SEQUENTIAL METHODS IN STATISTICS
BANACH CENTER PUBLICATIONS, VOLUME 16
PWN-POLISH SCIENTIFIC PUBLISHERS
WARSAW 1985

ON DOMINATION AND SUFFICIENCY IN SEQUENTIAL ANALYSIS

R. DÖHLER

Department of Mathematics, Technological University,
Karl-Marx-Stadt, G.D.R.

1. Introduction

Let $(\Omega, \mathscr{A}, \mathscr{P} = (P_\theta; \theta \in \Theta))$ be a statistical space and $\boldsymbol{F} = (\mathscr{F}_t)$ a nondecreasing family of sub-σ-algebras of $\mathscr{A}$. The main purpose of this paper is to answer the following questions. The terms in the formulation of these questions will be used here without detailed explanation on the assumption that their intuitive meaning is clear. Exact definitions will be given later on. Assume that τ is a random size of observation with respect to $\boldsymbol{F}$.

1. Suppose that $\mathscr{P}$ is dominated on $\mathscr{F}_t$ for every t and that $\mathscr{F}_\tau$ is an appropriate σ-algebra which describes the observation of τ and of the "process" $\boldsymbol{F}$ up to τ. Is $\mathscr{P}$ then dominated on $\mathscr{F}_\tau$?

2. Let X_t be, for every t, a sufficient statistic for $\mathscr{F}_t$. Under what conditions is (τ, X_τ) sufficient for $\mathscr{F}_\tau$?

As mentioned, we shall consider "random sizes of observation", i.e., we shall not restrict our attention to Markov times (which are defined by $(\tau \leqslant t) \in \mathscr{F}_t$, $\forall t$). Therefore first of all we have to define a σ-algebra $\mathscr{F}_\tau$, which will be conceived so as to describe the observation of an arbitrary random time τ and of the "process" $\boldsymbol{F}$ up to τ. After that there will be given a mathematical definition which tries to cover all possibilities of practical generation of random sizes of observation. To motivate the necessity of such a generalization of Markov times let us consider for example the following four situations:

a) Independent time: First we determine randomly according to a prescribed distribution function G a random variable τ with values in the interval $[0, \infty]$. After that we observe the process up to the realization t of τ. A reasonable statistical space to describe this situation should satisfy the following conditions: For every $\theta \in \Theta$ the time τ is independent of $\mathscr{F}_\infty$, and its distribution G is not a function of θ.

b) Let us first realize an independent time τ_1 as described in a) and then observe F up to a given Markov time τ_2. After stopping at τ_2 we carry out additional observations of size τ_1, i.e., the resulting random size of observation is $\delta = \tau_1 + \tau_2$, which is neither an independent nor a Markov time. Similar examples are $\tau_1 \vee \tau_2$, $\tau_1 \wedge \tau_2$.

c) Consider the sequential randomization procedure described by Bahadur ([1]): Let g_n, $n = 1, 2, \ldots$ be a sequence of random variables such that $0 \leqslant g_n \leqslant 1$ and every g_n is measurable with respect to $\mathscr{F}_n$. Then the sequential randomization proceeds as follows. If we have already conducted observation up to time n, we carry out an auxiliary experiment Z_n whose realizations are the numbers 0 or 1, such that the probability of $Z_n = 1$ is g_n. If the experiment yields 1 then we stop. Otherwise we proceed with the $(n+1)$th observation and the next step follows. Here g_n is interpreted as the conditional probability $P_\theta\big(\tau = n \,|\, \sigma((\tau \geqslant n) \cap \mathscr{F}_n)\big)$ of stopping at time n if we have already conducted observation up to n and know the past $\mathscr{F}_n$.

d) Let $\{\tau_d; d \in \Delta\}$ be a given set of Markov times (with respect to F), where $(\Delta, \mathscr{D})$ is a measurable space. First we determine the realization d of a random index D with prescribed distribution G defined on $(\Delta, \mathscr{D})$. Then we observe the process F up to τ_d. The resulting size τ_D of observation is no longer a Markov time.

A suitable statistical space to describe this method (which we shall not try to define) should guarantee the following properties of τ_D and D.

– D is independent of $\mathscr{F}_\infty$ and has the distribution G which is not a function of θ.

– Every event $(\tau_D \leqslant t)$ should be measurable with respect to the σ-algebra generated by D and $\mathscr{F}_t$.

We could invent a great number of further examples of random sizes of observation which are not Markov times. Therefore it is desirable to answer the questions posed initially not for Markov times only but as generally as possible.

The assertions stated later on can be found in [5] or are slight generalizations of those in [5]. Most of the proofs will be omitted.

This paper differs from [5] essentially in that here more emphasis is put on intuitive explanations and that a modification of a lemma of Sudakov is proved which turns out to be a consequence of the results on domination and sufficiency.

Notations. In the sequel let $(\Omega, \mathscr{A}, \mathscr{P})$ be the underlying statistical space, i.e., $(\Omega, \mathscr{A})$ a measurable space and $\mathscr{P} = \{P_\theta; \theta \in \Theta\}$ a set of probability measures defined on $(\Omega, \mathscr{A})$ and indexed by $\theta \in \Theta$, where Θ is an arbitrary set.

Denote by S the interval $[0, \infty)$ and by T the union $S \cup \{+\infty\}$. Next, let N be the set of positive integers and $\bar{N} = N \cup \{+\infty\}$. Every σ-algebra is automatically understood to be a subset of $\mathscr{A}$.

Suppose that $F = (\mathscr{F}_t)_{t \in T}$ is a nondecreasing collection of σ-algebras. In all our considerations we will not assume $\mathscr{F}_\infty$ to be equal to the σ-algebra generated by the $\mathscr{F}_t$, $t \in S$. If $\mathscr{F}$ is a σ-algebra and A a subset of Ω, then we admit both $A \cap \mathscr{F}$ and $\mathscr{F}_A$ as the notation for the trace $\{A \cap F; F \in \mathscr{F}\}$ of $\mathscr{F}$ on A. If A is a Borel subset of T then $\mathscr{B}_A$ denotes the class of Borel sets belonging to A.

Given a σ-algebra $\mathscr{F}$, the symbol $\bar{\mathscr{F}}$ denotes the closure $\{F \triangle A; F \in \mathscr{F}, A \in \mathscr{A}, P(A) = 0\}$ of $\mathscr{F}$ with respect to some probability measure P defined on $(\Omega, \mathscr{A})$.

If Q is another probability measure on $(\Omega, \mathscr{A})$ and the restriction $P | \mathscr{F}_A$ of P to some trace $\mathscr{F}_A$ (with $A \neq \emptyset$) is absolutely continuous with respect to $Q | \mathscr{F}_A$, then $\left.\frac{dP}{dQ}\right| \mathscr{F}_A$ is the Radon–Nikodym derivative of $P | \mathscr{F}_A$ with respect to $Q | \mathscr{F}_A$.

