The Annals of Probability
1978, Vol. 6, No. 5, 705-730

SPECIAL INVITED PAPER

SUFFICIENT STATISTICS AND EXTREME POINTS[1]

BY E. B. DYNKIN

Cornell University

A convex set M is called a simplex if there exists a subset M_e of M such that every $P \in M$ is the barycentre of one and only one probability measure μ concentrated on M_e. Elements of M_e are called extreme points of M. To prove that a set of functions or measures is a simplex, usually the Choquet theorem on extreme points of convex sets in linear topological spaces is cited. We prove a simpler theorem which is more convenient for many applications. Instead of topological considerations, this theorem makes use of the concept of sufficient statistics.

1. Introduction.

1.1. If M is a simplex in a finite-dimensional linear space, the set M_e of extreme points is finite, and to say that $\bar{P}$ is a barycentre of a probability measure μ concentrated on M_e means that

$$\bar{P} = \sum_{P \in M_e} \mu(P)P ,$$

where $\mu(P) \geqq 0$ for all $P \in M_e$ and $\sum_{P \in M_e} \mu(P) = 1$. The concept of a barycentre can be naturally extended to probability measures on spaces of functions and measures. Simplexes in such spaces play an important role in various fields of mathematics. Here are some examples:

1.1.A. The set of all probability measures invariant with respect to a measurable transformation T of a measurable space $(\Omega, \mathscr{F})$. (Extreme points are ergodic measures.)

1.1.B. The set of all Gibbs states specified by a given family of conditional distributions.

1.1.C. The set of all symmetric probability measures on a product space (with infinite number of factors). Extreme points are product measures.

1.1.D. The set of all Markov processes with a given transition function.

1.1.E. The set of all stationary probability distributions for a given stationary transition function.

1.1.F. The class of all normed excessive functions associated with a given transition function. A particular case is the class of all positive superharmonic functions h in a domain D of a Euclidean space normed by the condition $h(c) = 1$,

Received August 3, 1977.

[1] Research supported by NSF Grant No. MCS 77-03543.

AMS 1970 *subject classifications*. Primary 60-02; Secondary 60J50, 60K35, 82A25, 28A65.

Key words and phrases. Extreme points, sufficient statistics, Gibbs states, ergodic decomposition of an invariant measure, symmetric measures, entrance and exit laws, excessive measures and functions.

where c is a fixed point of D. This class is associated with the Brownian motion in D.

These classes were treated by many authors from different points of view. We mention here the works of Krylov and Bogolubov [13] (related to the class 1.1.A); Dobrushin [2] (class 1.1.B); de Finetti [9], [10]; and Hewitt–Savage [11] (1.1.C); Martin [15]; Doob [3]; and Hunt [12] (1.1.F).

In the present paper, all these classes of measures and functions and some others will be investigated by constructing suitable sufficient statistics.

1.2. The role of a special type of sufficient statistics (we call them H-sufficient) is revealed by Theorem 3.1. This theorem was first published in 1971 ([4], Section 2) in a slightly different form and without explicitly mentioning sufficient statistics. The theorem was applied to the class of all Markov processes with a given transition function (class 1.1.D) in [4] and to excessive measures and excessive functions (1.1.F) in [5].

We start with general definitions of a barycentre, extreme points, etc., in Section 2. Relations between H-sufficient statistics and decomposition into extreme points are investigated in Section 3. The main method of constructing H-sufficient statistics is a special kind of passage to the limit which is studied in Section 4. The rest of the paper is devoted to various applications. In particular, Sections 9—12 contain an improved version of the results on Markov processes published in [4] and [5]. The presentation is self-contained, but we refer to [5] for some technical details.

2. Convex measurable spaces.

2.1. Let $(M, \mathscr{B}_M)$ be an arbitrary measurable space. We say that a *convex structure* is introduced into M if a point P_μ, the barycentre of μ, is associated with each probability measure μ on $\mathscr{B}_M$. A space $(M, \mathscr{B}_M)$ provided with such a structure will be called a *convex measurable* space.

We say that P is an *extreme point* of M, and write $P \in M_e$, if P is not a barycentre of any measure μ except the measure concentrated on P. A convex measurable space M is called a *simplex* if M_e is measurable and each $P \in M$ is a barycentre of one and only one probability measure μ concentrated on M_e.

Let $(M, \mathscr{B}_M)$ and $(M', \mathscr{B}_{M'})$ be convex measurable spaces and let T be a mapping of M into M'. We say that *T preserves the convex structure* if T is measurable and transforms the barycentre of a measure μ into the barycentre of the measure

$$\mu'(\Gamma) = \mu(T^{-1}\Gamma), \quad \Gamma \in \mathscr{B}_{M'}.$$

We say that T is an *isomorphism* if it is invertible and T and T^{-1} preserve convex structure.

An axiomatic theory of convex measurable spaces can be developed but our task is rather an analysis of concrete spaces.

2.2. Let M be a collection of positive functions on an arbitrary set Z. (By a positive function we mean a function with values in an extended real half-line $[0, +\infty]$.) Let $\mathscr{B}_M$ be an arbitrary σ-algebra in M with the property:

2.2.A. For each $z \in Z$, the function $F_z(\varphi) = \varphi(z)$ is $\mathscr{B}_M$-measurable.

Let μ be a probability measure on $\mathscr{B}_M$. We define a *barycentre* φ_μ of μ by the formula

$$\varphi_\mu(z) = \int_M \varphi(z)\mu(d\varphi) . \tag{2.1}$$

If M contains the barycentres of all probability measures, it is a convex measurable space.

A measurable structure in M is called *natural* if it is determined by the minimal σ-algebra $\mathscr{B}_M$ with the property 2.2.A. Unless otherwise stipulated we consider in M the natural measurable structure, and we always consider in M the convex structure defined by formula (2.1).

Formula (2.1) makes sense also for finite nonprobabilistic measures μ. In this case, we call φ_μ a *generalized barycentre* of μ. If M contains all generalized barycentres, we say that M is a *convex cone*.

2.3. Now let M be a set of probability measures on a measurable space $(\Omega, \mathscr{F})$. The set M can be considered also as a class of positive functions on $\mathscr{F}$, and we can apply all the definitions of Subsection 2.2.

If M is a simplex, the formula

$$\bar{P}(A) = \int_{M_e} P(A)\mu(dP) \tag{2.2}$$

establishes a one-to-one correspondence between M and the set of all probability measures on M_e.

We consider one example. Let $M(\mathscr{F})$ be the class of all probability measures on a σ-algebra $\mathscr{F}$. It is easy to check, step by step, that:

(i) $M(\mathscr{F})$ is convex.

(ii) Measures $Q^\omega(A) = 1_A(\omega)$, $A \in \mathscr{F}$ are extreme points of $M(\mathscr{F})$.

(iii) Each $P \in M(\mathscr{F})$ is a barycentre of a measure μ defined by formula

$$\mu(\Gamma) = P\{\omega : Q^\omega \in \Gamma\} . \tag{2.3}$$

This measure is concentrated on the set $M_e(\mathscr{F})$ of extreme points of $M(\mathscr{F})$.

(iv) If P is an extreme point, then $P = Q^\omega$ for some ω.

(v) If μ is a measure concentrated on $M_e(\mathscr{F})$ and P is a barycentre of μ, then μ and P satisfy (2.3).

(vi) $M(\mathscr{F})$ is a simplex.

We prove all these statements in a much more general situation in Section 3.

2.4. We shall use the following abbreviations. If f is a function and $\mathscr{F}$ is a σ-algebra, then the expression $f \in \mathscr{F}$ means that f is $\mathscr{F}$-measurable and bounded. An expression Pf (or $P(f)$) means an integral of f with respect to a measure P.

Let M be a class of probability measures on $(\Omega, \mathscr{F})$. A set A is called *M-null* if $A \in \mathscr{F}$ and $P(A) = 0$ for all $P \in M$. We say that $A, B \in \mathscr{F}$ are *P-equivalent* if

$1_A = 1_B$ a.s. P. Two σ-algebras $\mathscr{A}$, $\mathscr{B} \subset \mathscr{F}$ are *M-equivalent* if, for each $P \in M$, every $A \in \mathscr{A}$ is P-equivalent to a $B \in \mathscr{B}$ and vice versa.

3. *H*-sufficiency and the decomposition into extreme points.

3.1. Let M be an arbitrary class of probability measures on a measurable space $(\Omega, \mathscr{F})$. We say that M is *separable* if $\mathscr{F}$ contains a countable family $\mathscr{A}$ separating the measures in M (which means that for each pair of different elements P_1, P_2 of M there exists $A \in \mathscr{A}$ such that $P_1(A) \neq P_2(A)$). The class $M(\mathscr{F})$ is separable if $\mathscr{F}$ is *countably generated* (i.e., generated by a countable family of sets).

A σ-algebra $\mathscr{F}^0 \subset \mathscr{F}$ is called *sufficient* for M if all measures $P \in M$ have a common conditional distribution relative to $\mathscr{F}^0$; in other words, if for each $\omega \in \Omega$ there exists a probability measure Q^ω on $\mathscr{F}$ such that, for each A, $Q^\omega(A)$ is $\mathscr{F}$-measurable and

$$P(A \mid \mathscr{F}^0) = Q^\omega(A) \quad \text{a.s.} \quad P \qquad \text{for all} \quad P \in M. \tag{3.1}$$

A sufficient σ-algebra will be called *H-sufficient* if, in addition,

$$Q^\omega \in M \quad \text{a.s.} \quad M \tag{3.2}$$

(which means that $P(Q^\omega \in M) = 1$ for all $P \in M$).

If $\mathscr{F}^1$ is M-equivalent to $\mathscr{F}^0$ and if $\mathscr{F}^0$ is sufficient (H-sufficient) for M, then so is $\mathscr{F}^1$.

