BAYESIAN APPROACH TO PARAMETER ESTIMATION: CONVERGENCE ANALYSIS

Anatoli Yashin
Adaptive Resource Policies

July 1983
WP-83-67

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

CONTENTS

BAYESIAN APPROACH TO PARAMETER ESTIMATION: CONVERGENCE ANALYSIS

1. INTRODUCTION

In spite of evident success in the analysis of many aspects of natural phenomena, uncertainty is still one of the most important features of the relations between human beings and natural systems.

The absence of exact knowledge about the structures, regularities, and peculiarities of system functions, the variety of unknown links between subsystems, errors of measurement, and sometimes the practical impossibility of measuring on the one hand, and the need to decide on appropriate control actions under incomplete information on the other, have prompted attempts to arrive at formal descriptions of uncertainties and to analyze their dynamic properties.

One of the most highly developed formal ways of dealing with the dynamic aspects of uncertainties is the theory of random

processes, but the practical application of this formal theory is often accompanied by long and informal procedures to identify when and how the basic assumptions and axioms of the theory may correspond to real situations. This aspect of the probabilistic method becomes especially important when we are dealing with statistical inference or data-processing problems. The Bayesian approach to statistical inference provides a way of taking into account the informal experience and intuition of the person dealing with a particular problem.

A detailed discussion of the Bayesian approach can be found in Savage (1954), Edwards et al. (1963), Box and Tiao (1973), Lindley (1974), and Peterka (1980). Theoretical research in this field has been stimulated mainly by the problems of estimation and control under incomplete information. The main conceptual difficulty in applying the Bayesian method is related to the interpretation of the *a priori* probability, although this difficulty is often overcome by using subjective measures of belief in a "rationally and consistently reasoning person" (Peterka 1980).

With this approach statistical analysis becomes a part of the man-machine interaction procedure. Such a concept widens the scope of the implementation of probabilistic methods for many situations with uncertainties, and provides a rational basis for decision-making. In particular, it is often used to solve identification problems as a first step in adaptive policy design for large-scale systems.

By systematically applying the Bayesian approach it is possible to produce a consistent theory with a formal structure from systems

identification. For example, combining the Bayesian approach with the results and methods of the general theory of processes developed during the last decade has enabled abstract theoretical results of the martingale theory to be applied, and their best implementation in practice to be determined (see Meyer 1966, 1976; Dellacherie 1972; Jacod 1979; Liptser and Shirjaev 1977; etc.).

One of the important characteristics of the Bayesian estimation procedure is its consistency property in parameter estimation, which often provides high-quality adaptive control algorithms. Many papers have been devoted to convergence analysis of Bayesian estimators (see, for example, Kiefer and Wolfowitz 1956; Kraft 1955; Le Cam and Schwartz 1960; Wald 1949; Ljung 1978; Freedman 1963; Doob 1949; Le Cam 1953, 1958).

Sufficient conditions for the convergence of such estimation algorithms was the subject of a paper by Baram and Sandell (1978), which included various assumptions about the properties of the observation process, parameter sets, and the correlation of parameters with the measuring process. Necessary and sufficient conditions of consistency of the parameter estimations for the diffusion observation process was the subject of a paper by Kitsul (1980).

The necessary and sufficient conditions of consistency for the discrete-time observation process and a denumerable set of parameter values were investigated in Yashin (1981). It turns out that the strong consistency property is often equivalent to the property of mutual singularity for some special family of probabilistic measures. The details of singularity conditions

given in Kabanov et al. (1978) make it possible to obtain the convenient conditions of convergence of the Bayesian estimation algorithm in adaptive filtration schemes (Kuznetzov and Yashin 1981; Kuznetzov et al. 1981). The main advantage of these conditions is that they can be checked before the measurement or observation process is begun.

The development of the results of Yashin (1981) will be twofold: first, a dissemination of the conditions in that paper on the wide class of continuous-time random observation processes; and second, an investigation of the consistency property for an uncountable set of parameter values.

This paper is devoted to the investigation of both of these problems. It turns out that in the case of continuous-time random processes, results similar to those of Yashin (1981) are true. However, the proof of the consistency property for an uncountable set of parameter values requires some additional conditions of restraint.

2. PRELIMINARIES

Let $(\Omega, \mathcal{H}, H, P)$ be the probabilistic space, where $H = (\mathcal{H}_t)_{t\geq 0}$ is a nondecreasing right-continuous family of σ-algebras $\mathcal{H}_t$, $t\geq 0$, $\mathcal{H}_t \subseteq \mathcal{H}$, $\mathcal{H}_\infty = \mathcal{H}$, and $\mathcal{H}_0$ is completed by the sets with a P-probability equal to zero. Consider the $\mathcal{H}_0$-measurable integrable random variable $\beta(\omega)$, which takes its values in some interval I of the real line. We will interpret β as the unknown, unobservable parameter of some dynamic system.

Let $\xi_t(\omega)$, $t\geq 0$ be a H-adapted continuous-time random process, taking its values in R^m with right-continuous, and having left limits, sampling paths. Denote by $\bar{H} = (\bar{\mathcal{H}}_t)_{t\geq 0}$, where $\bar{\mathcal{H}}_t = \bigcap_{u>t}\sigma\{\xi_s, s\leq u\}$. For simplicity we will use $\bar{\mathcal{H}}$ to denote the σ-algebra $\bar{\mathcal{H}}_\infty$. The process $\xi_t(\omega)$ will be interpreted as an observation process or the results of the measurement of some system variables.

Definition 1. The $\bar{H}$-adapted random process $\hat{\beta}_t(\omega)$, $t\geq 0$, is said to be a consistent (strongly consistent) estimation of the random variable β if $P\lim_{t\uparrow\infty}\hat{\beta}_t = \beta$ ($\lim_{t\uparrow\infty}\hat{\beta}_t = \beta$ P - a.s.)

We will deal with the properties of the conditional mathematical expectations $\bar{\beta}_t = E(\beta|\mathcal{H}_t)$ as an estimation of β. The simple necessary and sufficient conditions of consistency $\bar{\beta}_t = E(\beta|\bar{\mathcal{H}}_t)$ may be formulated in terms of the $\bar{\mathcal{H}}$-measurability of the random variable β.

Theorem 1. The estimation $\bar{\beta}_t$ is strongly consistent if and only if the random variable $\beta(\omega)$ is $\bar{\mathcal{H}}$-measurable.

The proof of this theorem follows from the Levy theorem about the regular martingale asymptotic behavior and evident property of $\bar{\mathcal{H}}$-measurable functions.

The relation between the consistency property of the Bayesian estimation of the random variable β and the same property of the arbitrary estimation may be determined by the following theorem.

Theorem 2. Let $\hat{\beta}_t$ be some arbitrary consistent $\bar{H}$-adapted estimation of β. Then the estimation $\bar{\beta}_t$ is strongly consistent.

Proof. According to the theorem's condition $P \lim \hat{\beta}_t = \beta$, for any $t \geq 0$, the random variable $\hat{\beta}_t$ is $\bar{\mathcal{H}}_t$-measurable; consequently β will also be $\bar{\mathcal{H}}$-measurable. According to the Levy theorem

$$\lim_{t \uparrow \infty} \bar{\beta}_t = E(\beta | \bar{\mathcal{H}}) \quad P\text{-a.s.}$$

and consistency follows from $\bar{\mathcal{H}}$-measurability of β.

If there is no information on the convergence of some non-Bayesian estimation, the proof of the $\bar{\mathcal{H}}$-measurability of β becomes more difficult. Fortunately there is another way of proving this property for β, as can be seen from the following.

Theorem 3. Let $\{\beta^n\}$, $n = 0,1,2,\ldots$ be the sequence of random variables such that $P \lim_{n \uparrow \infty} \beta^n = \beta$. Then, if the estimations of β^n are consistent for any $n = 0,1,2,\ldots$, the Bayesian estimation of β is strongly consistent.

