Applications of Mathematics

8

A. N. Shiryayev

Optimal Stopping Rules

Springer-Verlag
New York Heidelberg Berlin

A. N. Shiryayev

Optimal Stopping Rules

Translated by A. B. Aries

Springer-Verlag
New York Heidelberg Berlin

A. N. Shiryayev
Institute of Control Sciences
Moscow 117806
U.S.S.R.

AMS Subject Classifications: 60G40, 62L10, 62L15

Library of Congress Cataloging in Publication Data
Shiriaev, Al'bert Nikolaevich.
Optimal stopping rules.
(Applications of mathematics ; 8)
Translation of Statisticheskiĭ posledovatel'nyĭ analiz.
Bibliography: p.
Includes index.
1. Sequential analysis. 2. Optimal stopping
(Mathematical statistics) I. Title.
QA279.7.S5213 1977 519.5'4 77–11198
ISBN 0-387-90256-2

The original Russian edition STATISTICHESKY POSLEDOVATELNY ANALYZ (OPTIMALNYJE PRAVILA OSTANOVKI) was published in 1976 by Nauka.

Printed in the United States of America.

9 8 7 6 5 4 3 2 1

ISBN 0-387-90256-2 Springer-Verlag New York
ISBN 3-540-90256-2 Springer-Verlag Berlin Heidelberg

Applied Probability
Control
Economics
Information and Communication
Modeling and Identification
Numerical Techniques
Optimization

Applications of Mathematics

8

Preface

1

Along with conventional problems of statistics and probability, the investigation of problems occurring in what is now referred to as *stochastic theory of optimal control* also started in the 1940s and 1950s. One of the most advanced aspects of this theory is the theory of optimal stopping rules, the development of which was considerably stimulated by A. Wald, whose *Sequential Analysis*[1] was published in 1947.

In contrast to the classical methods of mathematical statistics, according to which the number of observations is fixed in advance, the methods of sequential analysis are characterized by the fact that the time at which the observations are terminated (*stopping time*) is random and is defined by the observer based on the data observed. A. Wald showed the advantage of sequential methods in the problem of testing (from independent observations) two simple hypotheses. He proved that such methods yield on the average a smaller number of observations than any other method using fixed sample size (and the same probabilities of wrong decisions). Furthermore, Wald described a specific sequential procedure based on his sequential probability ratio criterion which proved to be optimal in the class of all sequential methods.

By the *sequential method*, as applied to the problem of testing two simple hypotheses, we mean a rule according to which the time at which the observations are terminated is prescribed as well as the terminal decision as to which of the two hypotheses is true. It turns out that the problem of optimal terminal decision presents no particular difficulties and that the problem of finding the best sequential procedure can be reduced

[1]The Russian translation became available in 1960.

to that of finding the optimal stopping time for a Markov sequence constructed in a specific fashion (Sections 4.1, 4.2).

The necessity to use sequential methods did not seem very compelling in the problem of testing two simple hypotheses. However, the two problems given below require by their very nature, a sequential observation procedure, and associated optimal stopping times.

One such problem is the following optimal selection problem.

We are given n objects ordered in accordance with some common characteristic. We assume that the objects arrive in a random sequence. We wish to determine which object is the best one by pairwise comparison.

The problem is to optimize the selection scheme so as to maximize the probability of choosing the best object. (We assume that we have no access to the objects rejected.) We show in Section 2.3 that this problem can be also reduced to that of finding the optimal stopping time for a Markov chain.

The other problem (the so-called disruption problem: Sections 4.3, 4.4) is the following.

Let θ be a random variable taking on the values $0, 1, \ldots,$ and let the observations $\xi_1, \xi_2 \ldots$ be such that for $\theta = n$ the variables $\xi_1, \xi_2, \ldots, \xi_{n-1}$ are independent and uniformly distributed with a distribution function $F_0(x)$, and $\xi_n, \xi_{n+1}, \ldots$ are also independent and uniformly distributed with a distribution function $F_1(x) \neq F_0(x)$. (Thus, the probability characteristics change in the observable process at time θ.) The problem is how to decide by observing the variables $\xi_1, \xi_2, \ldots$ at which instant of time one should give the "alarm signal" indicating the occurrence of discontinuity or disruption (in probabilistic terms). But this should be done as to (on the one hand) avoid a "false alarm," and (on the other hand) so that the interval between the "alarm signal" and the discontinuity occurrence (when the "alarm signal" is given correctly) is minimal. By analogy with the previous problems, the solution of this problem can be also reduced to finding the optimal stopping time for some Markov random sequence.