THEOREM 3.1. *Let $\mathscr{F}^0$ be an H-sufficient σ-algebra for a separable class M. Then the set M_e of extreme points of M is measurable and each $P \in M$ is a barycentre of one and only one probability measure μ_P concentrated on M_e. If M is convex, it is a simplex.*

Let Q^ω be measures satisfying (3.1) *and* (3.2). *Then M_e is a subset of a set $\{Q^\omega\}$ and the measure μ_P is given by formula*

$$\mu_P(\Gamma) = P\{\omega : Q^\omega \in \Gamma\}. \tag{3.3}$$

A measure $P \in M$ belongs to M_e if and only if

$$P\{\omega : Q^\omega = P\} = 1. \tag{3.4}$$

PROOF. 1°. We start with the following elementary observation: If P is any probability measure on a σ-algebra $\mathscr{F}$ and if $\mathscr{F}^0$ is any subalgebra of $\mathscr{F}$, then the conditions (i), (ii), (iii) are equivalent:

(i) P is trivial on $\mathscr{F}^0$.
(ii) Each $\mathscr{F}^0$-measurable function Z is constant a.s. P.
(iii) $P\{P(A \mid \mathscr{F}^0) \neq P(A)\} = 0$ for each $A \in \mathscr{F}$.

2°. Denote by M_0 the set of all measures $P \in M$ which are trivial on $\mathscr{F}^0$. According to 1°, M_0 can be described by the condition (iii). Taking into account (3.1), we rewrite (iii) in the form

$$P\{Q^\omega(A) \neq P(A)\} = 0 \qquad \text{for all} \quad A \in \mathscr{F}. \tag{3.5}$$

Let $\mathscr{A}$ be a countable family of sets separating measures of M. Obviously (3.5) implies that

$$P\{Q^\omega(A) \neq P(A) \text{ for all } A \in \mathscr{A}\} = 0 . \tag{3.6}$$

Since P and Q^ω belong to M, (3.6) implies that

$$P\{Q^\omega \neq P\} = 0 . \tag{3.7}$$

It is clear that (3.5) follows from (3.7); hence each of the conditions (3.5), (3.6) and (3.7) characterizes the set M_0. The condition (3.5) can be rewritten also in the following form:

$$f_A(P) = 0 \qquad \text{for all } A \in \mathscr{A} , \tag{3.8}$$

where

$$f_A(P) = \int_\Omega Q^\omega(A)^2 P(d\omega) - P(A)^2 = \int [Q^\omega(A) - P(A)]^2 P(d\omega) . \tag{3.9}$$

Evidently, f_A is $\mathscr{B}_M$-measurable. Therefore $M_0 \in \mathscr{B}_M$. It follows from (3.7) that for each $P \in M_0$ there exists $\omega \in \Omega$ such that $P = Q^\omega$.

3°. Now we prove that

$$Q^\omega \in M_0 \quad \text{a.s.} \quad M . \tag{3.10}$$

It follows from (3.1) that $Q^\omega Y = P(Y \mid \mathscr{F}^0)$ a.s. P. Setting $Y_A = Q^\omega(A)^2$, we conclude from (3.9) that

$$f_A(Q^\omega) = Q^\omega Y_A - Y_A = P(Y_A \mid \mathscr{F}^0) - Y_A ,$$

and hence

$$Pf_A(Q^\omega) = 0 . \tag{3.11}$$

But it is clear from (3.9) that $f_A \geqq 0$. Therefore (3.11) implies that $f_A(Q^\omega) = 0$ a.s. P. We see that, for almost all ω, the measure Q^ω satisfies the condition (3.8) which implies that $Q^\omega \in M_0$.

4°. Let a measure μ_P be defined by formula (3.3). Then the formula

$$\int_M F(\tilde{P})\mu_P(d\tilde{P}) = \int_\Omega F(Q^\omega)P(d\omega) \tag{3.12}$$

holds for indicator functions $F = 1_\Gamma$, $\Gamma \in \mathscr{B}_M$. Standard arguments show that (3.12) is true for all bounded $\mathscr{B}_M$-measurable functions F. For $F(\tilde{P}) = \tilde{P}(A)$, $A \in \mathscr{F}$ the right side of (3.12) is equal to $P(A)$. Thus P is a barycentre of μ_P. According to 3°, μ_P is concentrated on M_0.

5°. Now let $P \in M$ be a barycentre of a measure μ concentrated on M_0. For every $\Gamma \subset M_0$, $\Gamma \in \mathscr{B}_M$

$$P\{Q^\omega \in \Gamma\} = \int_{M_0} \tilde{P}(Q^\omega \in \Gamma)\mu(d\tilde{P}) . \tag{3.13}$$

The left side is equal to $\mu_P(\Gamma)$. By (3.7) $\tilde{P}(Q^\omega \in \Gamma) = 1_\Gamma(\tilde{P})$ for $\tilde{P} \in M_0$. Therefore the right side of (3.13) is equal to

$$\int_{M_0} 1_\Gamma(\tilde{P})\mu(d\tilde{P}) = \mu(\Gamma) .$$

Hence $\mu_P = \mu$.

6°. Let $P \in M_e$. According to 4°, P is a barycentre of μ_P. Therefore μ_P is concentrated on P which means that $P\{Q^\omega \neq P\} = 0$, and $P \in M_0$ by 2°.

7°. Now let $\bar{P} \in M_0$ be a barycentre of a measure μ on M. According to 2°, $\bar{P}\{Q^\omega \neq \bar{P}\} = 0$. Hence μ is concentrated on the set $M' = \{P: P(Q^\omega \neq \bar{P}) = 0\}$. But if $P \in M'$, $C \in \mathscr{F}$, then $P(C) = PP(C \mid \mathscr{F}^0) = \int Q^\omega(C)P(d\omega) = \bar{P}(C)$. Therefore $M' = \{\bar{P}\}$ and μ is concentrated on $\bar{P}$. This proves that $\bar{P} \in M_e$.

3.2.

THEOREM 3.2. *Let a separable class M have an H-sufficient σ-algebra and let $\mathscr{F}^1$ be the class of all sets $A \in \mathscr{F}$ with the following property:*

$$(3.14) \qquad P(A) = 0 \quad or \quad P(A) = 1 \quad for\ all \quad P \in M_e\,.$$

Then a σ-algebra $\mathscr{F}^0$ is H-sufficient for M if and only if it is M-equivalent to $\mathscr{F}^1$.

PROOF. We need only to prove that each H-sufficient σ-algebra $\mathscr{F}^0$ is M-equivalent to $\mathscr{F}^1$. By Theorem 3.1, $\mathscr{F}^0 \subset \mathscr{F}^1$. Therefore it is sufficient to construct, for every fixed $P \in M$, $A \in \mathscr{F}^1$, a set $B \in \mathscr{F}^0$ which is P-equivalent to A. A function $Q^\omega(A)$ is P-equivalent to a $\mathscr{F}^0$-measurable function f. Sets $B = \{\omega : f(\omega) = 1\}$ and $C = \{\omega : f(\omega) = 0\}$ belong to $\mathscr{F}^0$, and

$$1_B + 1_C = 1 \quad \text{a.s.} \quad P\,,$$

$$P(BA) = P1_B Q^\omega(A) = P(B)\,, \qquad P(CA) = P1_C Q^\omega(A) = 0\,.$$

Hence $1_A = 1_A 1_B = 1_B$ a.s. P. Our theorem is proved.

Now suppose that a class M is a simplex and let $\mathscr{F}^1$ be defined by (3.14). It is clear that

$$(3.15) \qquad P(A \mid \mathscr{F}^1) = P(A) \qquad \text{for each} \quad P \in M_e\,.$$

Therefore $\mathscr{F}^1$ is H-sufficient for M_e (and consequently for M) if and only if a measurable mapping $\omega \to Q^\omega$ of $(\Omega, \mathscr{F}^1)$ into M_e exists such that $P(Q^\omega = P) = 1$ for all $P \in M_e$. In this case, every two measures of M_e are singular on $\mathscr{F}^1$ with respect to each other. If M_e is at most countable, this condition is not only necessary but also sufficient: It implies the existence of decomposition of Ω into the sets $\Omega_P \in \mathscr{F}^1$, $P \in M_e$ with the property $P(\Omega_P) = 1$, and the mapping Q^ω can be defined by formula $Q^\omega = P$ for $\omega \in \Omega_P$.

3.3. We discuss now the concept of H-sufficiency from a slightly different, more algebraic point of view.

A real-valued function $Q^\omega(A) = Q(\omega, A)$, $\omega \in \Omega$, $A \in \mathscr{F}$ is called a *Markov kernel* if, for each $\omega \in \Omega$, $Q(\omega, \cdot)$ is a probability measure and, for each $A \in \mathscr{F}$, $Q(\cdot, A)$ is an $\mathscr{F}$-measurable function. A linear operator on the space of bounded $\mathscr{F}$-measurable functions and a linear operator on the space $M(\mathscr{F})$ of all probability measures are associated with every Markov kernel Q. We denote them by the same letter and call them *Markov operators*. They are defined by the formulas

$$(3.16) \qquad Qf(\omega) = \int Q(\omega, d\omega')f(\omega') = Q^\omega(f)\,,$$

and

$$(PQ)(A) = \int P(d\omega)Q(\omega, A) . \tag{3.17}$$

We shall consider the first operator also on unbounded functions f (in this case Qf is defined only on a part of Ω). The second operator can be extended too: the formula (3.17) makes sense not only for $P \in M(\mathscr{F})$ but also for $P \in M(\mathscr{F}')$ if $\mathscr{F}'$ is a σ-algebra with the property that $Qf \in \mathscr{F}'$ for all $f \in \mathscr{F}$. Two Markov operators Q and Q' are called *M-equivalent* if $Qf = Q'f$ a.s. M for all $f \in \mathscr{F}$.

We say that a set $A \in \mathscr{F}$ is *Q-invariant* if $Q1_A = 1_A$.

LEMMA 3.1. *If all sets of a σ-algebra $\mathscr{F}^0$ are Q-invariant, then*

$$Q(gf) = gQf \tag{3.18}$$

for each $f \in \mathscr{F}$, $g \in \mathscr{F}^0$ and

$$P\{f \mid \mathscr{F}^0\} = P\{Qf \mid \mathscr{F}^0\} \quad \text{a.s.} \quad P \tag{3.19}$$

for every Q-invariant measure P and every $f \in \mathscr{F}$.

PROOF. It suffices to check (3.18) for $f = 1_A$, $g = 1_B$ where $A \in \mathscr{F}$, $B \in \mathscr{F}^0$. In this case

$$Q(gf) - gQf = (1 - g)Q(gf) - gQ[(1 - g)f]$$

and

$$0 \leqq (1 - g)Q(gf) \leqq (1 - g)Qg = 0 ,$$
$$0 \leqq gQ[(1 - g)f] \leqq g(1 - g) = 0 .$$

Formula (3.19) follows immediately from (3.18).