The proof of the theorem is the simple consequence of the property of the measurable functions.

Using the result of Theorem 3 we should concentrate our efforts on the findings of the appropriate sequence $\{\beta^n\}$, $n = 0,1,2,\ldots$, establishing consistency properties for $\bar{\beta}_t^n = E(\beta^n | \bar{\mathcal{H}}_t)$ for any $n = 0,1,2,\ldots$.

It is clear that the variables $\beta^n(\omega)$, $n = 0,1,\ldots$, should have a more simple structure than $\beta(\omega)$. We will use as such

variables piecewise constant functions with denumerable sets of values. It is well known that for any random integrable variable $\beta(\omega)$ there always exists a sequence of such random integrable variables $\{\beta^n\}$, $n = 0,1,2,\ldots$. The problem therefore is to prove the consistency property for all such β^n, $n = 0,1,2,\ldots$. We start with the investigation of this property for one random variable with a denumerable set of values in a continuous-time random observation process.

3. THE STRONG CONSISTENCY PROPERTY IN THE CASE OF A DENUMERABLE SET OF PARAMETER VALUES

Assume that parameter β takes the denumerable set of values $\{\beta_i\}$, $i \in \aleph$ and the σ-algebra $\mathcal{H}_t$ is generated by $\sigma(\beta)$ and $\bar{\mathcal{H}}_t$

$$\mathcal{H}_t = \sigma(\beta) \vee \bar{\mathcal{H}}_t$$

where $\sigma(\beta)$ is the σ-algebra in Ω generated by β, and $\aleph$ is some denumerable set.

Let P_t, $\bar{P}_t$, and $\bar{P}$ be the restrictions of measure P on σ-algebras $\mathcal{H}_t$, $\bar{\mathcal{H}}_t$, and $\bar{\mathcal{H}}$, respectively. Define the probabilistic measures $P^i(\cdot)$, $i \in \aleph$ on measurable space $(\Omega,\mathcal{H})$ using

$$P^i(A) = \frac{P(A \cap \{\beta = \beta_i\})}{p_i}$$

where $p_i = P(\beta = \beta_i)$, $i \in \aleph$, $\sum_{i \in \aleph} p_i = 1$, $A \in \mathcal{H}$

We will use P_t^i, $\bar{P}_t^i$, and $\bar{P}^i$ to denote the restrictions of measure P^i on σ-algebras $\mathcal{H}_t$, $\bar{\mathcal{H}}_t$, and $\bar{\mathcal{H}}$, respectively, and $\mathfrak{Z}_t^{kj}$ and $\bar{\Lambda}_t^j$ to denote the derivatives $\frac{d\bar{P}_t^k}{d\bar{P}_t^j}(\xi)$ and $\frac{d\bar{P}_t^j}{d\bar{P}_t}(\xi)$, if they exist.

Denote by $\bar{\Lambda}_t(\beta)$ the random process, defined by

$$\bar{\Lambda}_t(\beta) = \sum_{i \in \aleph} I(\beta = \beta_i) \bar{\Lambda}_t^i$$

Let $\pi_j(t) = P(\beta = \beta_j | \bar{\mathcal{H}}_t)$ be the *a posteriori* probability of the event $\{\beta = \beta_j\}$, given $\bar{\mathcal{H}}_t$, and $J_0 = -\sum_i p_i \ln p_i$ is the entropy of the random variable β.

The basic theorem about the consistency properties of the Bayesian estimation of β establishes the relationship between the following conditions:

(A) $\lim_{t \uparrow \infty} \bar{\beta}_t = \beta \quad P$ - a.s.

(B) $\lim_{t \uparrow \infty} \pi_j(t) = I(\beta = \beta_j) \quad P$ - a.s. , $\quad j \in \aleph$

(C) $\bar{P}_t^k \sim \bar{P}_t^j$, $\quad \forall k, j \in \aleph$, $\quad t \geq 0$

(D) $\bar{P}^k \perp \bar{P}^j$, $\quad \forall k, j \in \aleph$, $\quad k \neq j$

(E) $J_0 = E \ln \bar{\Lambda}_\infty(\beta)$

Theorem 4. Let (C) be true. Then

$$(E) \leftrightarrow (D) \leftrightarrow (B) \leftrightarrow (A)$$

Note that the analogs of Theorems 1-4 may be formulated for discrete-time stochastic processes ξ_n, $n \geq 0$. We will give here the formulation of Theorem 4 only.

Let $(\Omega, \mathcal{H}, H', P)$ be the probabilistic space where $H' = (\mathcal{H}'_n)_{n\geq 0}$ is a nondecreasing family of σ-algebras $\mathcal{H}'_n$, $n\geq 0$, $\mathcal{H}'_n \subset \mathcal{H}'_{n+1} \ldots \subseteq \mathcal{H}'$, and $\mathcal{H}'_0$ is completed by the P-zero sets. Assume that $\beta(\omega)$ is the $\mathcal{H}'_0$-measurable random variable which takes its values in some denumerable set $\aleph$. Let ξ_n, $n\geq 0$, be an H'-adapted discrete-time stochastic process taking its values in R^m. Denote by $\bar{H}' = (\bar{\mathcal{H}}'_n)_{n\geq 0}$, where $\bar{\mathcal{H}}'_n = \sigma\{\xi_m(\omega), m\leq n\}$.

Let P_n, $\bar{P}_n$, and $\bar{P}$ be the restrictions of measure P on σ-algebras $\mathcal{H}'_n$, $\bar{\mathcal{H}}'_n$ and $\bar{\mathcal{H}}' = \bar{\mathcal{H}}'_\infty$, respectively. The notations P^i_n, $\bar{P}^i_n$, ξ_n, $\bar{\lambda}_n$, and $\pi_j(n)$ are as defined above, with the natural changing of the index t to index n. The conditions (A), (B), (C), (D), and (E) may be rewritten as follows:

(A') $\lim\limits_{n\to\infty} \bar{\beta}_n = \beta \quad P-a.s.$

(B') $\lim\limits_{n\to\infty} \pi_j(n) = I(\beta=\beta_j) \quad P-a.s. \quad , \quad j\in\aleph$

(C') $\bar{P}^k_n \sim \bar{P}^j_n \quad , \quad \forall k, j\in\aleph \quad , \quad n\geq 0$

(D') $\bar{P}^k \perp \bar{P}^j \quad , \quad \forall k, j\in\aleph$

(E') $J_0 = E \ln\bar{\Lambda}_\infty(\beta)$

Theorem 4'. Let (C') be true. Then

$$(E') \leftrightarrow (D') \leftrightarrow (B') \leftrightarrow (A')$$

Theorem 4' becomes the simple corollary of Theorem 4 if we define σ-algebras $\mathcal{H}_t$, $t\geq 0$, by

$$\mathcal{H}_t = \mathcal{H}_n, \text{ for } n\leq t<n+1, \quad n\geq 0$$

To prove Theorem 4 some additional results will be useful.

4. SOME PROPERTIES OF CONDITIONAL MEASURES

The next assertion establishes a remarkable property of absolute continuous probability distributions.

Lemma 1. Let (C) be true. Then the following assertions are true

(a) For $\bar{P}^k$ and $\bar{P}^j$ - a.s. the following limits exist

$$\lim_{t\uparrow\infty} z_t^{kj} = z_\infty^{kj} \quad , \quad k, j\in\aleph$$

(b) The measures $\bar{P}^k$, $k \in \aleph$ have the Lebesque representations

$$\bar{P}^k(A) = \int_A z_\infty^{kj}\, d\bar{P}^j + \bar{P}^k(A\cap\{z_\infty^{kj}=\infty\}),\ A\in\bar{\mathcal{H}},\ k,j\in\aleph$$

(c) The following conditions are equivalent

$$\bar{P}^k \perp \bar{P}^j \leftrightarrow \bar{P}^k(Z_\infty^{kj}=\infty)=1 \leftrightarrow \bar{P}^j(Z_\infty^{jk}=\infty)=1, k,j\in\aleph$$

The proof of the assertions of Lemma 1 may be done in a similar way as in Kabanov et al. (1978), taking into account the equivalence property of $\bar{P}_t^k$, $\bar{P}_t^j$.