Instead of "$\mathscr{G}$, $\mathscr{F}$ are conditionally independent with respect to $\mathscr{D}$" we shall write shortly "$\mathscr{G}$, $\mathscr{F}$ are independent given $\mathscr{D}$". Every random variable τ with $\tau(\Omega) \subset T$ will be called random time.

2. The "past" $\mathscr{F}_\tau$ of the "process" F

Let τ be a random time.

We shall start with more or less intuitive considerations to define the "past". For these considerations we will assume that we are already given a σ-algebra, say $\mathscr{D}_\tau$, which describes the observation of τ and of the "past" of F up to τ.

Because of our formulation of the intuitive meaning of $\mathscr{D}_\tau$ this σ-algebra has to enclose $\sigma(\tau)$, i.e., $\sigma(\tau) \subset \mathscr{D}_\tau$. The problem now is to describe additionally the "past" up to τ. In order to get an inclusion from below and from above for this hypothetical σ-algebra $\mathscr{D}_\tau$, let us consider random events of the kind $(\tau \geqslant t)$ and $(\tau \leqslant t)$, $t \in T$.

$(\tau \geqslant t)$ The trace $(\tau \geqslant t) \cap \mathscr{D}_\tau$ would have to contain $(\tau \geqslant t) \cap \mathscr{F}_t$. For if we conduct observation at least up to t then the past $\mathscr{F}_t$ is known. From this statement together with $\sigma(\tau) \subset \mathscr{D}_\tau$ it would follow that

$$(\tau \geqslant t) \cap \mathscr{F}_t \subset \mathscr{D}_\tau. \tag{1}$$

$(\tau \leqslant t)$ The trace $(\tau \leqslant t) \cap \mathscr{D}_\tau$ would have to belong to $\sigma(\tau) \cap \mathscr{F}_t$. For if we conduct observation at most up to t, then (except information on τ) information on the process "at most up to t" is available,

$$(\tau \leqslant t) \cap \mathscr{D}_\tau \subset \sigma(\tau) \vee \mathscr{F}_t. \tag{2}$$

The inclusions (1) and (2) lead us to introduce the σ-algebras

$$\mathscr{F}_{\tau*} = \sigma\big((\tau \geqslant t) \cap \mathscr{F}_t;\ t \in T\big)$$

and

$$\mathscr{F}_\tau = \{A \in \mathscr{A};\ (\tau \leqslant t) \cap A \in \sigma(\tau) \cup \mathscr{F}_t,\ t \in T\},$$

which because of (1) and (2) would have to satisfy

$$\mathscr{F}_{\tau*} \subset \mathscr{G}_\tau \subset \mathscr{F}_\tau. \tag{3}$$

(The notation $\mathscr{F}_{\tau*}$ is taken from Chung and Doob, who mentioned this σ-algebra in [2].)

The inclusion $\mathscr{F}_{\tau*} \subset \mathscr{F}_\tau$ holds in fact, but in general $\mathscr{F}_{\tau*}$ and $\mathscr{F}_\tau$ are not equal, even in the case of a Markov time τ. We could easily find examples which demonstrate that $\mathscr{F}_{\tau*}$ in general does not describe the behaviour of the process $\boldsymbol{F}$ at time τ (see [2], Example 2). Therefore, instead of $\mathscr{F}_{\tau*}$, in the sequel we shall think of $\mathscr{F}_\tau$ as the "past" of the process $\boldsymbol{F}$, though $\mathscr{F}_\tau$ possibly contains too much information, as we can see in (3).

Let us summarize the arguments for working with $\mathscr{F}_\tau$:

– Let $\boldsymbol{F}^\tau = (\mathscr{F}_t^\tau)_{t \in T}$ be defined by $\mathscr{F}_t^\tau = \sigma\big((\tau \leqslant s);\ s \leqslant t\big) \vee \mathscr{F}_t,\ t \in T$. Then the past $\mathscr{F}_\tau$ of $\boldsymbol{F}$ up to τ is equal to the σ-algebra $\{A \in \mathscr{A};\ (\tau \leqslant t) \cap A \in \mathscr{F}_t^\tau,\ t \in T\}$ (see [5]). Suppose now that τ is a Markov time with respect to $\boldsymbol{F}$. Then $\mathscr{F}_t^\tau = \mathscr{F}_t,\ t \in T$, and we get $\mathscr{F}_\tau = \{A \in \mathscr{A};\ (\tau \leqslant t) \cap A \in \mathscr{F}_t,\ t \in T\}$, which means that our definition of $\mathscr{F}_\tau$ for an arbitrary random time is a generalization of the well-known corresponding σ-algebra in the case of a Markov time.

– Let the image $V = \tau(\Omega)$ of the random time τ be countable. Then there seems to be no doubt of the choice of $\mathscr{F}_\tau$ because of the following facts:

(i) The σ-algebras $\mathscr{F}_{\tau*}$ and $\mathscr{F}_\tau$ are equal (i.e., from (3) would follow $\mathscr{F}_{\tau*} = \mathscr{G}_\tau = \mathscr{F}_\tau$). They coincide with $\sigma\big((\tau = t) \cap \mathscr{F}_t;\ t \in V\big)$, and therefore the explicit representation of the events belonging to $\mathscr{F}_\tau$ is known:

$$\mathscr{F}_\tau = \{F = \bigcup_{t \in V} (\tau = t) \triangle F_t;\ F_t \in \mathscr{F}_t,\ t \in V\}.$$

(ii) A real-valued random variable Y is $\mathscr{F}_\tau$-measurable iff there exists a (real-valued) random process $(Y_t)_{t \in V}$, such that Y_t is $\mathscr{F}_t$-measurable, $t \in V$, and Y is equal to the value Y_τ of the process $(Y_t)_{t \in V}$ at time τ, i.e., $Y(\omega) = Y_{\tau(\omega)}(\omega)$, $\omega \in \Omega$.

