", *Proc. of the 2004 International Conference on Computational Intelligence for Modeling Control and Automation - CIMCA'2004,* ISBN 1740881885, p. 511-519.

n    Gasanov I. I., Raguimov, I. S., "On solution of stochastic control problem by the method of optimization on time series", Proc. of the *2003 Hawaii International Conference on Statistics and Related Fields* /ISSN 1539-7211, 7 pages

n    Raguimov, I. S., "On decision making in operations with stochastic factors", Proc. of the *2001 International Conference on Computational Intelligence for Modeling Control and Automation-CIMCA''2001* /ISBN 0858898470, p. 580-587
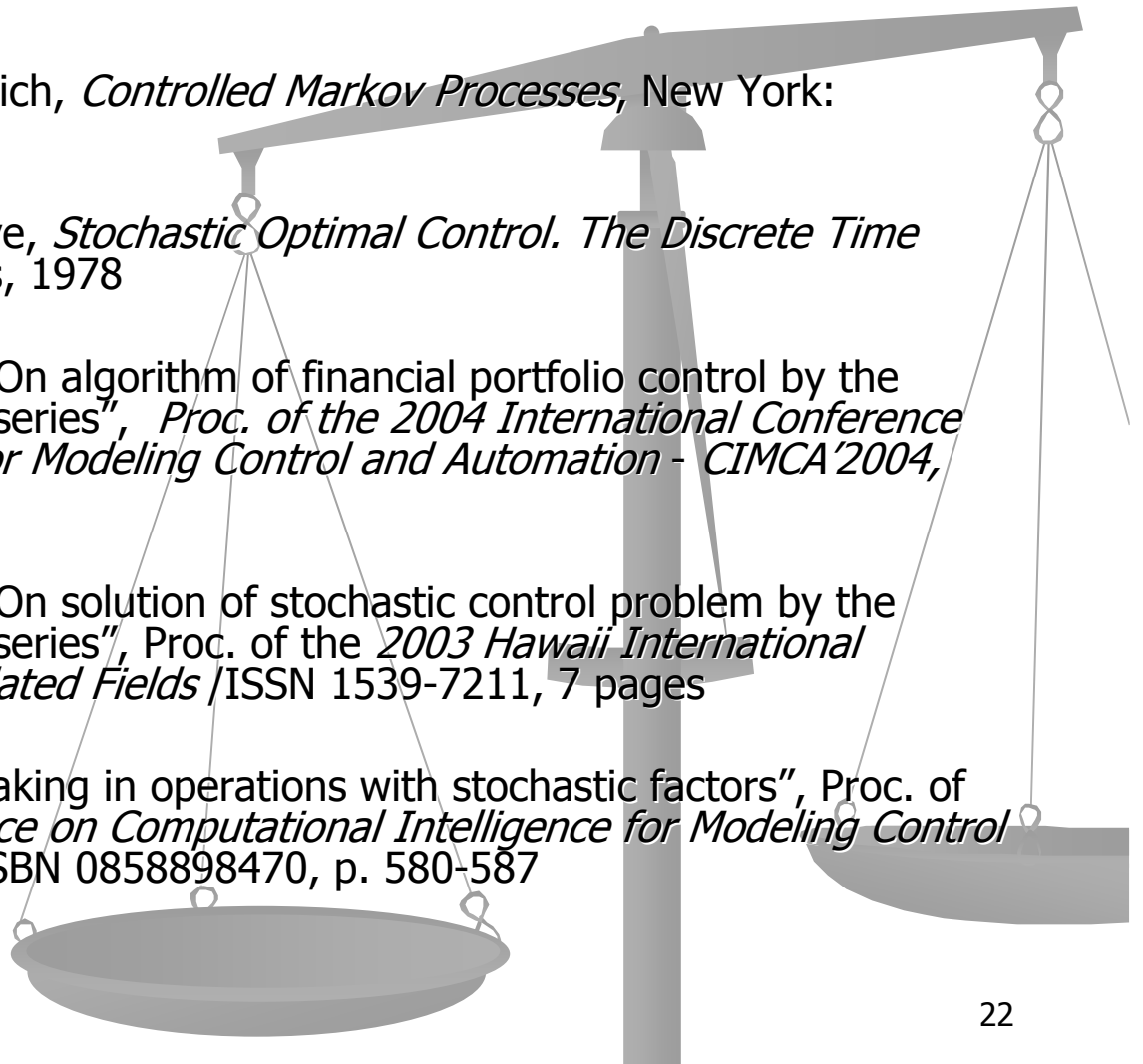
Any Questions or Comments?

THE END