Lemma 2. Let (C) be true. Then measures $\bar{P}_t^k$ and $\bar{P}_t$ are equivalent and

$$\bar{\Lambda}_t^k = \frac{d\bar{P}_t^k}{d\bar{P}_t} = \frac{\pi_k(t)}{p_k} \tag{1}$$

Proof. The property $P^i << P$ follows from the definition of the measures $P^i(\cdot)$, $i \in \aleph$. The properties $P_t^i << P_t$ and $\bar{P}_t^i << \bar{P}_t$ follow from the evident property $P^i << P$. The definition of $P^i(\cdot)$ yields the following formula for

$$\Lambda^i(\omega) \triangleq \frac{dP^i}{dP}$$

$$\Lambda^i(\omega) = \frac{I(\beta = \beta_i)}{p_i}$$

Let $y_t(\omega)$ be an arbitrarily bounded $\bar{\mathcal{H}}$-measurable function. We will use E_k to denote the operation of mathematical expectation with respect to measure P^k. We have

$$E_k y_t = E(\Lambda^k y_t) = E\left[y_t E(\Lambda^k | \bar{\mathcal{H}}_t)\right] = E\left(y_t \frac{\pi_k(t)}{p}\right)$$

The arbitrariness of y_t yields equality (1).

The absolute continuity of $\bar{P}_t$ with respect to $\bar{P}_t^k$, $k \in \aleph$ follows from the evident representation

$$\bar{P}_t(\cdot) = \sum_{i=1}^{\infty} p_i \bar{P}_t^i(\cdot) \tag{2}$$

Indeed, let $k \in \aleph$ be the arbitrary index and $A \in \bar{\mathcal{H}}_t$ be such that $P_t^k(A) = 0$. According to condition (C) for any other index $j \in \aleph$, $P^j(A) = 0$, and consequently, according to formula (2), $\bar{P}_t(A) = 0$, thus completing the proof.

Lemma 3. Let (C) be true. Then $\bar{P}^j$-a.s. for $k, j \in \aleph$, $k \neq j$, $t \geq 0$

$$\frac{\pi_k(t)}{\pi_j(t)} = \frac{p_k}{p_j} z_t^{kj} \tag{3}$$

and

$$\bar{P}^k\left\{0 < \frac{\pi_k(t)}{\pi_j(t)} < \infty\right\} = 1 = \bar{P}^j\left\{0 < \frac{\pi_k(t)}{\pi_j(t)} < \infty\right\} \tag{4}$$

Proof. From the definition of the processes z_t^{kj} we get

$$z_t^{kj} = \frac{d\bar{P}_t^k}{d\bar{P}_t^j}(\omega) = \frac{d\bar{P}_t^k}{d\bar{P}_t} \cdot \frac{d\bar{P}_t}{d\bar{P}_t^j} = \frac{\pi_k(t)}{p_k} \cdot \frac{p_j}{\pi_j(t)}, \ \bar{P}_t^j\text{-a.s.}$$

and formula (3) is true.

The equivalence of the measures $\bar{P}_t^k$ and $\bar{P}_t^j$, $k, j \in \aleph$, $k \neq j$ yields

$$\bar{P}_t^k(z_t^{kj} = 0) = \bar{P}_t^j(z_t^{kj} = 0) = 0$$

which yields (4).

Proof of Theorem 4.

(B) $\Rightarrow$ (D). From (B) and the condition $\sum_{i\in\aleph} \pi_i(t) = 1$ it follows that

$$\lim_{t\uparrow\infty} \frac{\pi_k(t)}{\pi_j(t)} I\{\beta = \beta_j\} = 0 \quad , \qquad k, j \in \aleph \ , \ k \neq j \qquad (5)$$

Taking (5) into account we can get

$$\lim_{t\uparrow\infty} z_t^{kj} I\{\beta = \beta_j\} = 0 \quad , \qquad P\text{-a.s.} \quad k, j \in \aleph \ , \ k \neq j$$

or

$$\bar{P}^j\{z_\infty^{kj} = 0\} = 1 \quad , \qquad k, j \in \aleph \ , \ k \neq j \qquad (6)$$

Using the Lebesque representation we get for $\bar{P}^k\{z_\infty^{kj} = 0\}$ [part (b) of Lemma 1], and using (6) we get

$$\bar{P}^k\{z_\infty^{kj} = 0\} = 0 \quad , \qquad k, j \in \aleph \ , \ k \neq j \qquad (7)$$

Comparing (6) and (7) we get (D).

(D) $\Rightarrow$ (B). Let Γ_{kj} be the singularity set for measures $\bar{P}^k$ and $\bar{P}^j$ such that $\bar{P}^k(\Gamma_{kj}) = 0$, and consequently $P^j(\Gamma_{kj}) = 1$. Using the Lebesque representation of measure $\bar{P}^k(\Gamma_{kj})$ we get

$$\bar{P}^k(\Gamma_{kj}) = \int_{\Gamma_{kj}} z_\infty^{kj} \, d\bar{P}^j + \bar{P}^k\left(\Gamma_{kj} \cap \{z_\infty^{kj} = \infty\}\right) = 0$$
$$k, j \in \aleph \ , \ k \neq j \qquad (8)$$

It follows from (8) that

$$\bar{P}^{j}(\mathfrak{Z}_{\infty}^{kj} = 0) = 1 \quad , \qquad k, j \in \aleph \quad , \quad k \neq j \tag{9}$$

Lebesque respresentation of the measure $\bar{P}^{k}(\mathfrak{Z}_{\infty}^{kj} = 0)$ yields

$$\bar{P}^{k}(\mathfrak{Z}_{\infty}^{kj} = 0) = 0 \quad , \qquad k, j \in \aleph \quad , \quad k \neq j \tag{10}$$

Consequently, the $\{\mathfrak{Z}_{\infty}^{kj} = 0\}$ coincide with the Γ_{kj} $\bar{P}^{k}$ and $\bar{P}^{j}$-a.s.
It follows from (5) and (9) that for any $k, j \in \aleph, k \neq j$

$$\lim_{t\uparrow\infty} \frac{\pi_k(t)}{\pi_j(t)} = 0 \quad , \qquad \bar{P}^{j}\text{-a.s.} \quad , \quad k, j \in \aleph \quad , \quad k \neq j$$

or

$$\lim_{t\uparrow\infty} \frac{\pi_k(t)}{\pi_j(t)} \quad I(\beta = \beta_j) = 0, \ \bar{P}\text{-a.s.} \quad , \quad k, j \in \aleph \quad , \quad k \neq j$$

Property (B) follows from the condition

$$\sum_{i\in\aleph} \pi_i(t) = 1$$

(B) $\Rightarrow$ (E). It follows from (2) and (B) that the following is true

$$I(\beta = \beta_i) \ln \bar{\Lambda}_{\infty}^{i} = I(\beta = \beta_i) \ln p_i \tag{11}$$

Summing both parts of (11) over i yields

$$\sum_{i\in\aleph} I(\beta = \beta_i) \ln \bar{\Lambda}_{\infty}(\beta) = -\sum_{i\in\aleph} I(\beta = \beta_i) \ln p_i \quad \bar{P}\text{-a.s.} \tag{12}$$

Averaging both parts of (12) over P yields

$$E \ln \bar{\Lambda}_{\infty}(\beta) = E \sum_{i\in\aleph} I(\beta = \beta_i) \ln p_i = E \lim_{k\uparrow\infty} \sum_{i=1}^{k} [-I(\beta = \beta_i) \ln p_i] \tag{13}$$