A Markov operator Q is called a *sufficient statistic* for M if there exists a σ-algebra $\mathscr{F}^0 \subset \mathscr{F}$ such that

$$P(f \mid \mathscr{F}^0) = Qf \quad \text{a.s.} \quad P \tag{3.20}$$

for all $P \in M$ and all $f \in \mathscr{F}$. If, in addition, (3.2) holds, we say that Q is *H-sufficient* for M. Obviously (3.20) is equivalent to (3.1).

If Q is a sufficient (or an H-sufficient) statistic for M, then so are all operators M-equivalent to Q.

THEOREM 3.3. *If a convex separable class M has an H-sufficient statistic, then there exists an H-sufficient statistic Q, such that*

$$Q(fQg) = QfQg \quad \textit{for all} \quad f, g \in \mathscr{F} \tag{3.21}$$

and M coincides with the class of all Q-invariant measures.

Every Markov operator Q with the property (3.21) *is H-sufficient for the class M of all Q-invariant measures. The corresponding H-sufficient σ-algebra $\mathscr{F}^0$ can be defined as the collection of all Q-invariant sets. A mapping $P \to P^0$, where P^0 is the restriction of the probability measure P to $\mathscr{F}^0$, is an isomorphism of M onto $M(\mathscr{F}^0)$. The inverse mapping is given by the formula $P = P^0Q$.*

PROOF. 1°. Let $\tilde{Q}^\omega$ be an H-sufficient statistic for M. By Theorem 3.1 $\Omega_1 = \{\omega : Q^\omega \bar{\in} M_e\}$ is an M-null set. Hence an operator

$$\begin{aligned} Qf &= \tilde{Q}f && \text{on } \Omega_1^c , \\ &= \tilde{Q}f(\omega^*) && \text{on } \Omega_1 , \end{aligned}$$

where ω^* is a fixed point of Ω_1^c, is H-sufficient for M too.

By (3.4), for all $\omega_1 \in \Omega_1^c$, $Q^{\omega_1}\{\omega : Q^\omega = Q^{\omega_1}\} = 1$, and

$$\begin{aligned} Q(fQg)(\omega_1) &= \int f(\omega)Q^\omega(g)Q^{\omega_1}(d\omega) = \int f(\omega)Q^{\omega_1}(g)Q^{\omega_1}(d\omega) \\ &= Q^{\omega_1}(g)Q^{\omega_1}(f) = Qf(\omega_1)Qg(\omega_1) . \end{aligned}$$

2°. It follows from (3.20) that $PQf = Pf$ for all $P \in M$, $f \in \mathscr{F}$. Therefore all $P \in M$ are Q-invariant. On the other hand, if P is Q-invariant, then

$$P(A) = \int P(d\omega)Q^\omega(A)$$

and $P \in M$ since $Q^\omega \in M$ for all $\omega \in \Omega$.

3°. Let Q be a Markov operator with the property (3.21) and let $\mathscr{F}^0$ be the totality of all Q-invariant sets. It is easy to see that $\mathscr{F}^0$ is a σ-algebra. By Lemma 3.1, all functions $f \in \mathscr{F}^0$ are Q-invariant. To prove the converse, we denote by H the class of all measurable transformations Φ of the real line such that $\Phi(f)$ is Q-invariant for every Q-invariant f. The class H contains linear functions and is closed under addition and monotone convergence. By virtue of (3.21), it is closed also under multiplication. Therefore it contains all bounded Borel functions, in particular, functions $\Phi(u) = 1_{u<c}$ for all constant c. Hence for each Q-invariant f, the sets $\{\omega : f(\omega) > c\}$ belong to $\mathscr{F}^0$, and f is $\mathscr{F}^0$-measurable.

4°. Setting $f = 1$ in (3.21), we see that $Q^2 = Q$. Hence $Qf \in \mathscr{F}^0$ for all $f \in \mathscr{F}$. The identity (3.21) implies that $Q(gf) = gQf$ for $f \in \mathscr{F}$, $g \in \mathscr{F}^0$. Hence, for each Q-invariant measure P,

$$P(gQf) = PQ(gf) = P(gf) ,$$

and (3.20) is satisfied; Q is a sufficient statistic for the class M of all Q-invariant measures and $\mathscr{F}^0$ is the corresponding sufficient σ-algebra. On the other hand, the identity $Q^2 = Q$ implies that $Q^\omega \in M$ of all ω, and Q is H-sufficient.

Since $Qf \in \mathscr{F}^0$ for all $f \in \mathscr{F}$, an equality $PQ = P$ implies that $P^0Q = P$ where P^0 is a restriction of $P \in M$ to $\mathscr{F}^0$. Obviously $P^0Q \in M$ for every $P^0 \in M(\mathscr{F}^0)$. Therefore we have a one-to-one correspondence between M and $M(\mathscr{F}^0)$. It is easy to check that this correspondence is an isomorphism in the sense of Subsection 2.1.

3.4. We shall prove that under certain circumstances sufficiency implies the H-sufficiency.

A family of Markov operators V_t satisfying the condition $V_sV_t = V_{s+t}$ for all s, t is called a one-parameter semigroup if t takes values on the positive real half-line, and it is called a one-parameter group if t takes values on the real line.

We say that V_t is measurable if, for each $f \in \mathscr{F}$, the function $V_t f(\omega)$ is measurable with respect to the pair t, ω (the measurable structure on Ω is given by $\mathscr{F}$ and on the real line by the σ-algebra $\mathscr{B}$ of all Borel sets).

THEOREM 3.4. *Let $\mathscr{V}$ be a finite or countable family of Markov operators or a measurable one-parameter semigroup or group in* $(\Omega, \mathscr{F})$ *and let $\mathscr{F}$ be countably generated. Suppose that $\mathscr{F}^0 \subset \mathscr{F}$ is sufficient for the class M of all $\mathscr{V}$-invariant measures and* (3.19) *holds for all $P \in M$, $Q \in \mathscr{V}$. Then $\mathscr{F}^0$ is H-sufficient for M.*

PROOF. Consider a Markov operator Q satisfying condition (3.20). To prove (3.2), we need only to check that for each $P \in M$ and each $f \in \mathscr{F}$

$$QVf = Qf \quad \text{for all} \quad V \in \mathscr{V} \quad \text{a.s.} \quad P\,. \tag{3.22}$$

(Indeed (3.22) implies that, for almost all ω, all measures $Q^\omega V$, $V \in \mathscr{V}$ coincide with Q^ω on a countable family of sets separating measures of $M(\mathscr{F})$ and therefore coincide everywhere.)

It follows from (3.19) that

$$QVf = Qf \quad \text{a.s.} \quad P \quad \text{for every} \quad V \in \mathscr{V}\,. \tag{3.23}$$

If $\mathscr{V}$ is at most countable, then (3.23) implies (3.22) and our theorem is proved.

In the case of a Markov semigroup or a group, we consider the set $A = \{(t, \omega) : Q^\omega V_t = Q^\omega\}$. It follows from (3.23) that for each t, $P\{\omega : (t, \omega) \in A\} = 1$. The set A belongs to $\mathscr{B} \times \mathscr{F}$. By Fubini's theorem there exists a set Ω_1 such that $P(\Omega_1) = 1$ and, if $\omega \in \Omega_1$, then, for almost all t, $(t, \omega) \in A$, that is, $Q^\omega V_t = Q^\omega$. Taking into account that $V_s V_t = V_{s+t}$ for all s, t, we easily prove that $Q^\omega V_t = Q^\omega$ for all $\omega \in \Omega_1$ and all t.

REMARK. Theorem 3.4 and its proof are valid for a group $\mathscr{G}$ of Markov operators if there exists a σ-algebra $\mathscr{B}_\mathscr{G}$ in $\mathscr{G}$ and a σ-finite measure λ on $\mathscr{B}_\mathscr{G}$ such that: (i) $Vf(\omega)$ is $\mathscr{B}_\mathscr{G} \times \mathscr{F}$-measurable for each $f \in \mathscr{F}$; (ii) $\lambda(VB) = \lambda(B)$ for each $B \in \mathscr{B}_\mathscr{G}$ and each $V \in \mathscr{G}$.

3.5. Let $\mathscr{F}_1$ and $\mathscr{F}_2$ be sufficient σ-algebras for a class M and let Q_1 and Q_2 be correspondent sufficient statistics. It is easy to see that $Q_1 Q_2 = Q_2 Q_1$ a.s. M for all $f \in \mathscr{F}$ if and only if $\mathscr{F}_1$ and $\mathscr{F}_2$ are conditionally independent given $\mathscr{F}^0 = \mathscr{F}_1 \cap \mathscr{F}_2$. In this case $\mathscr{F}^0$ is a sufficient σ-algebra for M and $Q_{12} = Q_1 Q_2$ and $Q_{21} = Q_2 Q_1$ are corresponding sufficient statistics.

Now let M be a convex class and let Q_1 and Q_2 be H-sufficient. The set $\Omega_0 = \{\omega' : Q_2^{\omega'} \bar{\in} M\}$ is M-null and therefore

$$Q_{12}^\omega(A) = \int_{\Omega_0^c} Q_1^\omega(d\omega') Q_2^{\omega'}(A)\,.$$

Hence $Q_{12}^\omega \in M$ if $Q_1^\omega \in M$, and Q_{12} is H-sufficient for M.

4. Asymptotic sufficiency.

4.1. We say that a sequence of Markov operators Q_n converges M-almost surely to a Markov operator Q and write $Q_n \to Q$ a.s. M if for each $P \in M$ and each $f \in \mathscr{F}$

$$(4.1) \qquad Qf = \lim Q_n f \quad \text{a.s.} \quad P.$$

A sequence Q_n is called an *asymptotically sufficient* statistic for M if there exists a sufficient statistic Q such that $Q_n \to Q$ a.s. M. If Q is H-sufficient, we say that Q_n is *asymptotically H-sufficient.*

To prove that a sequence Q_n is asymptotically sufficient, we use a concept of a support system.

4.2. A countable family W of bounded measurable functions in a measurable space $(\Omega, \mathscr{F})$ is called a *support system* if the following two conditions are satisfied:

4.2.A. If μ_n is a sequence of probability measures on $\mathscr{F}$ and if $\lim \int f\,d\mu_n = l(f)$ exists for each $f \in W$, then there is a probability measure μ such that $l(f) = \int f\,d\mu$ for all $f \in W$.