2

The present book deals with the general theory of optimal stopping rules for Markov processes with discrete and continuous time which enables us to solve, in particular, the problems mentioned above.

The general scheme of the book is the following.

Let $X = (x_n, \mathcal{F}_n, P_x)$, $n = 0, \ldots,$ be a Markov chain[2] with state space $(E, \mathcal{B})$. Here x_n is the state of the chain at time n, the σ-algebra $\mathcal{F}_n$ is interpreted as the totality of events observed before time n inclusively, and P_x is the probability distribution corresponding to the initial state x. Let us assume that if we stop the observations at time n we shall have the gain

[2]The basic probabilistic concepts are given in Chapter 1.

$g(x_n)$. Then the average gain corresponding to the initial state x is the mathematical expectation $M_x g(x_n)$.

Next, let τ be a random variable taking on the values $0, 1, \ldots$ and such that the event $\{\tau = n\} \in \mathcal{F}_n$ for each n.

We shall interpret τ as the instant of time at which the observations are terminated. Then the condition $\{\tau = n\} \in \mathcal{F}_n$ implies that the solution of the problem whether the observations should be terminated at time n depends only on the events observed until and including the time n.

We shall consider the gain $M_x g(x_\tau)$ corresponding to the stopping time τ and the initial state x (assuming that the mathematical expectation $M_x g(x_\tau)$ is defined).

Set

$$s(x) = \sup_\tau M_x g(x_\tau).$$

The function $s(x)$ is said to be a payoff and the time τ_ε such that $s(x) \leqslant M_x g(x_{\tau_\varepsilon}) + \varepsilon$ for all $x \in E$ is said to be an ε-optimal time. The main questions discussed in this book are: What is the structure of the function $s(x)$?; How can this function be determined?; When do the ε-optimal and optimal (*i.e.*, 0-optimal) times coincide?; What is their structure?

Chapter 2 deals with the investigation of these questions for various classes of the functions $g(x)$ and various classes of the times τ (taking, in particular, the value $+\infty$, as well) for the case of discrete time.

Here is a typical result of this chapter. Let us assume that the function $g(x)$ is bounded, $|g(x)| \leqslant C < \infty$, $x \in E$. Then we can show that the payoff $s(x)$ is the smallest excessive majorant of the function $g(x)$, *i.e.*, the smallest function $f(x)$ satisfying the conditions:

$$g(x) \leqslant f(x), \qquad Tf(x) \leqslant f(x),$$

where $Tf(x) = M_x g(x_1)$.

The time

$$\tau_\varepsilon = \inf\{n \geqslant 0 : s(x_n) \leqslant g(x_n) + \varepsilon\}$$

is ε-optimal for any $\varepsilon > 0$, and the payoff $s(x)$ satisfies the equation

$$s(x) = \max\{g(x), Ts(x)\}.$$

Chapter 3 deals with the theory of optimal stopping rules for Markov processes (with continuous time). Most results obtained in this chapter are similar, at least formally, to the pertinent results related to the case of discrete time. We should, however, note that rather advanced tools of the theory of martingales and Markov processes with continuous time have been used in this chapter.

Chapter 1 is of an auxiliary nature. Here, the main concepts of probability theory and pertinent material from the theory of martingales and Markov processes are given and properties of Markov times and stopping times are detailed. Chapter 4 deals with the applications of the results of

Chapters 2 and 3 to the solution of the problem of sequential testing of two simple hypotheses and the problem of disruption for discrete and continuous time.

3

The structure of the present edition is similar to that of 1969; nevertheless, there is substantial difference in content. Chapter 3, which deals with the case of continuous time, has been changed significantly to take into account new and recent results.

Chapter 2 contains some new results as well. Also, simpler proofs are given for some lemmas and theorems.

Finally, we note that the references given consist mainly of textbooks and monographs. References to sources of new results as well as supplementary material can be found in the Notes at the end of each chapter. Each chapter has its own numeration of lemmas, theorems, and formulas; in referring to the lemmas and theorems within each chapter the chapter number is omitted.[3]

To conclude the Preface I wish to express my gratitude to A. N. Kolmogorov for introducing me to the study of sequential analysis and for his valuable advice. I am also grateful to B. I. Grigelionis for many useful discussions pertaining to sequential analysis. I am indebted to G. Yu. Engelbert and A. Engelbert for many helpful comments and suggestions in preparing this edition for publication. I would like also to thank N. N. Moisejev who initiated the writing of this book.