Since the variables $J_k = \sum_{i=1}^{k} \left[- I(\beta = \beta_i) \ln p_i \right]$ increase monotonically as k grows, it is possible to change the orders of integration and to go to the limit in (13). This yields (E) because

$$E \ln \bar{\Lambda}_\infty(\beta) = - \lim_k \sum_{i=1}^{k} p_i \ln p_i = J_0$$

(E) $\Rightarrow$ (B). Condition (E) and formula (5) yield

$$E \sum_{i \in \aleph} I(\beta = \beta_i) \ln \pi_i(\infty) - E \sum_{i \in \aleph} I(\beta = \beta_i) \ln p_i = - \sum_{i \in \aleph} p_i \ln p_i \quad (14)$$

Since

$$E \sum_{i \in \aleph} I(\beta = \beta_i) \ln p_i = \sum_i p_i \ln p_i$$

it follows from (14) that

$$E \sum_{i \in \aleph} I(\beta = \beta_i) \ln \pi_i(\infty) = 0 \quad (15)$$

It is clear that equality (15) may be true if and only if

$$I(\beta = \beta_i) \ln \pi_i(\infty) = 0 \quad \bar{P} - \text{a.s.} \quad , \quad \forall_i \in \aleph$$

Taking into account the equality $\sum_{i \in \aleph} \pi_i(\infty) = 1$ we get property (B).

(B) $\Rightarrow$ (A). Property (B) yields that the indicators $I(\beta = \beta_i)$, $i \in \aleph$ are $\mathcal{H}$-measurable, and consequently the random variable β

is $\bar{\mathcal{H}}$-measurable. According to the Levi Theorem for regular martingales

$$\lim_{t\uparrow\infty} \bar{\beta}_t = \bar{\beta}_\infty = E(\beta|\bar{\mathcal{H}})$$

Since β is $\bar{\mathcal{H}}$-measurable,

$$E(\beta|\bar{\mathcal{H}}) = \beta \qquad \bar{P}\text{-a.s.}$$

(A) $\Rightarrow$ (B). Property (A) yields that random variable β is measurable and consequently $I(\beta = \beta_i)$, $i \in \aleph$ are $\bar{\mathcal{H}}$-measurable random variables. The processes

$$\pi_j(t) = E\left[I(\beta = \beta_i)|\bar{\mathcal{H}}_t\right]$$

are H-adapted regular martingales. Consequently, $\lim_{t\uparrow\infty} \pi_j(t) = \pi_j(\infty)$, $j \in \aleph$ exists $\bar{P}$-a.s. The $\bar{\mathcal{H}}$-measurability of the indicators $I(\beta = \beta_j)$ yields (B) and completes the proof of Theorem 4.

The results of Theorem 4 are too general to be implemented in practical convergence analysis of Bayesian algorithms. The applied statistician expects from statistical theory more convenient conditions which are formulated in terms of parameters and probabilistic characteristics of the systems and processes with which he deals. As will be seen later, such forms of conditions stem immediately from our results if we have some additional information about the observation process. We will consider the situation here when this information is concentrated in the semimartingale properties of the observable process ξ_t.

5. THE PROCESS ξ_t, $t \geq 0$ AS A SEMIMARTINGALE

The semimartingale is one of the key concepts of modern martingale theory. It accumulates the common properties of a wide class of random processes, which can be investigated in the framework of martingale techniques. This idea appeals to human intuition, which is inclined to represent dynamic processes describing natural phenomena as the sum of two components: slow (trend) and quick (noise). Before giving a formal definition we will introduce several new concepts.

Let the notations H, $\mathcal{H}_t$, P be as defined above in Section 2. We will use $M(H,P)$ to denote a class of H-adapted martingales with respect to measure P with regular (i.e., right-continuous and having left limits) sampling paths. The class of H-adapted, nondecreasing processes having a P-integrable variation with regular sampling paths will be denoted by $A^+(H,P)$. The notation $A(H,P) = A^+(H,P) - A^+(H,P)$ will be used for the class of arbitrary H-adapted regular processes with an integrable variation. In a similar way we can introduce the notation $V(H,P)$ for the class of H-adapted processes with a bounded variation. The class of continuous sampling path martingales will be denoted by $M^c(H,P)$. The notations $M_{loc}(H,P)$, $M^c_{loc}(H,P)$, $A_{loc}(H,P)$, and $V_{loc}(H,P)$ will be used for the classes of local martingales, continuous local martingales, the processes of locally integrable variation, and locally bounded variation, respectively. Predictable σ-algebra in $\Omega \times R_+$ generated by H-adapted processes will be denoted by $\pi(H)$, and $\bar{\sigma}$-algebra $\pi(H) \times B(R^m)$ in $\Omega \times R_+ \times R^m$ denoted by $\tilde{\pi}(H)$. $\pi(H)$-measurable processes will also be called H-predictable.

Definition 2. A random process $\xi = (\xi_t, \mathcal{H}_t)$ is called a semimartingale if one can identify the processes V and M such that

$$\xi = V + M$$

$$M \in M_{loc}(H,P) \tag{16}$$

$$V \in V_{loc}(H,P)$$

We will also use the concept of H-predictable projection of the random process.

Definition 3. The H-adapted process ${}^pX = ({}^pX_t)_{t\geq 0}$ is said to be an H-predictable projection of process X if, for any H-predictable non-negative function γ_t and arbitrary H-predictable non-decreasing process A, the following holds

$$E\int_0^\infty X_t\gamma_t dA_t = E\int_0^\infty {}^pX_t\gamma_t dA_t$$

The class of H-adapted semimartingales with respect to measure P will be denoted by S(H,P).

It is not hard to see that local martingales, supermartingales, and submartingales are semimartingales. Arbitrary processes with stationary independent increments are semimartingales. A process X with independent increments will be semimartingale if

$$f(t) = Ee^{i\lambda X_t}$$

is a function of locally bounded variation for any $\lambda \in R$ (Shirjaev 1980). The concept of a semimartingale is applicable to many processes governed by stochastic differential and integro-differential equations.

The class of semimartingales is invariant with respect to equivalent transformation of probabilistic measures and random change time transformations (Shirjaev 1980). Finally, if $X \in S(H,P)$ and $f = f(x)$ $x \in R$ is a twice continuously differentiable function, then the process

$$f = \left[f(X_t), \mathcal{H}_t \right]$$

is also semimartingale. Finally, any stochastic discrete-time process is semimartingale too.

In the next section we will give the singularity conditions for some probabilistic measures corresponding to semimartingales.

6. LOCAL ABSOLUTE CONTINUITY AND SINGULARITY OF PROBABILISTIC MEASURES

We start this section with an analysis of the properties of absolute continuity and singularity for local absolute continuous probability distributions (Kabanov et al. 1978).

Let probabilistic measures $\tilde{P}$ and P be defined on measurable space $(\Omega, \mathcal{H}, H)$, where all notations are the same as in Section 2. Assume that measures $\tilde{P}$ and P are locally equivalent ($\tilde{P} \overset{loc}{\sim} P$), and

the local density is given by

$$Z_t = \frac{d\tilde{P}_t}{dP_t}, \quad t \geq 0$$

which is the Radon-Nicodin derivative of measure $\tilde{P}_t$ with respect to P_t, where $\tilde{P}_t$ and P_t are the restrictions of $\tilde{P}$ and P to σ-algebras $\mathcal{H}_t$, $t \geq 0$. Notice that for any $t \geq 0$ $\tilde{P}(Z_t > 0) = P(Z_t > 0) = 1$.

We now introduce the process

$$M_t = \int_0^t Z_{s-}^{-1} \, dZ_s$$

It is easy to see that process M_t, $t \geq 0$, is H-local martingale and, by definition,

$$Z_t = Z_0 + \int_0^t Z_{s-} \, dM_s$$

Let $\mu^{(M)}(dt,dx)$ be the integer-valued random measure, corresponding to the jumps of M, and let $\nu^{(M)}(dt,dx)$ be its dual H(P)-predictable projection. Define

$$B_t(M) = \langle M^c \rangle_t + \int_0^t \int_{R \setminus \{0\}} \frac{x^2}{1 + |x|} \, d\nu^{(M)}, \quad t \geq 0$$

The following theorem was proved in Kabanov et al. (1978).