4.2.B. If a class H of real-valued functions contains W and is closed under addition, multiplication by constants, and bounded convergence, then H contains all bounded measurable functions. (We say that f_n converges boundedly to f if f_n converges pointwise to f and all the functions f_n are uniformly bounded.)

A measurable space $(\Omega, \mathscr{F})$ will be called a *B-space* if there exists a support system in $(\Omega, \mathscr{F})$. The unit interval $I = [0, 1]$ with the Borel measurable structure is an example of a B-space: a support system is formed by functions $1, x, x^2, \cdots, x^n, \cdots$.

A measurable space $(\Omega, \mathscr{F})$ is called a *Borel space* if it is isomorphic to a Borel subset of a complete separable metric space. It is well known (see, e.g., [7] or [19]) that all uncountable Borel spaces are isomorphic. By this fact it is easy to prove that all Borel spaces are B-spaces.

It follows from 4.2.B that a support system generates σ-algebra $\mathscr{F}$. Therefore, for any B-space $(\Omega, \mathscr{F})$, the σ-algebra $\mathscr{F}$ is countably generated and $M(\mathscr{F})$ is separable.

4.3.

LEMMA 4.1. *Let $(\Omega, \mathscr{F})$ be a B-space and let*

$$(4.2) \qquad P\{f \mid \mathscr{F}^0\} = \lim_{n\to\infty} Q_n f \quad \text{a.s.} \quad P$$

for every $P \in M$ and all $f \in \mathscr{F}$. Then Q_n is asymptotically sufficient for M and $\mathscr{F}^0$ is a sufficient σ-algebra for M.

PROOF. Put $\omega \in \Omega'$ if $\lim Q_n^\omega(f) = l^\omega(f)$ exists for all elements f of a support system W. If $\omega \in \Omega'$, then, by 4.1.A, there exists a probability measure Q^ω such that $Q^\omega(f) = l^\omega(f)$ for all $f \in W$. It follows from (4.2) that $P(\Omega') = 1$ and that

$$(4.3) \qquad P\{f \mid \mathscr{F}^0\} = Q^\omega(f) = Qf(\omega) \quad \text{a.s.} \quad P$$

for all $f \in W$, $P \in M$. By 4.1.B, (4.3) holds for all $f \in \mathscr{F}$. Therefore Q is sufficient for M. It follows from (4.2) and (4.3) that $Q_n \to Q$ a.s. M.

4.4. It follows from Theorem 3.1 that, if Q_n is an asymptotically H-sufficient statistic for M, then

$$Pf = \lim Q_n f \quad \text{a.s.} \quad P \tag{4.4}$$

for $P \in M_e$ and $f \in \mathscr{F}$. Formula (4.4) is valid also for all unbounded functions f for which (4.2) is true. In most applications, (4.2) and (4.4) hold for all P-integrable f.

Let us fix an arbitrary countable family W of bounded $\mathscr{F}$-measurable functions and define a convergence of measures by the condition that $P_n \to P$ if $P_n(f) \to P(f)$ for all $f \in W$. The formula (4.4) implies that each P in M_e is the limit of Q_n^ω for some ω.

4.5.

LEMMA 4.2. *Let M be a class of probability measures on a B-space $(\Omega, \mathscr{F})$. If $\mathscr{F}_1$, $\mathscr{F}_2$ are sufficient σ-algebras for M, then $\mathscr{F}^0 = \bar{\mathscr{F}}_1 \cap \bar{\mathscr{F}}_2$ is also sufficient for M. (Here $\bar{\mathscr{F}}_i$ is the minimal σ-algebra which contains $\mathscr{F}_i$ and all M-null sets.) If V_i is a sufficient statistic corresponding to $\mathscr{F}_i$ $(i = 1, 2)$, then formulas*

$$Q_1 = V_1, \qquad Q_{2k} = V_2 Q_{2k-1}, \qquad Q_{2k+1} = V_1 Q_{2k} \qquad \text{for} \quad k = 1, 2, \cdots \tag{4.5}$$

define an asymptotically sufficient statistic corresponding to $\mathscr{F}^0$.

PROOF. According to Lemma 4.1 it is sufficient to check formula (4.2). This formula follows from one result of Burkholder ([1], Theorem 4).

COROLLARY. *If $\mathscr{F}_i$ $i = 1, 2, \cdots$ are sufficient for M, then $\mathscr{F}^0 = \bigcap \bar{\mathscr{F}}_n$ is also sufficient for M.*

PROOF. By Lemma 4.2 all σ-algebras $\mathscr{A}_n = \bar{\mathscr{F}}_1 \cap \cdots \cap \bar{\mathscr{F}}_n$, $n = 1, 2, \cdots$ are sufficient for M. Let Q_n be corresponding sufficient statistics. Then

$$\lim Q_n f = \lim P\{f \mid \mathscr{A}_n\} = P\{f \mid \mathscr{F}^0\} \quad \text{a.s.} \quad P$$

for all $P \in M$ and all $f \in \mathscr{F}$.

5. Gibbs states.

5.1. Let L be a directed set, i.e., a partially ordered set with the property that for each two elements Λ_1, Λ_2 of L, there exists $\Lambda \in L$ such that $\Lambda > \Lambda_1$ and $\Lambda > \Lambda_2$. We consider two directed families indexed by L: a family of σ-algebras $\mathscr{F}_\Lambda \subset \mathscr{F}$ and a family of Markov operators Π_Λ in $(\Omega, \mathscr{F})$.

Following H. Föllmer, we say that $(\mathscr{F}_\Lambda, \Pi_\Lambda)$ is a *specification* in $(\Omega, \mathscr{F})$ if:

5.1.A. $\mathscr{F}_{\tilde{\Lambda}} \subset \mathscr{F}_\Lambda$ for $\tilde{\Lambda} > \Lambda$.

5.1.B. $\Pi_{\tilde{\Lambda}} \Pi_\Lambda = \Pi_{\tilde{\Lambda}}$ if $\tilde{\Lambda} > \Lambda$.

5.1.C. $\Pi_\Lambda f \in \mathscr{F}_\Lambda$ for $f \in \mathscr{F}$.

5.1.D. $\Pi_\Lambda f = f$ for $f \in \mathscr{F}_\Lambda$.

Concrete examples of specifications will be discussed in Sections 8 and 9.

A probability measure P on $(\Omega, \mathscr{F})$ is called a *Gibbs state specified by*

$\Pi = (\mathscr{F}_\Lambda, \Pi_\Lambda)$ if

$$P\{f \mid \mathscr{F}\} = \Pi_\Lambda f \quad \text{a.s.} \quad P \tag{5.1}$$

for each $f \in \mathscr{F}$ and $\Lambda \in L$.

We assume that the directed set Λ contains a *cofinal sequence* Λ_n, i.e., a sequence with the property that for every $\Lambda \in L$ there exists $\Lambda_n > \Lambda$.

Evidently, each Gibbs state P is a Π_Λ-invariant measure for all $\Lambda \in L$. On the other hand, if a probability measure P is invariant relative to the family $\{\Pi_\Lambda\}$, then, by 1.5.D, 1.5.C, and Lemma 3.1,

$$P\{f \mid \mathscr{F}_\Lambda\} = P\{\Pi_\Lambda f \mid \mathscr{F}_\Lambda\} = \Pi_\Lambda f \quad \text{a.s.} \quad P$$

and P is a Gibbs state. Now let P be invariant with respect to operators Π_{Λ_n} corresponding to a cofinal sequence Λ_n. For each $\Lambda \in L$ there exists a $\Lambda_n > \Lambda$ and, by 5.1.B, $\Pi_{\Lambda_n}\Pi_\Lambda = \Pi_{\Lambda_n}$. Therefore

$$P = P\Pi_{\Lambda_n} = P\Pi_{\Lambda_n}\Pi_\Lambda = P\Pi_\Lambda .$$

We see that the class $G(\Pi)$ of all Gibbs states specified by Π coincides with the class of all probability measures which are invariant with respect to a countable family Π_{Λ_n}.

5.2. We define the *tail σ-algebra* $\mathscr{F}^0$ as the intersection of all $\mathscr{F}_\Lambda$.

THEOREM 5.1. *Let $\Pi = (\mathscr{F}_\Lambda, \Pi_\Lambda)$ be a specification in a B-space $(\Omega, \mathscr{F})$. Then the tail σ-algebra $\mathscr{F}^0$ is H-sufficient for the class $G(\Pi)$ and, to each cofinal sequence Λ_n, there corresponds an asymptotically H-sufficient statistic Π_{Λ_n}.*

PROOF. It is clear that $\mathscr{F}_{\Lambda_n} \downarrow \mathscr{F}^0$. Therefore

$$\lim P\{f \mid \mathscr{F}_{\Lambda_n}\} = P\{f \mid \mathscr{F}^0\} \quad \text{a.s.} \quad P \tag{5.2}$$

for each probability measure P and each P-integrable f. If $P \in G(\Pi)$ then (5.2) implies (5.1). By Lemma 4.1, Π_{Λ_n} is an asymptotically sufficient statistic and $\mathscr{F}^0$ is a sufficient σ-algebra for $G(\Pi)$. Since

$$P\{\Pi_{\Lambda_n} f \mid \mathscr{F}^0\} = P\{P\{f \mid \mathscr{F}_{\Lambda_n}\} \mid \mathscr{F}^0\} = P\{f \mid \mathscr{F}^0\} \quad \text{a.s.} \quad P$$

for $P \in G(\Pi)$, $f \in \mathscr{F}$, the σ-algebra $\mathscr{F}^0$ is H-sufficient for $G(\Pi)$ by Theorem 3.4.

6. Shifts.

6.1. To each measurable transformation T of a space $(\Omega, \mathscr{F})$, there corresponds a Markov operator which transforms functions according to the formula

$$Tf(\omega) = f(T\omega)$$

and measures according to the formula

$$(PT)(A) = \int P(d\omega) 1_A(T\omega) = P(T^{-1}A) .$$

Markov operators of this kind will be called *shifts*.

LEMMA 6.1. *If T is a shift of a B-space $(\Omega, \mathscr{F})$, then*

$$Q_n = n^{-1} \sum_{k=0}^{n-1} T^k \tag{6.1}$$

is an asymptotically H-sufficient statistic for the class M of all T-invariant measures and the σ-algebra $\mathscr{F}^0$ *of all T-invariant sets is the corresponding H-sufficient σ-algebra.*

PROOF. By Birkhoff's ergodic theorem (see, e.g., [18], V-6), the relation (4.2) is satisfied, and Q_n is asymptotically sufficient for M by Lemma 4.1. H-sufficiency follows from Theorem 3.4 and Lemma 3.1.