Moscow
March 1977

A. N. SHIRYAYEV

[3]*Editor's Note*: The author's numbering scheme is illustrated by the following examples. References made *in* Chapter 2 *to* Chapter 2 might take the form Theorem 15, Lemma 15, Section 15, Subsection 15.5 (*i.e.*, the fifth subsection of Section 15). References made *in* Chapter 2 *to* Chapter 3 might take the form Theorem 3.15, Lemma 3.15, Section 3.15, Subsection 3.15.5. However, formula numbers begin with the chapter number, whether the reference is to a formula in the same or to another chapter, so that (2.15) signifies a reference (in Chapter 2 *or* in Chapter 3) to the fifteenth formula of Chapter 2. Finally, figures are numbered sequentially from start to finish of the entire work, whereas: theorems, lemmas, formulas, and footnotes are numbered sequentially by chapter; definitions are numbered sequentially by section; and remarks are numbered sequentially by subsection.

Applications of Mathematics

Volume1
W. H. Fleming and R. W. Rishel
Deterministic and Stochastic Optimal Control
1975. ix, 222p. 4 illus. cloth

Volume 2
G. I. Marchuk
Methods of Numerical Mathematics
1975. xii, 316p. 10 illus. cloth

Volume 3
A. V. Balakrishnan
Applied Functional Analysis
1976. x, 309p. cloth

Volume 4
A. A. Borovkov
Stochastic Processes in Queueing Theory
1976. xi, 280p. 14 illus. cloth

Volume 5
R. S. Liptser and A. N. Shiryayev
Statistics of Random Processes I
General Theory
1977. x, 394p. cloth

Volume 6
R. S. Liptser and A. N. Shiryayev
Statistics of Random Processes II
Applications
1977. x, 339p. cloth

Volume 7
N. N. Vorob'ev
Game Theory
Lectures for Economists
and Systems Scientists
1977. xi, 178p. 60 illus. cloth

Contents

Random processes: Markov times 1

1.1 Background material from the theory of probability

1.1.1

Let $(\Omega, \mathscr{F})$ be a measure space, i.e., a set Ω of points ω with a distinguished system $\mathscr{F}$ of its subsets forming a σ-algebra.

According to Kolmogorov's axiomatics the basis for all probability arguments is a probability space $(\Omega, \mathscr{F}, P)$ where $(\Omega, \mathscr{F})$ is a measure space and P is probability measure (probability) defined on sets from $\mathscr{F}$ and having the following properties:

$$P(A) \geq 0,\ A \in \mathscr{F} \qquad \text{(nonnegativity);}$$

$$P(\Omega) = 1 \qquad \text{(normability);}$$

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i) \qquad \text{(countable or } \sigma\text{-additivity);}$$

here $A_i \in \mathscr{F}$, $A_i \cap A_j = \varnothing$, $i \neq j$, where $\varnothing$ is the empty set.

The class of sets $\mathscr{F}^P$ is said to be the *completion of* $\mathscr{F}$ *with respect to measure* P if $\mathscr{F}^P$ contains the sets $A \subseteq \Omega$ when, for A_1 and $A_2 \in \mathscr{F}$, $A_1 \subseteq A \subseteq A_2$ and $P(A_2 - A_1) = 0$. The system of sets $\mathscr{F}^P$ is a σ-algebra, and the measure P extends uniquely to $\mathscr{F}^P$. A probability space $(\Omega, \mathscr{F}, P)$ is said to be *complete* if $\mathscr{F}^P$ coincides with $\mathscr{F}$.

Let $(\Omega, \mathscr{F})$ be a measure space and let $\overline{\mathscr{F}} = \bigcap_P \mathscr{F}^P$ where the intersection is taken over all probability measures P on $(\Omega, \mathscr{F})$. The system $\overline{\mathscr{F}}$ is a σ-algebra whose sets are said to be *absolutely measurable* sets in the space $(\Omega, \mathscr{F})$.