Theorem 5. Assume that $\tilde{P} \overset{loc}{<<} P$. Then

$$\tilde{P} << P \Leftrightarrow \tilde{P}\left[B_\infty(M) < \infty\right] = 1$$

$$\tilde{P} \perp P \Leftrightarrow \tilde{P}\left[B_\infty(M) = \infty\right] = 1$$

where $B_\infty(M) = \lim_{t \uparrow \infty} B_t(M)$

The equivalent formulations of the theorem are as follows:

$$\tilde{P} << P \Leftrightarrow \tilde{P}\left[\langle M^c\rangle_\infty + \int_0^\infty \int_E (1-\sqrt{1+x})^2 d\nu^{(M)} < \infty\right] = 1$$

$$\tilde{P} \perp P \Leftrightarrow \tilde{P}\left[\langle M^c\rangle_\infty + \int_0^\infty \int_E (1-\sqrt{1+x})^2 d\nu^{(M)} < \infty\right] = 1$$

or passing from M to Z,

$$\tilde{P} << P \Leftrightarrow \tilde{P}\left[B_\infty(Z) < \infty\right] = 1$$

$$\tilde{P} \perp P \Leftrightarrow \tilde{P}\left[B_\infty(Z) = \infty\right] = 1$$

where

$$B_t(Z) = \int_0^t (Z_{-s}^{-1})^2 d\langle Z^c\rangle_s + \int_0^t \int_E (1-\sqrt{1+xZ_{s-}^{-1}})^2 \nu^{(Z)}(ds,dx)$$

and $\nu^{(Z)}$ is the dual H-predictable projection of measure $\mu^{(Z)}$ corresponding to jumps of Z_t, $t \geq 0$.

These general results become more accessible for applications if they are reformulated in terms of characteristics and parameters corresponding to some particular processes. We will give these conditions for semimartingales in terms of their predictable characteristics (Kabanov et al. 1978).

Assume that the observable process ξ_t, $t \geq 0$ is semimartingale on probabilistic space $(\Omega, \mathcal{H}, H, P)$, where σ-algebra $\mathcal{H}$ and the family $H=(\mathcal{H}_t)$ $t \geq 0$ are as defined above in Section 2.

According to Kabanov et al. (1978) any H-adapted semimartingale may be represented in the form

$$\xi_t = \xi_0 + \alpha_t + m_t^c + \int_0^t \int_{|x|>1} x\, d\mu_s + \int_0^t \int_{|x|\leq 1} x\, d(\mu-\nu)_s$$

where

$\alpha_t \in A_{loc}(H,P) \cap \pi(H)$,

$m_t^c \in M_{loc}^c(H,P)$,

$\mu(ds,dx)$ is the measure of jumps ξ_t,

$\nu(ds,dx)$ is its dual H-predictable projection with respect to measure P.

Assume that process ξ_t, $t \geq 0$, is also semimartingale with respect to probabilistic measure $\tilde{P}$ that is on probabilistic space $(\Omega, \mathcal{H}, \tilde{P})$, and consequently may be represented by

$$\xi_t = \xi_0 + \tilde{\alpha}_t + \tilde{m}_t^c + \int_0^t \int_{|x|>1} x\, d\mu_s + \int_0^t \int_{|x|\leq 1} x\, d(\mu-\tilde{\nu})_s$$

where

$\tilde{\alpha}_t \in A_{loc}(\tilde{P},H)$,

$\tilde{m}_t^c \in M_{loc}^c(\tilde{P},H)$,

$\tilde{\nu}(dt,dx)$ is the dual H-predictable projection of $\mu(ds,dx)$ with respect to probabilistic measure $\tilde{P}$.

Let as above, $\bar{H} = (\bar{\mathcal{H}}_t)$, where $\mathcal{H}_t = \sigma\{\xi_s, s \leq t\}$, and P_t and $\tilde{P}_t$ are the restrictions of P and $\tilde{P}$ on σ-algebra $\bar{\mathcal{H}}_t$, $t \geq 0$. Denote by $\langle m \rangle_t$ ($\langle \tilde{m} \rangle_t$) the H-predictable square characteristic of the martingales m_t^c ($\tilde{m}_t^c$) respectively.

Let $(\tau_n)_{n\geq 0}$ be a sequence of stopping times with respect to H such that $\tau_n \uparrow \infty$ $\tilde{P}$-a.s. The processes $(\xi_{t\wedge\tau_n}, \mathcal{H}_t, P)_{t\geq 0}$ and $(\xi_{t\wedge\tau_n}, \mathcal{H}_t, \tilde{P})_{t\geq 0}$ are also semimartingales with triples of

characteristics

$$\alpha^{\tau_n},\ \langle m\rangle^{\tau_n},\ \nu^{\tau_n}$$

and

$$\tilde{\alpha}^{\tau_n},\ \langle \tilde{m}\rangle^{\tau_n},\ \tilde{\nu}^{\tau_n}$$

Definition 4. The measure P is said to have the property of (τ_n)-uniqueness if the triples $(\tilde{\alpha}^{\tau_n}, \langle \tilde{m}\rangle^{\tau_n}, \tilde{\nu}^{\tau_n})$ uniquely determine the restrictions $\tilde{P}_{\tau_n}$ of measure $\tilde{P}$ to the σ-algebras $\mathcal{H}_{\tau_n}$.

The next conditions will be useful in an analysis of the absolute continuity and singularity properties of probabilistic measures $\tilde{P}$ and P (see Kabanov et al. 1978).

I. $\tilde{P}_0 << P_0$

There exists an $(\bar{H})$-measurable function Y(t,x) such that

II. (a) $d\tilde{\nu} = Y d\nu$

(b) $\nu(\{t\}, E) = 1 \Rightarrow \tilde{\nu}(\{t\}, E) = 1,\ t \geq 0$

(c) $\langle m\rangle_t = \langle \tilde{m}\rangle_t,\ t \geq 0.$

There exists an H-predictable process γ_s such that

$$\text{(d)}\quad \tilde{\alpha}_t - \alpha_t - \int_0^t \int_{|x|\leq 1} x\left[Y(s,x)-1\right] d\nu = \int_0^t \gamma_s d\langle m\rangle_s,\ t \geq 0.$$

Define the H-predictable process B_t as follows:

$$B_t = \int_0^t \gamma_s d\langle m\rangle_s + \int_0^t \int_E \left[1-\sqrt{Y(s,x)}\right]^2 d\nu_s$$

$$+ \sum_{s\leq t} I(0<a_s<1) \left(1-\sqrt{\frac{1-\tilde{a}_s}{1-a_s}}\right)^2 (1-a_s)$$

where

$$a_t = \nu(\{t\}, E)$$

$$\tilde{a}_t = \tilde{\nu}(\{t\}, E)$$

III. (a) $P\{B_t<\infty\} = 1$, $t\geq 0$

(b) $P\{B_\infty<\infty\} = 1$

(c) $P\{B_\infty=\infty\} = 1$

Define the stopping times τ_n by

$$\tau_n = \inf \{t\geq 0: B_t\geq n\}$$

IV. The measure $\tilde{P}$ is (τ_n)-unique.

Theorem 6. (Kabanov et al. 1978) The following statements hold for the semimartingales $(\xi_t, \mathcal{H}_t, P)$ and $(\xi_t, \mathcal{H}_t, \tilde{P})$

1) I, II, III_b, IV $\Rightarrow \tilde{P} << P$

2) $\tilde{P} << P \Rightarrow$ I, II, III_b

3) If I, II, III_a, and IV hold, then $III_c \Leftrightarrow \tilde{P} \perp P$.