(iii) Let P be a probability measure on $(\Omega, \mathscr{A})$ and $E(\cdot)$ the expectation with respect to P. Then for every random variable $X \geqslant 0$ we have

$$E(X \mid \mathscr{F}_\tau) = \sum_{t \in V} \mathit{1}_{(\tau = t)} \frac{E(\mathit{1}_{(\tau = t)} X \mid \mathscr{F}_t)}{P(\tau = t \mid \mathscr{F}_t)}, \qquad P\text{-a.s.}$$

(iv) Let $(\Delta, \mathscr{D})$ be a measurable space. If $\mu = \mu(D, \omega)$ is a function defined on $\mathscr{D} \times \Omega$ which is measurable with respect to some σ-algebra $\mathscr{F}$ for every fixed $D \in \mathscr{D}$ and a probability measure for every fixed $\omega \in \Omega$, then μ is said to be an $\mathscr{F}$-measurable decision rule ([1]). We could show now the following representation of $\mathscr{F}_\tau$: Provided $(\Delta, \mathscr{D})$ is a Borel space as defined in [6], μ is an $\mathscr{F}_\tau$-measurable decision rule iff there exists a sequence μ_t, $t \in V$, of decision rules such that μ_t is $\mathscr{F}_t$-measurable and, for every $\omega \in \Omega$ and $D \in \mathscr{D}$, we have

$$\mu(D, \omega) = \mu_{\tau(\omega)}(D, \omega). \tag{4}$$

In [5] this representation of μ has been shown for random sizes of observation (as they will be defined later on) and almost surely only. The proof of (4) under the general conditions stated above will be published in *Wissenschaftliche Zeitschrift der Technischen Hochschule Karl-Marx-Stadt*.

– Let τ be an arbitrary random time and suppose that $(X_t)_{t \in T}$ is a progressively measurable random process (with respect to $\boldsymbol{F}$). Then the value X_τ of this process at time τ is measurable with respect to $\mathscr{F}_\tau$.

– Let $(X_t)_{t \in T}$ be a measurable random process and define $\boldsymbol{F}$ by $\mathscr{F}_t = \sigma(X_{s \wedge t};\, s \in T)$, $t \in T$. Denote by $\mathscr{G}_\tau^X$ the σ-algebra $\sigma(\tau, X_{s \wedge \tau};\, s \in T)$. In [10] under a weak condition (condition 1.11) on the set Ω of the underlying measurable space $(\Omega, \mathscr{A})$ the equality of $\mathscr{G}_\tau^X$ and $\mathscr{F}_\tau$ is shown for the case where τ is a Markov time. (In fact, more is shown: $\mathscr{F}_\tau = \sigma(X_{s \wedge \tau};\, s \in T)$.) This condition has been reformulated for arbitrary random times in [5].

A weaker property is valid for strong Markov processes. In the following let P be a probability measure on $(\Omega, \mathscr{A})$. If $(X_t)_{t \in T}$ is progressively measurable and the σ-algebras $\sigma(\tau, X_{s \vee \tau};\, s \in T)$, $\mathscr{F}_\tau$ are independent given $\sigma(\tau, X_\tau)$ under P, then $\bar{\mathscr{F}}_\tau$ is equal to $\bar{\mathscr{G}}_\tau^X$ ([5], Theorem 1).

As mentioned above, because of (3) the σ-algebra $\mathscr{F}_\tau$ is possibly too large. This might be the reason for the following fact: If we consider the collection $\bar{\boldsymbol{F}} = (\bar{\mathscr{F}}_t)_{t \in T}$ of closures under P of the σ-algebras $\mathscr{F}_t$, $t \in T$, and define the past of $\bar{\boldsymbol{F}}$ up to τ according to $\mathscr{B}_\tau = \{A \in \mathscr{A};\, (\tau \leqslant t) \cap A \in \sigma(\tau) \vee \bar{\mathscr{F}}_t,\ t \in T\}$, then $\mathscr{B}_\tau$ contains $\mathscr{F}_\tau$, but in general the closure of $\mathscr{F}_\tau$ is a strong subset of $\mathscr{B}_\tau$. (For an example see [5], p. 131). The equality $\mathscr{B}_\tau = \bar{\mathscr{F}}_\tau$ holds for instance in the case of strong Markov processes ([5], Theorem 2).

3. Random sizes of observation

Considering Bahadur's sequential method (Example c) of the introduction) let us first try to find out the essential common features of discrete random sizes of observation. The essence of Bahadur's method was as follows: We observe sequentially a random process $(X_n)_{n \in N}$. If we have realized already n observations $X_1, X_2, \ldots, X_n$, we decide – possibly with the aid of the n observed values, and with the aid of auxiliary experiments – whether to stop or to carry out the next observation.

But for this decision no information concerning the future X_{n+1}, $X_{n+2}, \ldots$ and no information on the unknown parameter θ is available, except the information enclosed in the past $X_1, X_2, \ldots, X_n$.

These statements lead us to the following definition:

DEFINITION 1. A random time τ with values in $\bar{N}$ is called a *random size of observation* if it satisfies the following requirements for every $t \in S$:

(i) The event $(\tau = t)$ and the σ-algebra $\mathscr{F}_\infty$ are independent given $\sigma((\tau \geqslant t) \cap \mathscr{F}_t)$ under every P_θ, $\theta \in \Theta$, i.e.,

$$P_\theta(\tau = t \mid \sigma((\tau \geqslant t) \cap \mathscr{F}_t)) = P_\theta(\tau = t \mid \sigma((\tau \geqslant t) \cap \mathscr{F}_\infty)), \qquad P_\theta\text{-a.s.}, \theta \in \Theta. \quad (5)$$

(ii) For the conditional probabilities $P_\theta(\tau = t \mid \sigma((\tau \geqslant t) \cap \mathscr{F}_t))$, $\theta \in \Theta$, there exists a version which is not a function of θ, i.e., a $\sigma((\tau \geqslant t) \cap \mathscr{F}_t)$-measurable random variable $\mathscr{P}(\tau = t \mid \sigma((\tau \geqslant t) \cap \mathscr{F}_t))$ such that

$$P_\theta(\tau = t \mid \sigma((\tau \geqslant t) \cap \mathscr{F}_t)) = \mathscr{P}(\tau = t \mid \sigma((\tau \geqslant t) \cap \mathscr{F}_t)), \qquad P_\theta\text{-a.s.}, \theta \in \Theta. \quad (6)$$

This definition does not make sense if τ is a random time whose image is the whole interval T, because (5) and (6) are fulfilled for every random time with continuous distribution functions. Therefore we shall make use of another definition, which for $\tau(\Omega) \subset \bar{N}$ turns out to be equivalent to Definition 1 ([4]).