Let $(\Omega, \mathscr{F})$ and $(E, \mathscr{B})$ be two measure spaces. The function $\xi = \xi(\omega)$ defined on Ω and taking on values in E is said to be *$\mathscr{F}/\mathscr{B}$-measurable* if the set $\{\omega : \xi(\omega) \in B\} \in \mathscr{F}$ for each $B \in \mathscr{B}$. In the theory of probability such functions are known as *random elements with values in E*. If $E = R$ and $\mathscr{B}$ is the σ-algebra of Borel subsets of R, the $\mathscr{F}/\mathscr{B}$-measurable functions $\xi = \xi(\omega)$ are said to be *random variables*. (The $\mathscr{F}/\mathscr{B}$-measurable functions are frequently referred to as *$\mathscr{F}$-measurable functions*).

1.1.2

If $\xi = \xi(\omega)$ is a nonnegative random variable, its *mathematical expectation* (denoted $M\xi$) will be, by definition, a *Lebesgue integral* $\int_\Omega \xi(\omega)P(d\omega)$.

The expectation $M\xi$ or the Lebesgue integral[1] of an arbitrary random variable $\xi = \xi(\omega)$ can be defined only in the case where one of the expectations $M\xi^+$ or $M\xi^-$ is finite (here $\xi^+ = \max(\xi, 0)$, $\xi^- = -\min(\xi, 0)$), and then is equal to $M\xi^+ - M\xi^-$.

The random variable ξ is said to be *integrable* if

$$M|\xi| = M\xi^+ + M\xi^- < \infty.$$

The Lebesgue integral $\int_A \xi(\omega)P(d\omega)$ over a set $A \in \mathscr{F}$, denoted by $M(\xi; A)$ as well, is, by definition, $\int_\Omega \xi(\omega)I_A(\omega)P(d\omega)$ where $I_A = I_A(\omega)$ is the *indicator* (characteristic function) of the set A:

$$I_A(\omega) = \begin{cases} 1, & \omega \in A, \\ 0, & \omega \notin A. \end{cases}$$

Thus, $M(\xi; A) = M(\xi I_A)$ and $M(\xi; \Omega) = M\xi$.

If $\mathscr{G}$ is a sub-σ-algebra of $\mathscr{F}$, and ξ is a random variable for which $M\xi$ exists (i.e., $M\xi^+ < \infty$ or $M\xi^- < \infty$), $M(\xi|\mathscr{G})$ denotes the *conditional expectation ξ with respect to $\mathscr{G}$*, i.e., any $\mathscr{G}$-measurable random variable $\eta = \eta(\omega)$ for which, for any $A \in \mathscr{G}$,

$$\int_A \xi(\omega)P(d\omega) = \int_A \eta(\omega)P(d\omega). \tag{1.1}$$

By virtue of the Radon–Nikodym theorem such a random variable $\eta(\omega)$ exists and can be defined from (1.1) uniquely on sets of P-measure zero.

In the case where $\xi(\omega) = I_A(\omega)$ is the indicator of a set A, the conditional expectation $M(I_A|\mathscr{G})$ is denoted by $P(A|\mathscr{G})$ and is said to be the *conditional probability of an event A with respect to $\mathscr{G}$*.

A sequence of random variables ξ_n, $n = 1, 2, \ldots,$ is said to be a *sequence convergent in probability to the random variable ξ* (in this case the notation $\xi_n \xrightarrow{P} \xi$ or $\xi = P\text{-}\lim \xi_n$ is used) if for each $\varepsilon > 0$

$$\lim_{n \to \infty} P\{|\xi_n - \xi| > \varepsilon\} = 0.$$

[1] The Lebesgue integral is frequently denoted by $\int_\Omega \xi(\omega)dP(\omega)$, $\int_\Omega \xi\, dP$, or $\int \xi\, dP$.

The sequence of random variables ξ_n, $n = 1, 2, \ldots$, is said to be *convergent with probability one*, or *convergent almost surely*, to the random variable ξ (the notation is $\xi_n \to \xi$ or $\xi_n \to \xi$ (P-a.s.)) if the set $\{\omega : \xi_n(\omega) \nrightarrow \xi(\omega)\}$ has P-measure zero. It is also said that $\xi_n \to \xi$ *on the set* A if $P(A \cap \{\xi_n \nrightarrow \xi\}) = 0$. In this case the notation is $\xi_n \to \xi$ $(A; (P\text{-a.s.}))$.

If $\xi_n \to \xi$ (P-a.s.) and $\xi_n \leq \xi_{n+1}$ (P-a.s.) we shall write $\xi_n \uparrow \xi$ or $\xi_n \uparrow \xi$ (P-a.s.). The convergence $\xi_n \downarrow \xi$ can be defined similarly.