The proof of this theorem may be found in Kabanov et al. (1978; Theorem 19).

The results of Theorem 6 are very useful in specifying the strong consistency conditions, as we will do in the next section.

7. CONSISTENCY CONDITIONS FOR BAYESIAN ESTIMATIONS WHEN OBSERVATIONS ARE SEMIMARTINGALES

The condition of absolute continuity and singularity of probabilistic measures $\tilde{P}$ and P formulated in Theorem 6 are given in terms of measure P, that is, in terms of an upper measure which is calculated in the likelihood ratio

$$Z_t = \frac{d\tilde{P}_t}{dP_t} \quad , \ t \geq 0$$

when it exists.

In practical situations, however, the properties of observable processes are usually defined by the measure P which is the lower measure in the likelihood ratio Z_t. In order to reformulate the results of Theorem 6 in terms of measure P, some auxilliary information about local martingale properties will be relevant.

Let $m_t \in M(H,P)$, $m_t > 0$ P-a.s., $t \geq 0$, and $E(m_t^{-1}) < \infty$ for any $t \geq 0$. Denote by $\mu(dt,ds)$ the integer-valued random measure, corresponding to jumps of m_t, and let $\nu(dt,dx)$ be the dual H(P)-predictable projection of $\mu(dt,dx)$. Denote also by $\mu'(dt,ds)$ the integer-valued random measure and the dual H(P)-predictable projection of the process $m'_t = m_t^{-1}$, and $\langle m^c \rangle_t$, $t \geq 0$ is the local square H(P)-predictable characteristic of the continuous part of the process m_t, $t \geq 0$. The formulas for local H(P)-predictable characteristics of the process ξ_t can be given as follows.

Lemma 4. The process $m'_t = m_t^{-1}$ is H(P)-submartingale. The process $\langle m^c \rangle'_t$, $t \geq 0$, and the measure $\nu(dt,dx)$ are characterized by

$$\int_0^t (m'^c_s)^{-2} d\langle m^c \rangle'_t = \int_0^t (m_s)^{-2} d\langle m^c \rangle_s$$

and

$$\int_0^t \int_E f(s,x)\ \nu(ds,dx) = \int_0^t \int_E f\left[s,\ \frac{-x}{m'_s(m'_s + x)}\right]\ \nu'(ds,dx)$$

Proof. The submartingale property of m'_t, $t \geq 0$, follows easily from the Jensen inequality for conditional mathematical expectations.

Using the $\mathcal{H}_0$-stochastic differentiation formula for $m'_t = m_t^{-1}$ we get

$$m'_t = \frac{1}{m_0} + \int_0^t - \left(\frac{1}{m_{s-}}\right)^2 dm_s + \int_0^t \left(\frac{1}{m_s}\right)^3 d\langle m^c \rangle_s$$

$$+ \sum_{s \leq t} \left[\frac{1}{m_s} - \frac{1}{m_{s-}} + \left(\frac{1}{m_{s-}}\right)^2 \Delta m_s\right]$$

It follows that

$$\langle m'^c \rangle_t = \int_0^t (m_s^c)^{-4} d\langle m'^c \rangle_s \quad t \geq 0$$

and consequently

$$\int_0^t (m'^c_s)^{-2} d\langle m'^c \rangle_s = \int (m_s^c)^{-2} d\langle m^c \rangle_s$$

This proves the first part of the lemma.

In order to prove the second part of the lemma, consider the arbitrarily bounded $\mathcal{H}_t$-measurable random variable η_t

and the $\tilde{\Pi}(H)$-measurable function $f(t,x)$, such that

$$E\int_0^t \int_E f(s,x)\ \nu(ds,dt) < \infty$$

for any $t\geq 0$.

We have

$$L_t = E\ \eta_t \int_0^t \int_E f(s,x)\nu(ds,dt) = E\int_0^t \int_E E(\eta_t|\mathcal{H}_{s-})\ f(s,x)\mu(ds,dx)$$

Notice that the jumps of processes m'_t and m_t are related by

$$\Delta m_t = -\frac{\Delta m'_t}{(m'_{t-}+\Delta)m'_{t-}}$$

Taking this into account for L_t, we can get

$$L_t = E\int_0^t \int_E E(\eta_t|\mathcal{H}_{s-})\ f\left[s,\ \frac{-x}{(m'_{s-}+x)m'_{s-}}\right]\ \mu'(ds,dx)$$

$$= E\ \eta + \int_0^t \int_E f\left[s,\ \frac{-x}{(m'_{s-}+x)m'_{s-}}\right]\ \nu'(ds,dx)$$

The arbitrariness of η_t yields the proof of the second part of the lemma.

8. EXAMPLES

(1) Assume that the observation process is a sequence of random variables $\left[X_n(\omega)\right]_{n\geq 0}$, taking their values in R adapted to some nondecreasing family of σ-algebras $\tilde{H} = \{\tilde{\mathcal{H}}_n\}$, $n=0,1,2,\ldots$.

Introduce the family of σ-algebras $H = (\mathcal{H}_t)_{t\geq 0}$ and the process $\xi_t(\omega)$ by

$$\mathcal{H}_t = \tilde{\mathcal{H}}_n \qquad \text{for} \quad n \leq t < n+1$$

$$\xi_t(\omega) = X_n(\omega) \qquad \text{for} \quad n \leq t < n+1$$

Let $\beta(\omega)$ be an $\mathcal{H}_0$-measurable integrable random variable taking its values in the set of non-negative integer numbers. $\mathcal{H}_t$ is defined in the normal way.

Denote by $\mu(ds,dx)$ the integer-valued random measure of jumps of the process ξ_t. The problem is to define the necessary and sufficient conditions for consistency of the estimation $\bar{\beta}_t = E(\beta|\bar{\mathcal{H}}_t)$. Let $V^{\beta}(ds,dx)$ be the dual H-predictable projection of μ. It can be easily shown that

$$V^{\beta}(\{m\},A) = E[\mu(\{m\},A)|\mathcal{H}_{m-1}] = P(\Delta X_m \in A|\mathcal{H}_{m-1}) \quad \text{if} \quad m = 1,2,\ldots$$

$$V^{\beta}(\{t\},A) = 0 \qquad \text{if } t \neq m = 1,2,\ldots$$

$$V^{i}(\{m\},A) = P^{i}(\Delta X_m \in A|\bar{\mathcal{H}}_{m-1}) \quad , \quad i \in \mathbb{N} \quad , \quad m = 1,2,\ldots$$

where $\Delta X_m = X_m - X_{m-1}$.

Denote by $\bar{Q}^k_m(A,B)$ the probabilistic measure on $\left[R \times \Omega,\ \sigma(R) \otimes \bar{\mathcal{H}}_{m-1}\right]$ which is defined as

$$\bar{Q}^k(dx,d\omega) = p^k(\Delta X_m \in dx|\bar{\mathcal{H}}_{m-1})\bar{P}^k_{m-1}(d\omega)$$

Assume that the measures $Q_m^k(.,.)$ and $Q_m^j(.,.)$ are equivalent and denote by $Y^{kj}(m,x)$ the derivative

$$Y^{k,j}(m,x) = \frac{dQ^k(dx,d\omega)}{dQ^j(dx,d\omega)}$$

where we omit for simplicity the symbol ω in $Y^{kj}(m,x)$. Let

$$a_m^k = \bar{P}^k(\Delta X_m \neq 0 \mid \bar{\mathcal{H}}_{m-1})$$

and

$$n_m^j = (1 - a_m^j) I(a_m^j = 1)$$

where $I(a_m^j = 1)$ is the indicator of the event $\{a_m^j = 1\}$. Assuming that P^j - a.s., the following inequality is true for any $t \geq 0$, $k \neq j$, $k,j \in \aleph$

$$\sum_{m \leq t} \left(\int_R \left[1 - \sqrt{Y^{kj}(m,x)}\right] V^j(\{m\},dx) + \left(1 - \sqrt{\frac{1 - a_m^k}{1 - a_m^j}}\right)^2 n_m^j \right) < \infty$$

Let also the measures $\bar{P}_0^k$ and $\bar{P}_0^j$ be equivalent for any $k, j \in \aleph$ and the event $\{a_m^k = 0\}$ yields the event $\{a_m^j = 0\}$ for any $k, j \in \aleph$.