COROLLARY. *Suppose that a shift T of a B-space* $(\Omega, \mathscr{F})$ *transforms into itself a class* $M \subset M(\mathscr{F})$ *and a σ-algebra* $\mathscr{F}^0 \subset \mathscr{F}$. *If* $\mathscr{F}^0$ *is H-sufficient for M, then the collection* $\mathscr{F}_T{}^0$ *of all T-invariant sets of* $\mathscr{F}^0$ *is H-sufficient for the class* M_T *of all T-invariant measures* $P \in M$.

This follows from Lemma 6.1 and 3.5 because the limit Q of operators (6.1) commutates with the conditioning with respect to $\mathscr{F}^0$.

THEOREM 6.1. *Let G be a finite or countable group of shifts of a B-space* $(\Omega, \mathscr{F})$. *The σ-algebra* $\mathscr{F}^0$ *of all G-invariant sets is H-sufficient for the class M of all G-invariant measures.*

PROOF. Denote by $\mathscr{A}_T$ the minimal σ-algebra containing all M-null sets and all T-invariant sets $A \in \mathscr{F}$. By Lemma 6.1, $\mathscr{A}_T$ is sufficient for M. By Lemma 4.2, an intersection $\mathscr{A}$ of $\mathscr{A}_T$ over all $T \in G$ is sufficient for M. Obviously, $\mathscr{F}^0 \subset \mathscr{A}$. On the other hand, if $A \in \mathscr{A}$, then $T1_A = 1_A$ a.s. M for each $T \in G$. The union B of $T^{-1}(A)$ over all $T \in G$ is G-invariant and $1_A = 1_B$ a.s. M. Hence $\mathscr{A}$ and $\mathscr{F}^0$ are M-equivalent and $\mathscr{F}^0$ is sufficient for M. H-sufficiency of $\mathscr{F}^0$ follows from Theorem 3.4 and Lemma 3.1.

REMARK. Theorem 6.1 holds for important classes of uncountable groups G. Suppose that there exists a countable subgroup G^1 of group G with the property that $\mathscr{F}^0$ is M-equivalent to the σ-algebra $\mathscr{F}^1$ of all G^1-invariant sets. As we know, $\mathscr{F}^1$ is sufficient for the class of all G^1-invariant measures. Hence $\mathscr{F}^1$ is sufficient for M, and $\mathscr{F}^0$ is sufficient for M too. By the remark at the end of 3.4, $\mathscr{F}^0$ is H-sufficient for M if G satisfies conditions (i), (ii).

Now let G be a locally compact group. Then condition (ii) is satisfied for Borel σ-algebra $\mathscr{B}_G$ and Haar measure λ. Condition (i) implies that, for each $P \in M$ and every square integrable f, $T \to Tf$ is a continuous mapping of G into $L^2(\Omega, P)$ (see, e.g., [17], Section 29). Using this fact, it is easy to prove that, if G has a countable everywhere dense subgroup G^1, then $\mathscr{F}^1$ is M-equivalent to $\mathscr{F}^0$ and $\mathscr{F}^0$ is H-sufficient for M.

The rôle of σ-algebra $\mathscr{F}^0$ for decomposition of invariant measures into extreme elements was discovered independently by Farrell [8] and Varadarajan [21]. The fact that $\mathscr{F}^0$ is a sufficient σ-algebra for M is proved in [8] also for a certain class of abelian semigroups.

6.2. We consider now a slightly wider class of operators than shifts.

THEOREM 6.2. *Let T be an invertible transformation of a B-space* $(\Omega, \mathscr{F})$, *let*

T and T^{-1} be measurable, and let $Y(\omega)$ be a strictly positive $\mathscr{F}$-measurable function. Let $Uf(\omega) = Y(\omega)f(T\omega)$. Then

$$V_n f = \frac{\sum_{-n}^{n} U^k f}{\sum_{-n}^{n} U^k 1}$$

is an asymptotically H-sufficient statistic for the class M of all U-invariant probability measures. The corresponding H-sufficient σ-algebra $\mathscr{F}^0$ consists of all T-invariant sets.

PROOF. We prove that

$$P\{f \mid \mathscr{F}^0\} = \lim_{n\to\infty} V_n f \quad \text{a.s.} \quad P \tag{6.2}$$

for every $P \in M$ and every P-integrable f.

Put $\gamma^{-1} = \sum U^k 1$, $\varphi = U^k f$, summing over all integers k. Let $\Omega_0 = \{\omega : \gamma = 0\}$, $\Omega_1 = \{\omega : \gamma > 0\}$. By the Chacon-Ornstein theorem (see, e.g., [18], V, 6.4) (6.2) holds on Ω_0 and, in order to prove that it holds on Ω_1, we need only check that

$$P\{f \mid \mathscr{F}^0\} = \varphi\gamma \quad \text{a.s.} \quad P \quad \text{on} \quad \Omega_1 . \tag{6.3}$$

The obvious relations $U\varphi = \varphi$, $U\gamma^{-1} = \gamma^{-1}$ imply that $T\varphi = \varphi Y^{-1}$, $T\gamma = \gamma Y$, and $T(\varphi\gamma) = \varphi\gamma$. Hence $\varphi\gamma$ is $\mathscr{F}^0$-measurable. On the other hand, $(Uf)g = U(fT^{-1}g)$ and therefore

$$P(gUf) = P(fT^{-1}g) \tag{6.4}$$

for all $P \in M$ and all positive $\mathscr{F}$-measurable f, g. It follows from this that

$$P(g\gamma U^k f) = P(gfT^{-k}\gamma)$$

for $g \in \mathscr{F}^0$ and $k = 0, \pm 1, \cdots$. Hence

$$P(g\gamma\varphi) = P(gf\alpha) , \tag{6.5}$$

where $\alpha = \sum T^{-k}\gamma$. Since α is $\mathscr{F}^0$-measurable, (6.5) implies that

$$\alpha P\{f \mid \mathscr{F}^0\} = \gamma\varphi . \tag{6.6}$$

Now α does not depend on f. Taking $f = 1$, we see that $\alpha = 1$, and (6.6) goes into (6.3). By Lemma 4.1, $\mathscr{F}^0$ is sufficient for M. Formula (6.4) implies (3.19), and $\mathscr{F}^0$ is H-sufficient for M by Theorem 3.4.

REMARK. Suppose that T_t is a one-parameter group of shifts and $U_t f(\omega) = Y_t f(T_t\omega)$ where $Y_{s+t} = Y_s T_s Y_t$. Then Theorem 6.2 holds with

$$V_t = \frac{\int_{-t}^{t} U_s\, ds}{\int_{-t}^{+t} U_s 1\, ds}$$

instead of V_n. This result was first proved by Yu. I. Kifer and S. A. Pirogov in an appendix to [5].

6.3.

THEOREM 6.3. *Let a class M of positive functions be a B-space and a simplex and let $k\varphi \bar{\in} M$ if $\varphi \in M$ and $k \neq 1$. Suppose that T is an automorphism of a* *cone* $M^* = \{k\varphi : \varphi \in M, k > 0\}$ *(which means that T and* T^{-1} *preserve generalized barycentres). Then the set* M_T *of all points* $\varphi \in M$ *such that* $T\varphi = \varphi$ *is also a simplex. This statement is true also for a one-parameter group of transformations* T_t.

PROOF. For each $\varphi \in M^*$ there exists one and only one positive number $k(\varphi)$ such that $\varphi/k(\varphi) \in M$. Put $Y(\varphi) = k(T\varphi)$ and $\bar{T}\varphi = T\varphi/Y(\varphi)$. Obviously $\bar{T}$ is an invertible transformation of the measurable space (M_e, B_{M_e}), and $\bar{T}$ and $\bar{T}^{-1}$ are measurable. Each $\varphi \in M$ can be uniquely represented in the form

$$\varphi = \int_{M_e} \tilde{\varphi}\mu(d\tilde{\varphi}) . \tag{6.7}$$

Hence

$$T\varphi = \int_{M_e} T\tilde{\varphi}\mu(d\tilde{\varphi}) = \int_{M_e} Y(\tilde{\varphi})\bar{T}\tilde{\varphi}\mu(d\tilde{\varphi}) = \int_{M_e} \psi\mu_1(d\psi)$$

where

$$\mu_1(d\psi) = Y(\bar{T}^{-1}\psi)\mu(\bar{T}^{-1}\,d\psi) .$$

It is clear that $T\varphi = \varphi$ if and only if $\mu_1 = \mu$ which is equivalent to the relation $\mu U = \mu$ where $UF(\psi) = Y(\psi)F(T\psi)$. By Theorem 6.2, the class of all U-invariant probability measures is a simplex, and formula (6.12) establishes an isomorphism of this class and M_T.

7. Symmetric measures.

7.1. In the rest of the paper we investigate various classes of measures on product spaces. We start with the necessary notations.

Let there be given an arbitrary set S and a set E_s associated with each s of S. We call a *configuration* and denote by x_S a collection of $x_s \in E_s$, $s \in S$. The *product space* E_S is the set of all configurations x_S. A space E_Λ of configurations x_Λ over Λ corresponds to each subset Λ of the set S.

Now let a σ-algebra $\mathscr{B}_s$ in E_s be fixed for each $s \in S$. We denote by $\mathscr{B}_\Lambda$ the minimal σ-algebra in E_Λ which contains sets $\{x_\Lambda : x_s \in \Gamma\}$ for all $s \in \Lambda$, $\Gamma \in \mathscr{B}_s$. To each probability measure P on $(E_S, \mathscr{B}_S)$ and each $\Lambda \subset S$, there corresponds a probability measure P_Λ on $(E_\Lambda, \mathscr{B}_\Lambda)$ defined by the formula

$$P_\Lambda(A) = P\{x_S : x_\Lambda \in A\} , \quad A \in \mathscr{B}_\Lambda . \tag{7.1}$$

A collection of measures P_Λ for all finite $\Lambda \subset S$ is called a *system of finite-dimensional distributions.* If $(E_s, \mathscr{B}_s)$ are Borel spaces, then (7.1) establishes a one-to-one correspondence between all probability measures P on $(E_S, \mathscr{B}_S)$ and all consistent systems of finite-dimensional distributions (Kolmogorov's theorem). In particular, to each family of probability measures p_s, $s \in S$, there corresponds a product measure P for which all finite-dimensional distributions P_Λ are the products of p_s, $s \in \Lambda$.