DEFINITION 2. A random time τ is called a *random size of observation* (with respect to $\boldsymbol{F}$; abbreviation r.o.) if for every $t \in S$ the following two conditions are fulfilled:

(i) The event $(\tau \leqslant t)$ and the σ-algebra $\mathscr{F}_\infty$ are independent given $\mathscr{F}_t$, i.e.,

$$P_\theta(\tau \leqslant t \mid \mathscr{F}_t) = P_\theta(\tau \leqslant t \mid \mathscr{F}_\infty), \qquad P_\theta\text{-a.s.}, \theta \in \Theta. \quad (7)$$

(ii) For the conditional probabilities $P_\theta(\tau \leqslant t \mid \mathscr{F}_t)$, $\theta \in \Theta$, there exists a version that is not a function of $\theta \in \Theta$:

$$P_\theta(\tau \leqslant t \mid \mathscr{F}_t) = \mathscr{P}(\tau \leqslant t \mid \mathscr{F}_t), \qquad P_\theta\text{-a.s.}, \theta \in \Theta. \quad (8)$$

Obviously any Markov time fulfils these conditions (take every $P_\theta(\cdot \mid \cdot) = I_{(\tau \leqslant t)}$). But we can say even more:

As a particular case this definition covers essentially the definition of Markov times: Take $\mathscr{F}_\infty = \mathscr{A}$; then conditions (i) and (ii) are equivalent to the fact that the event $(\tau \leqslant t)$ belongs to $\bar{\mathscr{F}}_t = \{F_t \triangle A;\ F_t \in \mathscr{F}_t,\ A \in \mathscr{A},\ P_\theta(A) = 0,\ \theta \in \Theta\}$ (cf. [5], Lemma 6).

For $\mathscr{P}$ consisting only of a single measure Siegmund ([11]) gives an equivalent definition and Pitman and Speed ([9]) use the same formulation as in condition (i). (Condition (ii) in this case is automatically fulfilled.) In their papers and in [5] they use the term "randomized time" instead of "random size of observation". In this paper we shall use the latter term to emphasize its statistical meaning.

If θ were regarded as a random variable (for instance in the Bayesian approach) then conditions (i) and (ii) could be summarized as

$$P\big(\tau \leqslant t \mid \sigma(\theta, \mathscr{F}_\infty)\big) = P(\tau \leqslant t \mid \mathscr{F}_t), \qquad P\text{-a.s.},$$

this equation being defined on a suitable probability space $(\Omega, \mathscr{A}, P)$. The last definition in terms of conditional independence only is mentioned by Dawid ([3], Section 7.4).

For any random time τ let $F^\tau = (\mathscr{F}_t^\tau)_{t\in T}$ be defined as in Section 2. We could easily show ([5], Theorem 3), that τ is an r.o. if and only if for every $t \in T$;

(i) $\mathscr{F}_t^\tau$ and $\mathscr{F}_\infty$ are independent given $\mathscr{F}_t$ under every P_θ, $\theta\in\Theta$,

(ii) $\mathscr{F}_t$ is sufficient for $\mathscr{F}_t^\tau$.

The first condition is equivalent to the assertion that $(\mathscr{F}_t, \mathscr{F}_t^\tau)_{t\in T}$ is a Markov process under every P_θ, $\theta\in\Theta$. For this property Bahadur ([1]) uses the term "transitivity" in the case of discrete time.

Every random time τ is Markov with respect to F^τ and, as has been mentioned in Section 2, the past $\mathscr{F}_\tau$ of F up to τ is equal to the "usual" past $\{A\in\mathscr{A}; (\tau \leqslant t) \cap A \in \mathscr{F}_t^\tau, t\in T\}$ of F^τ up to τ. Therefore the properties (7) and (8) would enable us to reduce all considerations to Markov times, as has been proposed in the case of a single probability measure by Pitman and Speed ([9]). (They do not define $\mathscr{F}_\tau$ for a general τ and reduce it to the Markov case.) Such an approach would simplify slightly the technique of the proofs. But in the proofs auxiliary random sizes of observation σ are used. The past of F up to σ and the past of F^τ up to σ are not equal in general. Therefore, in [5], in order to formulate all assertions in terms of the underlying family F only, the explicit reduction to Markov times has not been used. A simple consequence of (7) and (8) is the following. If τ is an r.o., then every Markov time σ with respect to F^τ is an r.o. with respect to F.

Let us now return briefly to the examples mentioned in the introduction. We shall find out if they fulfil the requirements of Definition 2.

a) Since τ and $\mathscr{F}_\infty$ are independent we have

$$P_\theta(\tau \leqslant t \mid \mathscr{F}_\infty) = P_\theta(\tau \leqslant t) = G(t), \qquad P_\theta\text{-a.s., } t\in T,\ \theta\in\Theta,$$

i.e., the conditional probability on the left-hand side is a constant and therefore $\mathscr{F}_t$-measurable. So the first condition of Definition 2 is fulfilled: $P_\theta(\tau \leqslant t \mid \mathscr{F}_\infty) = P_\theta(\tau \leqslant t \mid \mathscr{F}_t)$. The second condition follows immediately, if we take $\mathscr{P}(\tau \leqslant t \mid \mathscr{F}_t) = G(t)$.

b) If τ_1 is an independent time, then it is according to a) an r.o. If the time τ_2 is Markov with respect to F, so it is with respect to $F^{\tau_1} = (\mathscr{F}_t^{\tau_1})_{t\in T}$.

The sum $\sigma = \tau_1 + \tau_2$ is a Markov time with respect to F^{τ_1} (because both items are Markov). Therefore according to the last remark stated after Definition 2 it follows that σ is an r.o. with respect to F.

c) The stopping function g_n was understood to be the conditional probability of stopping at time n under the condition that we have already conducted observation up to n,

$$P_\theta\big(\tau = n \mid \sigma((\tau \geqslant n) \cap \mathscr{F}_n)\big) = g_n \cdot 1_{(\tau \geqslant n)}, \qquad n \in N,\ \theta \in \Theta.$$

We could easily construct a statistical space which satisfies this condition (see [1] or [4]). Then from the $\mathscr{F}_n$-measurability of g_n it follows that both requirements of the equivalent Definition 1 are fulfilled.

d) Let us only consider the first condition of Definition 2. Suppose that θ is fixed. Since D and $\mathscr{F}_\infty$ are independent and $\mathscr{F}_t \subset \mathscr{F}_\infty$, the σ-algebras $\sigma(D)$ and $\mathscr{F}_\infty$ are independent given $\mathscr{F}_t$ (see for instance [5], Lemma 5), and it follows that $\sigma(D) \vee \mathscr{F}_t$ and $\mathscr{F}_\infty$ are independent given $\mathscr{F}_t$([5], Lemma 4). The event $(\tau_D \leqslant t)$ belongs to $\sigma(D) \vee \mathscr{F}_t$. Therefore it is independent of $\mathscr{F}_\infty$ given $\mathscr{F}_t$, and (i) is fulfilled. In a similar manner we could prove condition (ii) of Definition 2.