Then, from the results of Kabanov et al. (1978), it follows that $\bar{P}_t^k(\cdot)$ and $\bar{P}_t^j$ are equivalent for any $k, j \in \aleph$ and $t \geq 0$. The conditions of singularity for the measures $\bar{P}^k(\cdot)$ and $\bar{P}^j(\cdot)$ may also be represented with the help of the results in Kabanov et al. (1978), taking into account the equivalence of measures $\bar{P}_t^k$ and $\bar{P}_t^j$:

$$\sum_{m=1}^{\infty} \left(\int_R \left[1 - \sqrt{Y^{kj}(m,x)}\right] V^j(\{m\},dx) + \left(1 - \sqrt{\frac{1 - a_m^k}{1 - a_m^j}}\right)^2 \cdot n_m^j \right) = \infty$$

for any $k, j \in \aleph$, $k \neq j$, P^j- a.s. This is also a condition of consistency of the Bayesian estimation $\bar{\beta}_t$.

(2) Let the process ξ_t be the Markovian jumping process on any probabilistic space $(\Omega, \bar{\mathcal{H}}, P^j)$, $j \in \aleph$, which is characterized by the family of functions $\lambda^j_{\alpha\gamma}$, $\alpha, \gamma \in \Gamma$ where Γ is some denumberable set on R, $j \in \aleph$.

Let the processes $\lambda^j_{\alpha\gamma}(t)$ be the measurable functions of t for any $\alpha, \gamma \in \Gamma$, $j \in \aleph$ and let the following conditions be true:

i) $0 \leq \lambda^j_{\alpha\gamma}(t) < \infty$

ii) $\sum_{\gamma \in \Gamma} \lambda^j_{\alpha\gamma}(t) \leq 0$

iii) $\sup_{\alpha \in \Gamma} \int_0 | \lambda^j_{\alpha\alpha}(s) | \, ds < \infty$

Assume also that measures $\bar{P}^k_0(\cdot)$ and $\bar{P}^j_0(\cdot)$ are equivalent and the following conditions are true for any $t \geq 0$ and $k, j \in \aleph$, $\bar{P}^j$ - a.s.

i) $$\int_0^t I(\xi_{s-} = \alpha) \lambda^k_{\alpha\gamma}(s) ds = \int_0^t I(\xi_{s-} = \alpha) \lambda^j_{\alpha\gamma} \lambda^{k+}_{\alpha\gamma}(s) \lambda^k_{\alpha\gamma}(s) ds$$

ii) $$\int_0^t \sum_{\alpha, \gamma \in \Gamma} \left(1 - \sqrt{\lambda^j_{\alpha\gamma}(s) \lambda^{k+}_{\alpha\gamma}(s)}\right)^2 I(\xi_{s-} = \alpha) \lambda^k_{\alpha\gamma}(s) \, ds < \infty$$

where

$$\lambda^{k+}_{\alpha\gamma} = \begin{cases} (\lambda^k_{\alpha\gamma})^{-1} & \text{if} \quad \lambda^k_{\alpha\gamma} > 0 \\ 0 & \text{if} \quad \lambda^k_{\alpha\gamma} = 0 \end{cases}$$

Then the condition of singularity of measures $\bar{P}^k$ and $\bar{P}^j$ will be

$$\int_0^\infty \sum_{\alpha,\gamma\in\Gamma} \left(1 - \sqrt{\lambda^k_{\alpha\gamma}(\lambda^j_{\alpha\gamma})^{-1}}\right)^2 I\{\xi_s = \alpha\}\, \lambda^j_{\alpha\gamma}(s)\, ds = \infty \quad \bar{P}^j\text{-a.s.}$$

According to Theorem 4 this is equivalent to the almost certain convergence of Bayesian estimation.

(3) Let observation be a continuous-time diffusion-type process:

$$\xi_t = \xi_0 + \int_0^t A(\beta,s,\omega)\, ds + \int_0^t B(s,\omega)\, dw_s$$

where w_s is the Wiener process on $(\Omega,\mathcal{H},P)$, which is H-adapted and, as before, $\mathcal{H}_t = \sigma(\beta) \vee \bar{\mathcal{H}}_t$.

Assume that for any k and j $k,j \in \aleph$ the measures $P^k(\cdot)$ and $P^j(\cdot)$ are equivalent and P^j-a.s. the following inequality is true for any $t \geq 0$ and $k,j \in \aleph$

$$\int_0^t \left[A(k,s,\omega) - A(j,s,\omega)\right]^2 ds < \infty \qquad \bar{P}^j\text{-a.s.}$$

Then for any $t \geq 0$ the measures P^k_t and P^j_t are equivalent and the strong consistency property is equivalent to the P^j of the integral (Kitsul 1980)

$$\int_0^\infty \left[A(k,s,\omega) - A(j,s,\omega)\right]^2 ds = \infty$$

(4) Assume that $\xi_t(\omega)$ is the multivariant point process that is the sequence of $(T_n, X_n)_{n\geq 1}$, where T_n are the stopping times with respect to $H = (\mathcal{H}_t)_{t\geq 0}$, $\mathcal{H}_t = \sigma(\beta) \vee \bar{\mathcal{H}}_t$, such that the following conditions hold

i) $T_1 > 0$

ii) $T_{n+1} > T_n$, if $T_n < \infty$

iii) $T_{n+1} = T_n$, if $T_n = \infty$

and X_n are $\mathcal{H}_{T_n}$-measurable random variables taking their values in $[R, \mathcal{B}(R)]$. The random variable β is as defined above.

The multivariant point process can be represented with the help of the integer-valued random measure $\mu(\cdot)$ on $(]0,\infty[, R)$

$$\mu(]0,t], \Gamma) = \sum_{n\geq 1} I(T_n \leq t)\, I(X_n \in \Gamma), \quad \Gamma \in \mathcal{B}(R)$$

Let $V^i(dt,dx)$ be the dual $\bar{H}$-predictable projection of μ on $(\Omega, \bar{\mathcal{H}}, \bar{P}^i)$, $i \in \aleph$. Denote by $a_t^j = V^j(\{t\}, R \setminus \{0\})$ and assume that for any $k,j \in \aleph$ the event $\{a_t^k = 0\}$ yields the event $\{a_t^j = 0\}$ and P_0^k is equivalent to P_0^j. Assume that there is a function $Y^{kj}(\omega,t,x)$ such that

$$V^k(dt,dx) = Y^{kj}(\omega,t,x)\, V^j(dt,dx) \qquad \bar{P}^j\text{-a.s.}$$

and for any $t \geq 0$

$$\int_0^t \left(1-\sqrt{Y^{kj}(s,x)}\right)^2 V^j(ds,dx) + \sum_{s\leq t} I(0 < a_s^j < 1) \left(1-\sqrt{\frac{1-a_s^k}{1-a_s^j}}\right)(1-a_s^j)$$

Then it follows from Kabanov et al. (1978) that the measure $\bar{P}_t^k$ is equivalent to $\bar{P}_t^j$ for any $k, j \in \aleph$.