1.1.3

We shall give the main theorems on passage to the limit under the sign of a Lebesgue integral (expectation).

Theorem 1 (Monotone convergence). *If* $\xi_n \uparrow \xi$ (P-a.s.) *and* $M\xi_1^- < \infty$, *then*

$$M\xi_n \uparrow M\xi. \tag{1.2}$$

If $\xi_n \downarrow \xi$ (P-a.s.) *and* $M\xi_1^+ < \infty$, *then*

$$M\xi_n \downarrow M\xi. \tag{1.3}$$

Theorem 2 (Fatou's lemma). *If* $\xi_n \geq \eta$, $n = 1, 2, \ldots$, *and* $M\eta > -\infty$, *then*[2]

$$M \varliminf \xi_n \leq \varliminf M\xi_n. \tag{1.4}$$

If $\xi_n \leq \eta$, $n = 1, 2, \ldots$, *and* $M\eta < \infty$, *then*

$$\varlimsup M\xi_n \leq M \varlimsup \xi_n \tag{1.5}$$

Theorem 3 (Lebesgue's theorem on *dominated* convergence). *Let* $\xi_n \xrightarrow{P} \xi$ *and let there exist an integrable random variable* η *such that* $|\xi_n| \leq \eta$, $n = 1, 2, \ldots$. *Then* $M|\xi| < \infty$ *and*

$$M|\xi_n - \xi| \to 0, \qquad n \to \infty. \tag{1.6}$$

Remark 1. Theorems 1–3 still hold if the expectations are replaced by conditional expectations in the formulations of these theorems.

The following is the generalization of Lebesgue's theorem on dominated convergence.

Theorem 4. *Let* $\xi_n \xrightarrow{P} \xi$, $M|\xi_n| < \infty$, $n = 1, 2, \ldots$. *Then* $M|\xi| < \infty$ *and*

$$M|\xi_n - \xi| \to 0, \qquad n \to \infty, \tag{1.7}$$

if and only if the random variables ξ_n, $n = 1, 2, \ldots$, *are uniformly integrable.*[3]

[2] We denote by $\varliminf \xi_n$ (or $\liminf \xi_n$, the lower limit of the sequence ξ_n, $n = 1, 2, \ldots$, i.e., $\sup_n \inf_{m \geq n} \xi_m$. Similarly, the upper limit $\varlimsup \xi_n$ (or $\limsup \xi_n$) is $\inf_n \sup_{m \geq n} \xi_m$.

[3] The family of random variables $\{\xi_\alpha, \alpha \in \mathfrak{A}\}$ is said to be *uniformly integrable* if

$$\lim_{x \to \infty} \sup_{\alpha \in \mathfrak{A}} \int_{\{|\xi_\alpha| > x\}} |\xi_\alpha| \, dP = 0.$$

Remark 2 (Generalized Fatou's lemma). Let $\xi_n \geq \eta_n$, $n = 1, 2, \ldots$, where the variables η_n, $n = 1, 2, \ldots$, are uniformly integrable and $\eta_n \xrightarrow{P} \eta_\infty$. Then (compare with (1.4))

$$M \underline{\lim}\, \xi_n \leq \underline{\lim}\, M\xi_n.$$

To prove this we shall note that by virtue of (1.4)

$$\begin{aligned} &M \underline{\lim}\, \xi_n - M\eta_\infty = M(\underline{\lim}\, \xi_n - \eta_\infty) \\ &= M \underline{\lim}(\xi_n - \eta_n) \leq \underline{\lim}(M\xi_n - M\eta_n) \leq \underline{\lim}\, M\xi_n - \lim M\eta_n. \end{aligned}$$

But by virtue of (1.7) $\lim M_{\eta_n} = M_{\eta_\infty}$, hence $M \underline{\lim}\, \xi_n \leq \underline{\lim}\, M\xi_n$.

1.1.4

Let $T = [0, \infty)$, $\overline{T} = T \cup \{\infty\}$, $N = \{0, 1, \ldots\}$, $\overline{N} = N \cup \{\infty\}$. The family of $\mathscr{F}/\mathscr{B}$-measurable functions (random elements) $X = \{\xi_t(\omega)\}$, $t \in T\,(t \in N)$, is said to be a *random process with continuous (discrete) time with values in E*. The random process with discrete time is said to be a *random sequence* as well.