Let us finish this section with a simple consequence of Definition 2. Suppose that $V = \tau(\Omega)$ is countable. If $X \geqslant 0$ is a random variable, then according to Section 2 its conditional expectation under $\mathscr{F}_\tau$ is

$$E_0(X \mid \mathscr{F}_\tau) = \sum_{t\in V} 1_{(\tau = t)} \frac{E_\theta(1_{(\tau = t)} \cdot X \mid \mathscr{F}_t)}{P_\theta(\tau = t \mid \mathscr{F}_t)}.$$

If X is $\mathscr{F}_\infty$-measurable and τ is an r.o., then it follows – essentially from the conditional independence property (i) of τ – that

$$E_0(X \mid \mathscr{F}_\tau) = \sum_{t\in V} 1_{(\tau = t)} \frac{P_\theta(\tau = t \mid \mathscr{F}_t) E_\theta(X \mid \mathscr{F}_t)}{P_\theta(\tau = t \mid \mathscr{F}_t)} =$$

$$= \sum_{t\in V} 1_{(\tau = t)} E_\theta(X \mid \mathscr{F}_t), \qquad P_\theta\text{-a.s.},\ \theta \in \Theta.$$

As a particular case we get the following well-known fact: If τ is a Markov time, then the last equation holds for every random variable $X \geqslant 0$, because a Markov time remains an r.o. with respect to F if we replace $\mathscr{F}_\infty$ by $\mathscr{A}$ (see the remark after Definition 2).

4. Domination in sequential analysis

Let $A \neq \emptyset$ be a random event and $\mathscr{F}$ a σ-algebra. We shall say that the family $\mathscr{P}$ of the underlying statistical space $(\Omega, \mathscr{A}, \mathscr{P})$ is dominated on the trace $\mathscr{F}_A$ if there exists a σ-finite measure μ on $(A, \mathscr{F}_A)$ such that all restrictions $P_\theta | \mathscr{F}_A$, $\theta \in \Theta$, are absolutely continuous with respect to μ. The measure μ will be called the *dominating measure* on $\mathscr{F}_A$. In the sequel A will be the event $(\tau < \infty)$ and $\mathscr{F}$ will be the σ-algebra $\mathscr{F}_\tau$.

Throughout this section let us make the following assumption:

$$\text{For every } t \in S \text{ the family } \mathscr{P} \text{ is dominated on } \mathscr{F}_t. \tag{9}$$

We shall not assume domination on $\mathscr{F}_\infty$, because this would be in contradiction to the fact that in the case of independent observations domination on $\mathscr{F}_\infty$ does not hold. (This is a simple consequence of the strong law of large numbers, see [5], p. 116).

Suppose that τ is an arbitrary r.o. in the sense of Definition 1 and denote by $\mathscr{F}_{\tau<\infty}$ the trace $(\tau < \infty) \cap \mathscr{F}_\tau$. Let us state precisely the first question of the introduction.

– If assumption (9) is fulfilled, is then $\mathscr{P}$ dominated on $\mathscr{F}_{\tau<\infty}$?

– Assume that the answer to the first question is affirmative, and suppose that μ_t is a dominated measure on $\mathscr{F}_t$, $t \in S$, and $f_{\theta,t} = \left.\dfrac{dP_\theta}{d\mu_t}\right| \mathscr{F}_t$, and $\mu_{\tau<\infty}$ is a dominating measure on $\mathscr{F}_{\tau<\infty}$. Is it then possible to choose μ_t, $f_{\theta,t} (t \in S)$, and $\mu_{\tau<\infty}$ in such a way that the Radon–Nikodym derivative of $P_\theta | \mathscr{F}_{\tau<\infty}$ with respect to $\mu_{\tau<\infty}$ is equal to $f_{\theta,\tau}$, where $f_{\theta,\tau}$ is the function on $(\tau < \infty)$ defined by

$$f_{\theta,\tau}(\omega) = f_{\theta,\tau(\omega)}(\omega), \qquad \omega \in (\tau < \infty).$$

Let us start with the construction of $\mu_{\tau<\infty}$:

It is well known ([7], Lemma 7) that $\mathscr{P}$ is dominated on $\mathscr{F}_t$ if and only if there exists a finite or countable subset $\mathscr{P}_t^*$ of $\mathscr{P}$ which is equivalent to $\mathscr{P}$ on $\mathscr{F}_t$, i.e., if F_t belongs to $\mathscr{F}_t$, then

$$P_\theta(F_t) = 0, \quad \forall P_\theta \in \mathscr{P}_t^* \quad \Leftrightarrow \quad P_\theta(F_t) = 0, \quad \forall P_\theta \in \mathscr{P}.$$

By $\mathscr{P}^*$ we shall denote the union of all sets $\mathscr{P}_n^*$, $n \in N$, which is again finite or countable. This subset $\mathscr{P}^*$ of $\mathscr{P}$ is now equivalent to P on every σ-algebra $\mathscr{F}_t$, $t \in S$ (see [5]). Let P^* be a strict convex linear combination of the measures belonging to $\mathscr{P}^*$:

$$P^* = \sum_{P_\theta \in \mathscr{P}^*} c_\theta P_\theta, \qquad 0 < c_\theta \leqslant 1, \sum c_\theta = 1. \tag{10}$$

This specially constructed privileged measure will be the base of all further considerations. Notice that in the construction of P^* no random time plays any role.

THEOREM 1. *Let assumption* (9) *be fulfilled and let* P^* *be a privileged measure defined according to* (10). *Then for every random size of observation* τ *(with* $(\tau < \infty) \neq \emptyset$*) the family* $\mathscr{P}$ *is dominated on* $\mathscr{F}_{\tau<\infty}$ *by the restriction of* P^* *to* $\mathscr{F}_{\tau<\infty}$.

Thus the answer to the first part of the question is positive without any restriction, and the way to construct a dominating measure $\mu_{\tau<\infty}$ on $\mathscr{F}_{\tau<\infty}$ is extremely simple: We have to choose P^* once for ever and the restriction $P^* | \mathscr{F}_{\tau<\infty}$ is dominating for every r.o. τ.

THEOREM 2. *Let assumption* (9) *be fulfilled and suppose that for some* $\theta\in\Theta$ *there exist versions* $f_{\theta,t}$ *of the Radon–Nikodym derivatives* $\left.\frac{dP_\theta}{dP^*}\right| \mathscr{F}_t$, $t\in S$, *such that every path of the random process* $f_\theta = (f_{\theta,t})_{t\in S}$ *is right continuous. Then for every random size of observation* τ *(with* $(\tau<\infty)\neq\emptyset$*) the "value* $f_{\theta,\tau}$ *of the process* f_θ *at time* $\tau<\infty$*" is a version of the Radon-Nikodym derivative of* $P_\theta | \mathscr{F}_{\tau<\infty}$ *with respect to* $P^* | \mathscr{F}_{\tau<\infty}$, i.e., *the function* $f_{\theta,\tau}$ *defined by* $f_{\theta,\tau}(\omega) = f_{\theta,\tau(\omega)}(\omega)$ for $\omega\in(\tau<\infty)$ *is* $\mathscr{F}_{\tau<\infty}$*-measurable and satisfies*

$$P_\theta((\tau<\infty)\cap F) = \int_{(\tau<\infty)\cap F} f_{\theta,\tau}\,dP^*, \qquad F\in\mathscr{F}_\tau.$$

The right-continuity condition concerning f_θ does not seem to be too strong. It is fulfilled for example if F is right continuous, because (f_θ, F) is a martingale on $(\Omega, \mathscr{A}, P^*)$.