A system of random variables on the probability space $(E_S, \mathscr{B}_S, P)$ is given by the formula

$$X_s(\omega) = x_s \quad \text{for} \quad \omega = x_S , \quad s \in S .$$

These random variables are independent if and only if P is a product measure.

Sets $\{x_S : x_\Lambda \in A\}$, Λ finite, $A \in \mathscr{B}_\Lambda$, are called *cylinders*. Two measures on $\mathscr{B}_S$ are identical if they coincide on all cylinders.

7.2. In this section we assume that $(E_s, \mathscr{B}_s) = (E, \mathscr{B})$ does not depend on s and we write E^S for E_S and $\mathscr{B}^S$ for $\mathscr{B}_S$. Any transformation g of S induces a transformation $x_S' = T_g x_S$ of the space E^S given by the formula $x_s' = x_{gs}$. Put $g \in G$ if g is invertible and $gs \neq s$ only for a finite number of s. Measures, measurable sets, and functions invariant with respect to the family of operators T_g, $g \in G$, will be called *symmetric*.

Let Γ be a finite subset of S. Denote by G^Γ the totality of all $g \in G$ such that $gs = s$ outside G. Denote by $\mathscr{F}^\Gamma$ the class of all elements of $\mathscr{B}^S$-invariant relative to T_g, $g \in G^\Gamma$. Let V_Γ be an arithmetic mean of operators T_g, $g \in G^\Gamma$. It is easy to see that

$$P\{f \mid \mathscr{F}^\Gamma\} = V_\Gamma f \quad \text{a.s.} \quad P \tag{7.2}$$

for each symmetric measure P and each P-integrable f.

THEOREM 7.1. *Let S be a countable set, $(E, \mathscr{B})$ a Borel space and M the class of all symmetric measures on $(E^S, \mathscr{B}^S)$. Then*:

(a) *M is a simplex*;

(b) *a measure P is an extreme point of M if and only if P is a product of identical probability measures $p_s = p$, $s \in S$ (in other words, if X_s, $s \in S$ are identically distributed independent random variables)*;

(c) *the class $\mathscr{F}^0$ of all symmetric sets is an H-sufficient σ-algebra for M and V_{Γ_n} is the corresponding asymptotically H-sufficient statistic if $\Gamma_n \uparrow S$.*

PROOF. The fact that $\mathscr{F}^0$ is H-sufficient for M follows immediately from Theorem 6.1. If $\Gamma_n \uparrow S$, then $\mathscr{F}^{\Gamma_n} \downarrow \mathscr{F}^0$ and (7.2) implies that

$$P\{f \mid \mathscr{F}^0\} = \lim V_{\Gamma_n} f \quad \text{a.s.} \quad P$$

for all $P \in M$ and all $f \in \mathscr{F}$. By Lemma 4.1, V_{Γ_n} is an asymptotically sufficient statistic for M. The statement (c) is proved. By Theorem 3.1, (c) implies (a).

It remains to prove (b). Let $S = \{0, 1, \cdots, n, \cdots\}$ and $\Gamma_n = \{0, 1, \cdots, n-1\}$. By virtue of (c) and (4.4)

$$Pf = \lim V_{\Gamma_n} f \quad \text{a.s.} \quad P \tag{7.3}$$

for $P \in M_e$ and $f \in \mathscr{B}^S$. In order to prove that P is a product measure, it suffices to check that, for all m and all $A \in \mathscr{B}^{\Gamma_n}$, $B \in \mathscr{B}$,

$$P\{x_{\Gamma_m} \in A,\, x_m \in B\} = P\{x_{\Gamma_m} \in A\} P\{x_m \in B\}\,. \tag{7.4}$$

It follows from (7.3) that

$$P\{1_A(x_{\Gamma_m}) P 1_B(x_m)\} = \lim_{n \to \infty} P\{1_A(x_{\Gamma_m}) V_{\Gamma_n} 1_B(x_m)\}\,. \tag{7.5}$$

Evidently $V_{\Gamma_n} 1_B(x_m) = n^{-1} \sum_{k=1}^n 1_B(x_k)$ for $n \geqq m$. Since

$$P\{x_{\Gamma_m} \in A,\, x_k \in B\} = P\{x_{\Gamma_m} \in A,\, x_m \in B\}$$

for all $k \geqq m$, (7.5) implies (7.4).

Now we prove that each product measure $\bar{P} \in M$ belongs to M_e. Since $\bar{P}$ is a barycentre of a probability measure μ concentrated on M_e, we have

$$\bar{P}[\varphi(x_0)\varphi(x_1)] = \int_{M_e} P[\varphi(x_0)\varphi(x_1)]\mu(dP) \tag{7.6}$$

for every $\varphi \in \mathscr{B}$. Here $\bar{P}$ and P are symmetric product measures and therefore (7.6) is equivalent to

$$[\bar{P}\varphi(x_0)]^2 = \int_{M_e} P\varphi(x_0)^2\mu(dP) ,$$

which implies that

$$\int_{M_e} [P\varphi(x_0) - \bar{P}\varphi(x_0)]^2\mu(dP) = 0 . \tag{7.7}$$

It follows from (7.7) that $\mu\{P: P \in M_e, P = \bar{P}\} = 1$. Thus $\bar{P} \in M_e$.

7.3. The statements (a) and (b) of Theorem 7.1 are true for uncountable S too. Indeed, if Λ is a countable subset of S, then the measure P_Λ, introduced by (7.1), characterizes a symmetric measure P uniquely because it defines all finite-dimensional distributions of P. Hence the mapping $P \to P_\Lambda$ is a one-to-one mapping of M onto the set of all symmetric measures on $(E^\Lambda, \mathscr{B}^\Lambda)$. This mapping preserves the convex structure, and P is a product measure if and only if P_Λ is a product measure also.

The statement (c) has to be modified as follows. Let Λ be an arbitrary countable subset of S. Denote by $\mathscr{F}^0$ the collection of all sets of the form $A \times E^{S\setminus\Lambda}$ where A is a symmetric subset of E^Λ. Then $\mathscr{F}^0$ is an H-sufficient σ-algebra for M.

8. Stochastic fields.

8.1. Let $(E_S, \mathscr{B}_S)$ be a product of spaces $(E_s, \mathscr{B}_s)$, $s \in S$, and let L be a collection of subsets of S ordered by inclusion. Denote by $\mathscr{F}_\Lambda$ a σ-algebra in E_S generated by random variables X_s, $s \in S\setminus\Lambda$. Assume that, for each $\Lambda \in L$, a measure $p_\Lambda(\cdot \mid x_{S\setminus\Lambda})$ is given on $(E_\Lambda, \mathscr{B}_\Lambda)$ which depends on $x_{S\setminus\Lambda}$, and put

$$\Pi_\Lambda f(x_S) = \int_{E_{S\setminus\Lambda}} f(x_{S\setminus\Lambda}\, y_\Lambda) p_\Lambda(dy_\Lambda \mid x_{S\setminus\Lambda}) . \tag{8.1}$$

(We denote by $x_{S\setminus\Lambda}\, y_\Lambda$ a configuration which coincides with y_Λ over Λ and with $x_{S\setminus\Lambda}$ over $S\setminus\Lambda$.) If $\Pi = (\mathscr{F}_\Lambda, \Pi_\Lambda)$ is a specification (i.e., if 5.1.A—5.1.D are satisfied), we say that p is a *specifying function*.

We say that (X_s, P) is a *stochastic field specified by* p if

$$P\{X_\Lambda \in A \mid X_{S\setminus\Lambda}\} = p_\Lambda(A \mid X_{S\setminus\Lambda}) \quad \text{a.s.} \quad P \tag{8.2}$$

for each $\Lambda \in L$ and each $A \in \mathscr{B}_\Lambda$. Obviously (8.2) is equivalent to (5.1). Hence Theorem 5.1 can be applied to the set of all stochastic fields specified by p.

8.2. Let S be a countable set and let L be the collection of all finite subsets of S. To each real-valued function $U(\Gamma, x_\Gamma)$, $\Gamma \in L$, $x_\Gamma \in E_\Gamma$, there corresponds a specifying function

$$p_\Lambda(C \mid x_{S\setminus\Lambda}) = Z^{-1} \int_C [\exp \textstyle\sum U(\Gamma, x_\Gamma)] \prod_{s\in\Lambda} \lambda_s(dx_s) , \tag{8.3}$$

where Γ runs over all finite subsets of S such that $\Gamma \cap \Lambda \neq \emptyset$, λ_s is a measure on $(E_s, \mathscr{B}_s)$ and Z is independent of x_Λ and can be calculated from the condition $p_\Lambda(E_\Lambda \mid x_{S\setminus\Lambda}) = 1$. (The only restriction on U is convergence of series in (8.3).)

Now suppose that S is a graph. A specifying function p is called Markov if $p_\Lambda(\cdot \mid x_{S\setminus\Lambda})$ depends only on $x_{\partial\Lambda}$ where $\partial\Lambda$ is a collection of all points of $S\setminus\Lambda$ which have neighbors in Λ. A function (8.3) is Markov if and only if the inequality $U(\Gamma, x_\Gamma) \neq 0$ implies that each two points s_1, s_2 of Γ are neighbors.

Proofs of all statements of subsection 8.2 can be found, for example, in [20].

8.3. Now suppose $(E_s, \mathscr{B}_s) = (E, \mathscr{B})$ does not depend on s. Let L be the collection of all finite subsets of a countable set S. Consider the family $\{\mathscr{F}^\Gamma, \Gamma \in L\}$ of σ-algebras in E^S which has been defined in 7.2. Obviously $\mathscr{F}^\Gamma \supset \mathscr{F}^{\tilde{\Gamma}}$ if $\Gamma \subset \tilde{\Gamma}$. Suppose that, for each $\Gamma \in L$, a measure $p_\Gamma(\cdot \mid x_S)$ on $(E^\Gamma, \mathscr{B}^\Gamma)$ is given depending on x_S and such that operators

$$\Pi_\Gamma f(x_S) = \int_{E^S} f(x_{S\setminus\Gamma}) p_\Gamma(dy_\Lambda \mid x_S)$$

satisfy conditions 5.1.B—5.1.D. Theorem 5.1 can be applied to the class of all probability measures P satisfying the condition

$$P\{X_\Gamma \in A \mid \mathscr{F}^\Gamma\} = p_\Gamma\{A \mid X_S\} \quad \text{as.} \quad P \qquad \text{for all} \quad \Gamma \in L \quad \text{and all} \quad A \in \mathscr{B}^\Gamma .$$

The tail σ-algebra $\mathscr{F}^0 = \bigcap \mathscr{F}^\Gamma$ coincides with a collection of all symmetric subsets of E^S.