For fixed $\omega \in \Omega$ the function of time $\xi_t(\omega)$, $t \in T$ (or $t \in N$) is said to be a *trajectory corresponding to the elementary outcome* ω.

Each random process $X = \{\xi_t(\omega)\}$, $t \in Z$ (where $Z = T$ in the case of continuous time and $Z = N$ in the case of discrete time), is naturally associated with σ-algebras $\mathscr{F}_t^\xi = \sigma\{\omega : \xi_s, s \leq t\}$, the smallest σ-algebras containing algebras $\mathscr{A}_t^\xi$ generated by sets $\{\omega : \xi_s \in \Gamma\}$, $s \leq t$, $\Gamma \in \mathscr{B}$.

The random process $X = \{\xi_t(\omega)\}$, $t \in T$, is said to be a *measurable process* if for any $\Gamma \in \mathscr{B}$ the set

$$\{(\omega, t) : \xi_t(\omega) \in \Gamma\} \in \mathscr{F} \times \mathscr{B}(T),$$

where $\mathscr{B}(T)$ is the σ-algebra of Borel sets on $T = [0, \infty)$.

The measurable random process $X = \{\xi_t(\omega)\}$, $t \in T$, is said to be a *process adapted to the family of σ-algebras* $F = \{\mathscr{F}_t\}$, $t \in T$, if for each $t \in T$ and $\Gamma \in \mathscr{B}$

$$\{\omega : \xi_t(\omega) \in \Gamma\} \in \mathscr{F}_t.$$

For short we shall denote these processes by $X = (\xi_t, \mathscr{F}_t)$, $t \in T$, or simply $X = (\xi_t, \mathscr{F}_t)$.

The random process $X = (\xi_t, \mathscr{F}_t)$, $t \in T$, is said to be a *progressively measurable process* if for each $t \in T$ and $\Gamma \in \mathscr{B}$

$$\{(\omega, s) : \xi_s(\omega) \in \Gamma, s \leq t\} \in \mathscr{F}_t \times \mathscr{B}([0, t]),$$

where $\mathscr{B}([0, t])$ is the σ-algebras of Borel sets on $[0, t]$.

Each progressively measurable process is measurable and adapted. The converse also holds in a precise sense: if the real[4] process $X = \{\xi_t(\omega)\}$, $t \in T$, is measurable and adapted to $F = \{\mathscr{F}_t\}$, $t \in T$, it permits a progressively

[4] This is a process with values in R or $\overline{R}$.

measurable modification[5] ([72], chap. 4, p. 42). Each real adapted process with right (left) continuous trajectories is a progressively measurable process ([72], chap. 4, p. 43).

1.2 Markov times

1.2.1

In the present section we shall define and discuss the properties of Markov times, which play a decisive role in the theory of optimal stopping rules. The discussion will deal only with the case of continuous time. The definitions and results can be carried almost automatically to the case of discrete time—in which case as a rule they become simpler.

Let $(\Omega, \mathscr{F})$ be a measure space, let $T = [0, \infty)$, and let $F = \{\mathscr{F}_t\}$, $t \in T$, be a nondecreasing sequence of sub-σ-algebras, i.e., $\mathscr{F}_s \subseteq \mathscr{F}_t \subseteq \mathscr{F}$ for $s \leq t$.

Definition 1. The random variable $\tau = \tau(\omega)$ with values in $\bar{T} = [0, \infty]$ is said to be a *Markov time with respect to the system* $F = \{\mathscr{F}_t\}$, $t \in T$,[6] if for each $t \in T$

$$\{\omega : \tau(\omega) \leq t\} \in \mathscr{F}_t.$$

Markov times can be interpreted as random variables independent of the "future."

Definition 2. The Markov time $\tau = \tau(\omega)$ defined in a probability space $(\Omega, \mathscr{F}, P)$ is said to be a *stopping time* or a *finite Markov time* if

$$P\{\tau(\omega) < \infty\} = 1.$$

With each Markov time $\tau = \tau(\omega)$ we may associate the aggregate $\mathscr{F}_\tau$ of the sets $A \in \mathscr{F}$ for which $A \cap \{\omega : \tau(\omega) \leq t\} \in \mathscr{F}_t$ for all $t \in T$. It is easy to verify that $\mathscr{F}_\tau$ is a σ-algebra.