Let L_t: $\Theta\times\Omega\to S$ be the likelihood function on $\mathscr{F}_t$ (for $t\in S$), defined by $L_t(\theta,\omega) = f_{\theta,t}(\omega)$, $\theta\in\Theta$, $\omega\in\Omega$. Then Theorem 1 means in other words that from the existence of the likelihood functions L_t, $t\in S$, for fixed sizes of observation it always follows that there exists a likelihood function $L_{(\tau)}$: $\Theta\times(\tau<\infty)\to S$ for a procedure with an arbitrary random size of observation τ.

Theorem 2 says essentially that under weak conditions on L_t, $t\in S$, the likelihood function $L_{(\tau)}$ can be calculated by the help of L_t, $t\in S$:

$$L_{(\tau)}(\theta,\omega) = L_{\tau(\omega)}(\theta,\omega), \qquad \theta\in\Theta,\ \omega\in(\tau<\infty).$$

One simple consequence of the last statement is the following: If $\hat\theta_t$ is a maximum-likelihood estimator for fixed size t, $t\in S$, then a maximum-likelihood estimator $\hat\theta_{(\tau)}$ for the procedure with a random size of observation τ is given by $\hat\theta_{(\tau)}(\omega) = \hat\theta_{\tau(\omega)}(\omega)$, $\quad\omega\in(\tau<\infty)$.

5. Sufficiency in sequential analysis

As usual, we shall say that a σ-algebra $\mathscr{G}$ is sufficient for another σ-algebra $\mathscr{F}$ if $\mathscr{G}$ belongs to $\mathscr{F}$ and if for every event $F\in\mathscr{F}$ there exists a version $\mathscr{P}(F|\mathscr{G})$ of the conditional probabilities $P_\theta(F|\mathscr{G})$, which does not depend on $\theta\in\Theta$.

A statistic X is said to be *sufficient* if $\sigma(X)$ is sufficient.

Throughout this section let $X=(X_t)_{t\in T}$ be a random process with values in a measurable space $(E,\mathscr{E})$ such that every statistic X_t is sufficient for $\mathscr{F}_t$, $t\in T$.

If τ is a random time, denote by (τ, X_τ) the mapping with values in the measurable space $(T\times E, \mathscr{B}_T \otimes \mathscr{E})$ defined on Ω by

$$(\tau, X_\tau)(\omega) = \big(\tau(\omega), X_{\tau(\omega)}(\omega)\big).$$

We shall now try to find a general answer to the second question posed at the beginning of the paper: If τ is an r.o. in the sense of Definition 2, under what conditions is the function (τ, X_τ) a sufficient statistic for $\mathscr{F}_\tau$?

First we shall consider the case of an r.o. τ with a countable (or finite) image $V = \tau(\Omega)$ in a slightly more general speech: Suppose that, for every $t \in V$, we are given an σ-algebra $\mathscr{G}_t$ sufficient for $\mathscr{F}_t$. Denote by $\mathscr{G}_\tau$ the σ-algebra generated by all traces $(\tau = t) \cap \mathscr{G}_t$, $t \in V$. Roughly speaking, this σ-algebra $\mathscr{G}_\tau$ describes the observation of τ and of the "value" of the random process $(\mathscr{G}_t)_{t\in V}$ at time τ. This explanation is motivated by the fact that $\mathscr{G}_\tau$ is equal to $\sigma(\tau, X_\tau)$ if every $\mathscr{G}_t$ is generated by X_t (see [5]).

We shall now formulate a modification of Theorem 8.1. [1], using the introduced general notion of an r.o. and the σ-algebras $\mathscr{F}_\tau$ and $\mathscr{G}_\tau$.

THEOREM 3. *Let τ be a random size of observation whose image $V = \tau(\Omega)$ is countable (or finite), and let $(\mathscr{G}_t)_{t\in V}$ be the sequence of sufficient σ-algebras introduced above. Then the σ-algebra $\mathscr{G}_\tau$ is sufficient for $\mathscr{F}_\tau$.*

If we did not suppose in this theorem that τ is a random size of observation, then we would not get the sufficiency of $\mathscr{G}_\tau$ for $\mathscr{F}_\tau$ (see for instance Example 4, [5], p.132).

Let us now return to the case of arbitrary random sizes of observation. Obviously in the case of continuous time the formulation of Theorem 3 in terms of the σ-algebra $\mathscr{G}_\tau = \sigma\big((\tau = t) \cap \sigma(X_t); t \in T\big)$ would not make sense, because every event $(\tau = t)$ could have probability zero. Therefore instead of $\mathscr{G}_\tau$ we shall consider $\sigma(\tau, X_\tau)$. Let us start with an example which will help us to discuss the necessary conditions for the sufficiency of (τ, X_τ).

EXAMPLE. Suppose that Θ consists only of the numbers 1 and 2; Z is a random variable whose realizations are the numbers -1 and $+1$ and whose distribution depends on $\theta \in \Theta$ (for instance by $P_\theta(Z = 1) = \theta/3$, $\theta \in \Theta$); V is a random variable with values in the interval $(0, \infty)$ whose distribution function $G(x)$ is continuous and does not depend on $\theta \in \Theta$; Z is independent of V under both P_θ, $\theta \in \Theta$.

One could easily construct a statistical space $(\Omega, \mathscr{A}, \mathscr{P})$ which satisfies these assumptions. (For details see [5], Example 2, p. 129. We shall discuss this example here only intuitively and without strong proofs.)

Consider now the random process $X = (X_t)_{t\in T}$ defined by $X_t = Z|\cos(\pi\cdot t/2V)|$ for $t \in S$ and $X_\infty = Z$, and let $\mathscr{F}_t = \sigma(X_s; s \leqslant t) = \sigma(V, Z)$, $t \in T$. This process starts at $X_0 = Z$ and turns at V for the first time to zero. The information on Θ contained in $\mathscr{F}_t$ is concentrated on $\sigma(Z)$, because the distribution of V does not depend on θ.