9. Markov processes with a given transition function.

9.1. A stochastic field (X_s, P), $s \in S$, is called a *stochastic process* if S is a subset of a real line. The case when S is an interval is the most important.

We denote by $\mathscr{F}_{\leqq s}$ the σ-algebra in E_S generated by X_t, $t \leqq s$, $t \in S$. The notations $\mathscr{F}_{<s}$, $\mathscr{F}_{>s}$, $\mathscr{F}_{\geqq s}$ have an analogous meaning.

A real-valued function $p(s, x; t, \Gamma)$, $s < t \in S$, $x \in E_s$, $\Gamma \in \mathscr{B}_t$, is called a *Markov transition function* if $p(s, x; t, \cdot)$ is a probability measure, $p(s, \cdot; t, \Gamma)$ is a $\mathscr{B}_s$-measurable function, and

$$(9.1) \qquad p(s, x; u, \Gamma) = \int_{E_t} p(s, x; t, dy) p(t, y; u, \Gamma)$$

for all $s < t < u \in S$, $x \in E_s$, $\Gamma \in \mathscr{B}$.

Starting from a transition function p, we define a specification $\Pi(p)$ in the following way. We consider a family of finite-dimensional distributions

$$(9.2) \qquad \begin{aligned} &p(s_1, dx_1, \cdots, s_n, dx_n) \\ &\quad = p(s, x; s_1, dx_1)p(s_1, x_1; s_2, dx_2) \cdots p(s_{n-1}, x_{n-1}; s_n, dx_n), \\ &\qquad\qquad s_1 < s_2 < \cdots < s_n \in \Lambda_s = S \cap (s, +\infty) \end{aligned}$$

and denote by $P_{s,x}$ the corresponding probability measure on $\mathscr{F}_{>s}$. We define L as the totality of all sets Λ_s, $s \in S$, and put

$$(9.3) \quad \mathscr{F}_{\Lambda_s} = \mathscr{F}_{\leqq s}, \quad \Pi_{\Lambda_s} f(x_S) = \int f(x_{S\setminus\Lambda_s} y_{\Lambda_s}) P_{s,x_s}(dy_{\Lambda_s}) = P_{s,x_s} f(x_{S\setminus\Lambda_s} X_{\Lambda_s}) .$$

Then $\Pi(p) = (F_\Lambda, \Pi_\Lambda)$ is a specification. (Formula (9.3) is a particular case of (8.1) with $p_{\Lambda_s}(A \mid x_{S\setminus\Lambda_s}) = P_{s,x_s}\{x_{\Lambda_s} \in A\}$.) We will use an abbreviation Π_s for the operator Π_{Λ_s}.

We say that (X_s, P) is a *Markov process with a transition function p* and we write $P \in M(p)$ if

$$P\{A \mid \mathscr{F}_{\leqq s}\} = P_{s,X_s}(A) \quad \text{a.s.} \quad P \tag{9.4}$$

for each $s \in S$, $A \in \mathscr{F}_{>s}$. Obviously $M(p)$ coincides with the set of all Gibbs states specified by $\Pi(p)$.

9.2. THEOREM 9.1. *Let $r = \inf S$. If $r \in S$, then Π_r is an H-sufficient statistic for $M(p)$ and the corresponding H-sufficient σ-algebra is generated by X_r. If $r \bar{\in} S$, then an intersection $\mathscr{F}^0$ of all σ-algebras $\mathscr{F}_{<s}$, $s \in S$ is an H-sufficient σ-algebra for $M(p)$ and, to each sequence $s_n \downarrow r$, $s_n \in S$ these corresponds an asymptotically H-sufficient statistic Π_{s_n}.*

PROOF. In the case $r \in S$, we need only check that

$$P\{f \mid X_r\} = \Pi_r f \quad \text{a.s.} \quad P \tag{9.5}$$

if $P \in M(p)$ and $f \in \mathscr{B}_s$. It is sufficient to prove this only for functions of the form $f(x_S) = \varphi(x_r)\psi(x_{\Lambda_r})$. But for such f the left side of (9.5) is equal to

$$\varphi(X_r)P\{\psi(X_{\Lambda_r}) \mid \mathscr{F}_{\leqq r}\} = \varphi(X_r)P_{r,X_r}\psi(X_{\Lambda_r})$$

and, by the definition of Π_r, the right side of (9.5) is the same.

Suppose now that $r \bar{\in} S$. Then to each $s_n \downarrow r$, $s_n \in S$ there corresponds a cofinal sequence Λ_{s_n}, and, if (E_S, B_S) is a B-space, we can apply Theorem 5.1. This is the case if S is countable.

If S is not countable, we consider a countable subset $\Lambda = \{s_n\}$ of S where $s_n \downarrow r$ and we use the same trick as in 7.3 replacing each measure P by P_Λ.

Since $\mathscr{F}_{\leqq s_m} \downarrow \mathscr{F}^0$, we have

$$P\{f \mid \mathscr{F}^0\} = \lim_{m\to\infty} \Pi_{s_m} f \quad \text{a.s.} \quad P \tag{9.6}$$

for each $P \in M(p)$ and each P-integrable f, and our theorem will be proved if we construct a Markov operator Q with the properties

$$P\{f \mid \mathscr{F}^0\} = Qf \quad \text{a.s.} \quad P\,, \qquad Q^\omega \in M(p) \quad \text{a.s.} \quad P\,, \tag{9.7}$$

for each $P \in M(p)$. It follows from (9.4) that, for $P \in M(p)$,

$$P\{f \mid \mathscr{F}^0\} = P\{P\{f \mid \mathscr{F}_{\leqq s}\} \mid \mathscr{F}^0\} = P\{P_{s,X_s} f \mid \mathscr{F}^0\} \quad \text{a.s.} \quad P\,.$$

Therefore if

$$P\{\varphi(X_{s_n}) \mid \mathscr{F}^0\} = Q\varphi(X_{s_n}) \quad \text{a.s.} \quad P \in M(p) \tag{9.8}$$

for all $n = 1, 2, \cdots$ and all $\varphi \in \mathscr{B}_{s_n}$, then (9.7) is true for all n and all $f \in \mathscr{F}$, and hence, it is true for all $f \in \mathscr{F}$.

Denote by $\tilde{M}(p)$ the class of all Markov processes on $(E_\Lambda, \mathscr{B}_\Lambda)$ with the transition function p and by $\tilde{\mathscr{F}}^0$ the corresponding tail σ-algebra. It follows from (9.5) and an analogous formula for $\tilde{\mathscr{F}}^0$ that

$$(9.9) \qquad P\{\varphi(x_{s_n}) \mid \tilde{\mathscr{F}}^0\} = P\{\varphi(x_{s_n}) \mid \mathscr{F}^0\} \quad \text{a.s.} \quad P \in M(p) .$$

The mapping $P \to P_\Lambda$ is a one-to-one mapping of $M(p)$ onto $\tilde{M}(p)$. As we know, there exists a Markov operator $\tilde{Q}$ in (E_Λ, B_Λ) such that $\tilde{Q}^\omega \in \tilde{M}(p)$ a.s. $\tilde{M}(p)$ and

$$(9.10) \qquad P\{\varphi(x_{s_n}) \mid \tilde{\mathscr{F}}^0\} = \tilde{Q}\varphi(x_{s_n}) \quad \text{a.s.} \quad P$$

for each $P \in \tilde{M}(p)$ and each $\varphi \in \mathscr{B}_{s_n}$. Denote by Q^ω a measure of class $M(p)$ which corresponds to $\tilde{Q}^\omega$. It follows from (9.9) and (9.10) that Q satisfies (9.8).

10. Entrance and exit laws.

10.1. Let $p(s, x; t, \Gamma)$ be a Markov transition function. Put

$$P_t^s h^t(x) = \int_{E_t} p(s, x; t, dy) h^t(y) ,$$
$$(\nu_t P_t^s)(\Gamma) = \int_{E_t} \nu_s(dx) p(s, x; t, \Gamma) .$$

(Here h^t is a $\mathscr{B}_t$-measurable function with the values in the extended half-line $[0, +\infty]$; ν_t is a measure on $\mathscr{B}_t$.)

We say that ν is an *entrance law* if $\nu_s P_t^s = \nu_t$ for all $s < t \in S$ and we say that h is an *exit law* if $P_t^s h^t = h^s$ for all $s < t \in S$.

If ν is an entrance law and h is an exit law, then the value of $\nu_t(h_t)$ does not depend on t and we denote it by $\{\nu, h\}$. If $\{\nu, h\} = 1$, then the formula

$$(10.1) \qquad \begin{aligned} &p(t_1, dx_1, \cdots, t_n, dx_n) \\ &\quad = \nu_{t_1}(dx_1) p(t_1, x_1; t_2, dx_2) \cdots p(t_{n-1}, x_{n-1}; t_n, dx_n) h^{t_n}(x_n) , \\ &\qquad\qquad t_1 < \cdots < t_n \in S \end{aligned}$$

defines a family of consistent finite-dimensional distributions, and we denote by P_ν^h the corresponding probability measure on $(E_S, \mathscr{B}_S)$.

Let $\infty > h^s(x) > 0$ for all s, x.