An obvious interpretation of the σ-algebra $\mathscr{F}_\tau$ is the following. By $\mathscr{F}_t$ we shall understand the totality of events related to some physical process and observed before time t. (For example, let $\mathscr{F}_t = \sigma\{\omega : \xi_s, s \leq t\}$ be a σ-algebra generated by values ξ_s, $s \leq t$, of an observable process $X = \{\xi_t\}$, $t \in T$). Then $\mathscr{F}_\tau$ is the totality of events to be observed over the random time τ.

[5] The process $\tilde{X} = \{\tilde{\xi}_t(\omega)\}$, $t \in T$, is said to be a modification of the process $X = \{\xi_t(\omega)\}$, $t \in T$, if, for each $t \in T$, $P(\omega : \tilde{\xi}_t(\omega) \neq \xi_t(\omega)) = 0$.

[6] The words "with respect to the system $F = \{\mathscr{F}_t\}$, $t \in T$" will be omitted when ambiguity is impossible.

Definition 3. The system of σ-algebras $F = \{\mathscr{F}_t\}$, $t \in T$, is said to be a *right continuous* system[7] if for each $t \in T$

$$\mathscr{F}_t = \mathscr{F}_{t+},$$

where $\mathscr{F}_{t+} = \bigcap_{s>t} \mathscr{F}_s$.

Lemma 1. *Let τ be a Markov time. Then the events $\{\tau < t\}$ and $\{\tau = t\}$ belong to $\mathscr{F}_t$ for each $t \in T$.*

Proof follows immediately from the fact that

$$\{\tau < t\} = \bigcup_{k=1}^{\infty} \left\{\tau \leq t - \frac{1}{k}\right\} \quad \text{and} \quad \left\{\tau \leq t - \frac{1}{k}\right\} \in \mathscr{F}_{t-1/k} \subseteq \mathscr{F}_t.$$

Lemma 2. *Let a family $F = \{\mathscr{F}_t\}$, $t \in T$, be right continuous and let $\tau = \tau(\omega)$ be a random variable with values in $\bar{T} = [0, \infty]$ such that $\{\tau < t\} \in \mathscr{F}_t$ for all $t \in T$. Then τ is a Markov time, i.e., $\{\tau \leq t\} \in \mathscr{F}_t$, $t \in T$.*

PROOF. If $\{\tau < t\} \in \mathscr{F}_t$, then $\{\tau \leq t\} \in \mathscr{F}_{t+\varepsilon}$ for each $\varepsilon > 0$. Consequently,

$$\{\tau \leq t\} \in \bigcap_{\varepsilon > 0} \mathscr{F}_{t+\varepsilon} = \mathscr{F}_{t+} = \mathscr{F}_t.$$

The above lemma implies that in the case of right continuous families $F = \{\mathscr{F}_t\}$, $t \in T$, we need only to prove that $\{\tau < t\} \in \mathscr{F}_t$, $t \in T$, in order to verify whether the random variable τ is a Markov time.

In general the condition "$\{\tau < t\} \in \mathscr{F}_t, t \in T$" is weaker than the condition "$\{\tau \leq t) \in \mathscr{F}_t, t \in T$." To convince ourselves that this is the fact we shall put $\Omega = T$. Let $\mathscr{F}$ be a σ-algebra of Lebesgue sets on T,

$$x_t(\omega) = \begin{cases} 0, & t \leq \omega, \\ 1, & t > \omega, \end{cases}$$

and let $\mathscr{F}_t = \sigma\{\omega : x_s(\omega),\ s \leq t\}$. Then the random variable $\tau(\omega) = \inf\{t \geq 0 : x_t(\omega) = 1\}$ satisfies the condition $\{t < t\} \in \mathscr{F}_t$, whereas $\{\tau \leq t\} \notin \mathscr{F}_t$, $t \in T$. □

Remark. Let $t \in N = \{0, 1, \ldots\}$ and let $\tau = \tau(\omega)$ be a random variable with values in $\bar{N} = \{0, 1, \ldots, \infty\}$. Then the condition "$\{\tau \leq n\} \in \mathscr{F}_n, n \in N$" is equivalent to the condition "$\{\tau < n\} \in \mathscr{F}_n, n \in N$."

Lemma 3. *If τ and σ are Markov times, then $\tau \wedge \sigma \equiv$ $\min(\tau, \sigma)$, $\tau \vee \sigma =$ $\max(\tau, \sigma)$, and $\tau + \sigma$ are also Markov times.*

[7] This definition is no longer meaningful in the case of discrete time $t \in N = \{0, 1, \ldots\}$.