For $t \in T$ we have: $(X_t > 0) \subset (Z = 1)$, $(X_t < 0) \subset (Z = -1)$, $P_\theta(X_t = 0) = 0$, $\theta \in \Theta$. Therefore, except in an event of probability zero, Z is known if we know X_t, i.e., X_t is sufficient for $\mathscr{F}_t = \sigma(V, Z)$, $t \in T$.

Let τ be equal to V; then X_τ is equal to zero and $\mathscr{F}_\tau$ equals $\sigma(V, Z)$. But the σ-algebra $\sigma(\tau, X_\tau)$ is not sufficient for $\mathscr{F}_\tau$: though $\mathscr{F}_\tau = \sigma(V, Z)$ contains information on θ, the σ-algebra $\sigma(\tau, X_\tau) = \sigma(V)$ does not.

Let us point out some features of the considered example.

– $\mathscr{P}$ is dominated on $\mathscr{F}_t$, $t \in T$.

– The process X is continuous and every statistic X_t is sufficient for $\mathscr{F}_t$.

– τ is a Markov time with respect to F.

The example shows that these conditions do not imply the sufficiency of (τ, X_τ). Therefore we shall add another not too strong condition concerning the right-continuity of the process of Radon–Nikodym derivatives. In order to formulate this condition, let us make a remark on the representation of the process of Radon–Nikodym derivatives.

Suppose that – in addition to the sufficiency assumption on X stated at the beginning of the section – condition (9) is fulfilled, i.e., $\mathscr{P}$ is dominated on $\mathscr{F}_t$, $t \in S$. Further, let P^* be the measure constructed in (10). Then, according to a well-known theorem of Halmos and Savage ([7], Theorem 1), for every $t \in S$ and $\theta \in \Theta$ there exists a version $f_{\theta,t}$ of the Radon–Nikodym derivative of $P_\theta | \mathscr{F}_t$ with respect to $P^* | \mathscr{F}_t$ which is $\sigma(X_t)$-measurable and which therefore has a representation $f_{\theta,t} = g_{\theta,t}(X_t)$, where $g_{\theta,t} = g_{\theta,t}(x)$ is a measurable real-valued function defined on $(E, \mathscr{E})$. We shall denote the functions g_θ, $\theta \in \Theta$, defined on $T \times E$ by $g_\theta(t, x) = g_{\theta,t}(x)$ as "likelihood-generating functions".

THEOREM 4. *Let $\mathscr{P}$ be dominated on $\mathscr{F}_t$ for every $t \in S$. Suppose, in addition to the assumptions on X stated at the beginning of this section, that*

– *$(E, \mathscr{E})$ is a metric space and $\mathscr{E}$ its σ-algebra of Borel sets,*

– *every path of X is right-continuous.*

If there exists a collection g_θ, $\theta \in \Theta$, of likelihood-generating functions such that for every $\theta \in \Theta$

$$g_\theta(t, x) = \lim_{\substack{s \to t \\ y \to x}} g_\theta(s, y), \qquad t \in S,\ x \in E, \tag{11}$$

then for every random size of observation τ the mapping (τ, X_τ) is a sufficient statistic for $\mathscr{F}_\tau$.

The formulation of this theorem is a slight generalization of Theorem 11, [5], which is formulated only for real-valued random processes X. The proof would be the same as in [5]. We shall only sketch it here. It is based on a slight modification of the above-mentioned Theorem 1, [7]:

According to Theorems 1 and 2 of Section 4 the Radon–Nikodym derivative of $P_\theta | \mathscr{F}_{\tau<\infty}$ with respect to $P^* | \mathscr{F}_{\tau<\infty}$ exists and is equal to the function $f_{\theta,\tau}$ defined in Theorem 2, which because of the definition of g_θ turns out to be $g_\theta(\tau, X_\tau)$. Since g_θ is continuous in the sense of (11), it is $\mathscr{B}_S \otimes \mathscr{E}$-measurable. Thus $g_\theta(\tau, X_\tau)$ is measurable with respect to $(\tau < \infty) \cap \sigma(\tau, X_\tau)$. Therefore, according to Theorem 1, [7], the random function (τ, X_τ) is sufficient for $\mathscr{F}_\tau$. (The exclusion of the event $(\tau = \infty)$ in the proof does not affect the result, see [5].)

Theorem 4 can be reformulated without topological assumptions about the process $\boldsymbol{X}$, but then $(\Omega, \mathscr{A})$ has to be a Blackwell space (as defined by Meyer [8]):

THEOREM 5. *Let $(\Omega, \mathscr{A})$ be a Blackwell space, and let $\mathscr{P}$ be dominated on $\mathscr{F}_t$ for every $t \in S$. In addition to the assumptions on $\boldsymbol{X}$ stated at the beginning of the section, suppose the σ-algebra $\mathscr{E}$ to be countably generated and x to be progressively measurable with respect to $\boldsymbol{F}$.*

If there exists a collection g_θ, $\theta \in \Theta$, of likelihood-generating functions such that for every $\theta \in \Theta$ and every $\omega \in \Omega$ the mapping $h_{\theta,\omega}$: $S \to S$ defined by $h_{\theta,\omega}(t) = g_\theta(t, X_t(\omega))$ is right-continuous, then for every random size of observation τ the function (τ, X_τ) is a sufficient statistic for $\mathscr{F}_\tau$.

This theorem is a corrected version of Theorem 12, [5]. The proof differs from that of Theorem 4 only in the following argument. We cannot assume a priori that $g_\theta(\tau, X_\tau)$ is $(\tau < \infty) \cap \sigma(\tau, X_\tau)$-measurable, because we have not assumed that g_θ is continuous in the sense of (11). Nevertheless this measurability follows from the fact that $(\Omega, \mathscr{A})$ is a Blackwell space and from the conclusion (b) to Theorem III, 17, [8].

Let us finish this section with the following two remarks.

1. In Section 4 we mentioned that, as a consequence of the strong law of large numbers, domination on $\mathscr{F}_\infty$ in general cannot be assumed. But, on the other hand, it is the strong law of large numbers too which guarantees the existence of a sufficient statistic X_∞ for $\mathscr{F}_\infty$ (see [5]). Thus the assumptions of Theorems 3, 4 and 5 are not in contradiction with practical situations.

2. In Definition 2 we took into account any kind of randomization in generating the r.o. τ. This is the essential reason why we cannot expect (τ, X_τ) to be minimal sufficient for $\mathscr{F}_\tau$, as we can see in the following trivial example. Let every $\mathscr{F}_t$ be equal to the σ-algebra $\{\emptyset, \Omega\}$, $t \in T$, and suppose that τ is a nondegenerated random time whose distribution does not depend on $\theta \in \Theta$. Finally assume $X_t \equiv 0$, $t \in T$. Then

– τ is an r.o. with respect to $\boldsymbol{F}$.
– $\mathscr{F}_\tau = \sigma(\tau)$,
– $\sigma(\tau, X_\tau) = \sigma(\tau)$.