The condition that is equivalent to an almost certain convergence of $\bar{\beta}_t$ to β is

$$\left(1-\sqrt{Y^{kj}(s,x)}\right)^2 V^j(ds,dx) + \sum_{s\le t} I(0 < a_s^j < 1) \left(1-\sqrt{\frac{1-a_s^k}{1-a_s^j}}\right)^2 (1-a_s^j) = \infty$$

$$P^j - \text{a.s.} \quad \text{for any } k,j \in \aleph,\ k \neq j$$

9. THE UNCOUNTABLE SET OF PARAMETER VALUES

Consider now the case when β takes its values in some interval I of the real line. Let $\{\beta^n\}$ be the sequence of piece-wise constant functions of ω such that

$$P \lim \beta^n(\omega) = \beta(\omega)$$

Denote by $\bar{P}^x(\cdot)$, $x \in I$ the family of probabilistic measures on $\bar{\mathcal{H}}$ which are defined by the equalities

$$\bar{P}^x(A) = P(A \mid \beta = x) \quad , \qquad A \in \bar{\mathcal{H}}$$

Denote by $\bar{P}_t^x(\cdot)$ the restrictions of P^x on $\bar{\mathcal{H}}_t$.

Theorem 5. For any $x,y \in I$ let the measures $P_t^x(\cdot)$ and $P_t^y(\cdot)$ be equivalent and for any sets $A,B \in \mathcal{B}(I)$, $A \cap B = \phi$ the measures $P^A(\cdot) = \frac{1}{\lambda(A)} \int_A \bar{P}^x(\cdot)\lambda(dx)$ and $P^B(\cdot) = \frac{1}{\lambda(B)} \int_B \bar{P}^x(\cdot)\lambda(dx)$ be orthogonal. Then the estimation $\bar{\beta}_t$ is stongly consistent.

Before proving this theorem we will give some additional statements.

Lemma 4. For any $x,y \in I$ let the measures $\bar{P}_t^x$ and $\bar{P}_t^y$ be equivalent and $\lambda(\cdot)$ some probabilistic measure on $(I, \mathcal{B}(I))$. Then for any sets $A,B \in \mathcal{B}(I)$, $A \cap B = \phi$, the measures $\bar{P}_t^A(\cdot) = \int_A \bar{P}^x(\cdot)\lambda(dx)$ and $\bar{P}_t^B(\cdot) = \int_B \bar{P}_t^x\lambda(dx)$ are equivalent.

Proof. Let $\Gamma \in \bar{\mathcal{K}}$ be such that $\bar{P}_t^A(\Gamma) = 0$. Then $\bar{P}_t^x(\Gamma) = 0$ λ - a.s., and consequently $\bar{P}^B(\Gamma) = 0$.

Lemma 5. Let the conditions of Theorem 5 be true. Then for any n the estimations $\bar{\beta}_t^n$ are strongly consistent.

Proof. According to the choice of the sequence of $\{\beta^n\}$ for any n the random variable β^n has a denumerable set of values. According to the conditions of Theorem 5 the measures $\bar{P}^{ni}(\cdot)$ and $\bar{P}^{nk}(\cdot)$, where $P^{ni}(\cdot) = \frac{P(\cdot \cap \{\beta^n = \beta_i^n\}}{P(\beta^n = \beta_i^n)}$, are orthogonal. It follows from Lemma 5 that the measures $P_t^{ni}(\cdot)$ and $P_t^{nk}(\cdot)$ are equivalent. The result of Theorem 5 then follows from Theorem 4.

10. CONCLUSION

This paper represents the results for the strong consistency property of Bayesian estimation in two cases: a denumerable and uncountable parameter set and wide class of continuous-time stochastic observation processes. In the case of the denumerable set of parameter values the necessary and

sufficient conditions of consistency are formulated in terms of absolute continuity and singularity of some special family of conditional probabilistic measures. In the case of an uncountable parameter set the sufficient condition of strong consistency is formulated. The results of consistency may be specified when more details of the properties of random observation processes are available.

REFERENCES

Baram, Y., and N.R. Sandell (1978) Consistent estimation of finite parameter sets with application to linear system identification. *IEEE Trans. Automat. Contr.* 23(3).

Box, G.E.P., and G.C. Tiao (1973) *Bayesian Inference in Statistical Analysis.* New York: Addison-Wesley.

Dellacherie, C. (1972) *Capacities et Processus Stochastiques (Capacities and Stochastic Processes).* Berlin: Springer.

Doob, J.L., (1949) Le Calcul des Probabilites et ses Applications (Application of the theory of martingales). *Colloque International du CNRS.* 13:22-28.

Edwards, W., H. Lindman, and L.J. Savage (1963) Bayesian statistical inference for psychological research. *Psychol. Rev.* 70:193-242.

Freedman, D.A. (1963) On the asymptotic behaviour of Bayes estimates in the discrete case. *Ann. Math. Statistics.* 34:1386-1403.

Jacod, J. (1979) Calcule Stochastique et Probleme de Martingales (Stochastic Calculus and Problems of Martingales). *Lecture Notes in Math.* 714. Berlin: Springer.

Kabanov, Y.M., R.S. Lipster, and A.N. Shirjaev (1978) Absolutnaya Neprerivnost i Singularnost Localno Absolutno Neprerivnich Verojatnostnich Raspredeleny (Absolute Continuity and Singularity of Locally Absolutely Continuous Probabilistic Distributions). *Mathematichesky Sbornik* 107:149, 3:11.

Kiefer, J., and J. Wolfowitz (1956) Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters. *Ann. Math. Statistics* 27:887-906.

Kitsul, P.I. (1980) Necessary and sufficient conditions of consistency estimations of the parameters of diffusion processes. *Proc. XIII European Meeting of Statisticians.* Brighton: Stonebridge Press.

Kraft, C. (1955) Some conditions for consistency and uniform consistency of statistical procedures. *University of California Pub. Statist.* 2:125-142.

Kuznetzov, N.A., A.V. Lubkov, A.V. and A.I. Yashin (1981) About consistency of Bayesian estimates in adaptive Kalman filtration scheme. *Automatic and remote control* 4:47-56.

Kuznetzov, N.A., and A.I. Yashin (1981) On consistent parameter estimation in adaptive filtering. *Problems of Control and Information Theory.* 10(5):317-327.

Le Cam, L. (1958) Les Properietes Asymptotiques des Solutions de Bayes (The Asymptotic Properties of Bayesian Solutions). *Publ. de l'Inst. de Statist. de l'Univ. de Paris.* 7:17-35.

Le Cam, L. (1953) On some asymptotic properties of maximum likelihood estimates and related Bayes' Estimates. *University of California Publ. Statist.* 1:277-330.

Le Cam, L. and L. Schwartz (1960) A necessary and sufficient condition for the existence of the consistent estimates. *Ann. Math. Statistics* 31:140-150.

Lindley, D.V. (1975) The future of statistics - a Bayesian 21st century (lecture at the Conference on Directions for Mathematical Statistics, Canada 1974). *Supp. Adv. Appl. Prob.*, 7:106-115.

Liptser, R.S., and A.N. Shirjaev (1977) *Statistics of Random Processes. I and II.* Berlin: Springer.

Ljung, L. (1978) Convergence analysis of parametric identification methods. *IEEE Trans. Automat. Contr.* 23(4).

Meyer, P.A. (1966) *Probability and Potentials.* Waltham: Blaisdell.

Meyer, P.A. (1976) Un Cours sur les Integrales Stochastiques (A Course on Stochastic Integrals), Seminare de Probabilities, X. (University of Strasbourg, 1974/5). *Lecture Notes in Math.* 511:245-224. Berlin: Springer.

Peterka, V. (1980) Bayesian approach to system identification. *Trends and Progress in System Identification* edited by P. Eykhoff. Oxford: Pergamon Press.

Savage, L.J. (1954) *The Foundations of Statistics*. New York: Wiley.

Shirjaev, A. (1980) Martingales: recent results and applications *Proc. XIII European Meeting of Statisticians*. Brighton: Stonebridge Press.

Wald, A. (1949) Note on the consistency of the maximum likelihood estimate. *Ann. Math. Statistics* 20:595-601.

Yashin, A.I. (1981) Sostojatelnost Bayesovskich Otcenok Parametros (consistency of Bayesian Parameter Estimates). *Problemy Peredachy Informacii* 1:62-72.