But since $\{\emptyset, \Omega\}$ is minimal sufficient for $\mathscr{F}_\tau$, the σ-algebra $\sigma(\tau, X_\tau)$ is not minimal sufficient.

6. Generalization of a lemma of Sudakov

The theorems presented in Sections 4 and 5 enable us to formulate a well-known lemma of Sudakov ([12]) under more general conditions. Suppose that $X = (X_t)_{t \in T}$ is the process and (τ, X_τ) the mapping considered in Section 5.

LEMMA. *Let all conditions of Theorem* 4 *including* (11) *be fulfilled, and let τ be a random size of observation. Denote by m_θ the measure induced by the mapping (τ, X_τ) on the measurable space $(S \times E, \mathscr{B}_S \otimes \mathscr{E})$ given P_θ, $\theta \in \Theta$* (i.e., $m_\theta(B) = P_\theta\big((\tau, X_\tau)^{-1}(B)\big)$, $B \in \mathscr{B}_S \otimes \mathscr{E}$), *and by m^* the measure induced by (τ, X_τ) given P^*.*

Then we have for every $\theta \in \Theta$

(i) *m_θ is absolutely continuous with respect to m^*;*

(ii) *the Radon–Nikodym derivative of m_θ with respect to m^* is determined by*

$$\frac{dm_\theta}{dm^*}(t, x) = g_\theta(t, x), \qquad t \in S,\ x \in E.$$

Instead of the domination of $\mathscr{P}$ on every $\mathscr{F}_t$, $t \in S$, in [12] it is assumed that there exists a $\theta_0 \in \Theta$ such that for every $t \in S$ the family $\{m_{\theta,t};\ \theta \in \Theta\}$ of distributions induced by X_t on $(E, \mathscr{E})$ is dominated by $m_{\theta_0,t}$. We could easily show that this assumption together with the sufficiency of X_t for $\mathscr{F}_t$ implies the domination of $\mathscr{P}$ on $\mathscr{F}_t$ by $P_{\theta_0} | \mathscr{F}_t$, $t \in S$. This means that our domination assumption (9) is not stronger than the corresponding one in [12]; it is even weaker, because in (10) we did not suppose that the dominating measure P^* is a particular measure P_{θ_0} belonging to $\mathscr{P}$. Our formulation of the lemma differs from that in [12] further in the admission of arbitrary random sizes of observation instead of Markov times and in the consideration of general processes X. Finally, we suppose neither that $\mathscr{F}_t$ is equal to $\sigma(X_s;\ s \leqslant t)$, $t \in S$, nor that $P_\theta(\tau = \infty) = 0$, $\theta \in \Theta$ (i.e., we admit $m_\theta(S \times E) < 1$).

Proof of the lemma. For the sake of shortness we shall write $\sigma(\tau, X_\tau)$ instead of $(\tau < \infty) \cap \sigma(\tau, X_\tau)$. According to Theorem 1, $\mathscr{P}$ is dominated on $\mathscr{F}_{\tau<\infty}$ by $P^* | \mathscr{F}_{\tau<\infty}$. Applying Theorem 2, we get

$$\frac{dP_\theta}{dP^*}\Big| \mathscr{F}_{\tau<\infty} = g_\theta(\tau, X_\tau), \qquad P^*\text{-a.s.},\ \theta \in \Theta.$$

(The right-hand side is understood to be defined on $(\tau < \infty)$.)

As mentioned in the proof of Theorem 4 under condition (11) of that theorem, the mapping g_θ is $\mathscr{B}_S \otimes \mathscr{E}$-measurable, and therefore $g_\theta(\tau, X_\tau)$ is $\sigma(\tau, X_\tau)$-measurable. Thus $g_\theta(\tau, X_\tau)$ is a version of the Radon–Nikodym derivative of $P_\theta | \sigma(\tau, X_\tau)$ with respect to $P^* | \sigma(\tau, X_\tau)$, i.e., for every $B \in \mathscr{B}_S \otimes \mathscr{E}$ we have

$$P_\theta\big((\tau, X_\tau)^{-1}(B)\big) = \int_{(\tau, X_\tau)^{-1}(B)} g_\theta(\tau, X_\tau)\, dP^*.$$

By the integral transformation formula (for instance [7], Lemma 3) we get the desired equation, which concludes the proof,

$$m_\theta(B) = \int_B g_\theta(t, x)\, dm^*. \qquad \blacksquare$$

Acknowledgment. The author wishes to thank Dr. H.-J. Engelbert for helpful discussions, which led to the formulation of Theorem 5.

References

[1] R. R. Bahadur, *Sufficiency and statistical decision functions*, Ann. Math. Statist. **25** (1954), 423–462.

[2] K. L. Chung and J. L. Doob, *Fields, optionality and measurability*, Amer. J. Math. **87** (1965), 397–424.

[3] A. P. Dawid, *Conditional independence in statistical theory* (*with discussion*), J. Roy. Statist. Soc. Ser. B **41** (1979), 1–31.

[4] R. Döhler, *Untersuchungen zu randomisierten Zeiten*, Dissertation, Technische Hochschule Karl-Marx-Stadt, Karl-Marx-Stadt 1975.

[5] —, *Dominierbarkeit und Suffizienz in der Sequentialanalyse*, Math. Operationsforsch. Statist. Ser. Statist. **12** (1981), 101–134.

[6] E. B. Dynkin and A. A. Jushkevitch, *Controlled Markov processes and their applications*, Nauka, Moscow 1975, in Russian.

[7] P. R. Halmos and L. J. Savage, *Applications of the Radon–Nikodym theorem to the theory of sufficient statistics*, Ann. Math. Statist. **20** (1949), 225–241.

[8] P. A. Meyer, *Probabilités et potentiel*, Hermann, Paris 1966.

[9] J. W. Pitman and T. P. Speed, *A note on random times*, Stochastic Process. Appl. **1** (1973), 369–374.

[10] A. N. Shiryayev, *Statistical Sequential Analysis*, 2nd ed., Nauka, Moscow 1967, in Russian.

[11] D. O. Siegmund, *Some problems in the theory of optimal stopping rules*, Ann, Math. Statist. **38** (1967), 1627–1640.

[12] V. N. Sudakov, *On measures defined by Markov times*, Scientific Seminar LOMI **12** (1969), 157–164, in Russian.

Presented to the semester
Sequential Methods in Statistics
September 7–December 11, 1981