Let R_p^h be the class of all entrance laws ν normed by the condition $\{\nu, h\} = 1$ with natural measurable and convex structures. Let $M(p^h)$ be the class of all Markov processes with the transition function

$$(10.2) \qquad p^h(s, x; t, dy) = h^s(x)^{-1} p(s, x; t, dy) h^t(y) .$$

It is easy to see that $\nu \to P_\nu^h$ is an isomorphism of convex measurable spaces R_ν^h and $M(p^h)$. According to Theorem 9.1 and 3.1, the space R_ν^h is a simplex. If ν is an extreme point of R_p^h, and if $P_\nu^h[\varphi^t(X_t)] < \infty$, then by 4.4, for every sequence $s_n \downarrow r_1$,

$$(10.3) \qquad P_\nu^h[\varphi^t(X_t)] = \lim P^h_{s_n, X_{s_n}} \varphi^t(X_t) \quad \text{a.s.} \quad P_\nu^h .$$

Formula (10.3) implies that for each ν_t-integrable f

$$(10.4) \qquad \nu_t(f) = \lim h^{s_n}(X_{s_n})^{-1} \int_{E_{s_n}} p(s_n, X_{s_n}; t, dy) f(y) \quad \text{a.s.} \quad P_\nu^h .$$

10.2. Now we investigate a class S_ν^p of all exit laws normed by the condition $\{\nu, h\} = 1$ under the following additional assumption:

10.2.A. If $\nu_t(\Gamma) = 0$, then $p(s, x; t, \Gamma) = 0$ for all $s \in t$, $x \in E_s$.

We proved in [6] (see Lemma 4.2) that a density $\rho(s, x; t, y) = p(s, x; t, dy)/\nu_t(dy)$ can be selected in such a way that

$$\int_{E_t} \rho(s, x; t, y)\nu_t(dy)\rho(t, y; u, z) = \rho(s, x; u, z) \tag{10.5}$$

for all $s < t < u$, $x \in E_s$, $z \in E_u$ and

$$\int \rho(s, x; t, y)\nu_s(dx) = 1 \qquad \text{for all} \quad s\,, \quad x\,. \tag{10.6}$$

The formula

$$\hat{p}(s, dx; t, y) = \nu_s(dx)\rho(s, x; t, y) \tag{10.7}$$

defines a backward transition function. Starting from $\hat{p}$, we define probability measures $P^{u,x}$ on $\mathscr{F}_{<u}$ exactly in the same way as measures $P_{s,x}$ were defined with the help of forward transition function p. We say that (X_s, P) is a Markov process with a backward transition function $\hat{p}$ if

$$P\{A \,|\, \mathscr{F}_{\geqq u}\} = P^{u,X_u}(A) \quad \text{a.s.} \quad P \qquad \text{for} \quad A \in \mathscr{F}_{<u}\,.$$

We consider a measurable structure in S_ν^p generated by functions $F(h) = \nu_s(\varphi h^s)$, $s \in S$, $\varphi \in \mathscr{B}_s$. It was proved in [4] (Lemma 4.2) that $h^s(x)$ is measurable with respect to the pair h, x, and hence the condition 2.2.A is satisfied. It is easy to check that the mapping $h \to P_\nu^h$ is an isomorphism of S_ν^p onto the class $M(\hat{p})$ of all Markov processes with the backward transition function $\hat{p}$ defined by (10.7). Now we use Theorem 3.1 and propositions dual to Theorem 9.1 and to formula (4.4), and we conclude that S^p is a simplex and that

$$\nu_t(h^t\varphi^t) = \lim P^{u_n, X_{u_n}}\varphi^t(x_t) = \lim \int \varphi^t(x)\hat{p}(t, dx; u_n, X_{u_n}) \quad \text{a.s.} \quad P_\nu^h \tag{10.8}$$

if h is an extreme point of S_ν^p, if $u_n \uparrow r_2$ and if $P_\nu^h|\varphi^t(X_t)| < \infty$.

It follows from (10.8) and (10.7) that

$$\int h^t(x)\varphi^t(x)\,\nu_t(dx) = \lim \int \rho(t, x; u_n, X_{u_n})\varphi^t(x)\nu_t(dx) \quad \text{a.s.} \quad P_\nu^h$$

if $h^t\varphi^t$ is ν_t-integrable. Applying the last formula to

$$\begin{aligned} \varphi^t(x) &= \rho(s, x; t, y) \qquad &&\text{for} \quad t > s\,, \\ &= 0 &&\text{for} \quad t \leqq s\,, \end{aligned}$$

we see that, if h is extreme and if $h^s(x) < \infty$, then

$$h^s(x) = \lim \rho(s, x; u_n, X_{u_n}) \quad \text{a.s.} \quad P_\nu^h\,.$$

REMARK. S. E. Kuznecov [14] has proved that the assumption 10.2.A is not only sufficient but also necessary for the class S_ν^p to be a simplex.

11. Excessive measures and excessive functions.

11.1. In this section, the results of Section 10 will be extended to wider classes of measures and functions associated with a transition function p. Let

S coincide with the set of all real numbers. An *excessive function* h and an *excessive measure* ν are defined, respectively, by conditions

$$P_t^s h^t \leqq h^s\,, \qquad P_t^s h^t \uparrow h^s \quad \text{as} \quad t \downarrow s$$

and

$$\nu_s P_t^s \leqq \nu_s\,, \qquad \nu_s P_t^s \uparrow \nu_t \quad \text{as} \quad s \uparrow t\,.$$

It is convenient to replace in the definition of a Markov transition function the condition $p(s, x; t, E_t) = 1$ by a weaker condition $p(s, x; t, E_t) \leqq 1$. An immediate gain is that an extended class is invariant with respect to transformation $p \to p^h$ defined by formula (10.2) for each strictly positive finite excessive function h.

Let ν be an excessive measure and h be an excessive function. We put $\{\nu, h\} = +\infty$ if $\nu_t(h^t) = +\infty$ for some t. If $\nu_t(h^t) < \infty$ for all t, we define $\{\nu, h\}$ as a supremum of sums

$$\nu_{t_1}(h^{t_1}) + \textstyle\sum_{k=2}^{m} [\nu_{t_k}(h^{t_k}) - \nu_{t_{k-1}}(P_{t_k}^{t_{k-1}} h^{t_k})]$$

over all finite subsets $t_1 < t_2 < \cdots < t_m$ of S. (This is consistent with the definition given in Section 10 for the case of an entrance law ν and an exit law h.)

The crucial point is the construction of a probability measure P_ν^h corresponding to a triple p, ν, h such that $\{\nu, h\} = 1$. As in Section 10, we start from formula (10.1). However P_ν^h will be defined not on $(E_S, \mathscr{B}_S)$ but on a different space (Ω, F). In order to construct this space, we add to E_s two extra points a_s and b_s and denote by $\bar{\mathscr{B}}_s$ a σ-algebra in $\bar{E}_s = E_s \cup a_s \cup b_s$ generated by $\mathscr{B}_s$ and the one-point sets $\{a_s\}$ and $\{b_s\}$. We define Ω as a subset of the product space $(\bar{E}_S, \bar{\mathscr{B}}_S) = \Pi_{s \in S} (E_s, \bar{\mathscr{B}}_s)$, namely, x_S of E_S belongs to Ω if there exist two real numbers $\alpha < \beta$ such that

$$x_s = a_s \quad \text{for} \quad s \leqq \alpha\,, \qquad x_s \in E_s \quad \text{for} \quad s \in (\alpha, \beta)\,, \qquad x_s = b_s \quad \text{for} \quad s \geqq \beta\,.$$

The random variables $\alpha(\omega)$ and $\beta(\omega)$ are called *the birth time* and *the death time*. To each $s \in S$ there corresponds a function X_s on Ω defined by the formula

$$X_s(\omega) = x_s \qquad \text{for} \quad \omega = x_S\,,$$

and we denote by $\mathscr{F}$ the σ-algebra in Ω generated by X_s, $s \in S$.

We proved in [5] that, if $\{\nu, h\} = 1$, then there exists one and only one probability measure P_ν^h on (Ω, F) such that, for every $t_1 < \cdots < t_n \in S$, $\Gamma_1 \in \mathscr{B}_{t_1}, \cdots$, $\Gamma_n \in \mathscr{B}_{t_n}$,

$$P_\nu^h\{\alpha < t_1, X_{t_1} \in \Gamma_1, \cdots, X_{t_n} \in \Gamma_n, \beta > t_n\} = p(t_1, \Gamma_1, \cdots, t_n, \Gamma_n)\,,$$

where the right side is defined by formula (10.1).

To each $s \in S$, $x \in E_s$ there corresponds an excessive measure

$$\begin{aligned} \nu_t^{s,x}(\Gamma) &= p(s, x, t, \Gamma) \qquad &\text{for} \quad t > s\,, \\ &= 0 &\text{for} \quad t \leqq s\,. \end{aligned}$$

We say that a probability measure P on $\mathscr{F}$ defines *a Markov process* (X_s, P)

with a transition function p^h if, for all $s \in S$ and $A \in \mathscr{F}_{>s}$,

$$P\{A \mid \mathscr{F}_{\leqq s}\} = P^h_{s,X_s}(A) \quad \text{a.s. } P \qquad \text{on} \quad \{\omega: \alpha < s < \beta\} .$$

Let $\Lambda = \{t_1 < \cdots < t_m\}$ be a finite subset of S and let $t_0 = -\infty$, $t_{m+1} = +\infty$. Put

$$\alpha_\Lambda = t_{k+1} \quad \text{if} \quad t_k \leqq \alpha < t_{k+1} , \qquad \beta_\Lambda = t_k \quad \text{if} \quad t_k < \beta \leqq t_{k+1} ,$$
$$k = 0, 1, \cdots, m ;$$

$$\Pi_\Lambda f(\omega) = \Pi_{(\alpha_\Lambda(\omega), +\infty)} f(\omega) .$$

Theorem 9.1 can be extended to processes with random birth and death times as follows.

THEOREM 11.1. *Put* $A \in \mathscr{F}^0$ *if* $\{A, \alpha \leqq s\} \in \mathscr{F}_{\leqq s}$ *for all* $s \in S$. *Then* $\mathscr{F}^0$ *is an H-sufficient σ-algebra for* $M(p^h)$. *To each increasing sequence of finite sets* Λ_n *with a union everywhere dense in S, there corresponds an asymptotically H-sufficient statistic* Π_{Λ_n}.

Now all the results of Section 10 can be easily carried over to excessive measure and functions. We have to replace s_n by α_{Λ_n} in (10.4) and u_n by β_{Λ_n} in (10.6).

11.2. We proved in [5] that the space of all p-excessive measures is a Borel space (the main point is that each p-excessive measure is defined uniquely by the values ν_t for rational t). Therefore all simplexes $M(p)$, $R_p{}^h$, $S_\nu{}^p$ investigated in Sections 9, 10 and 11 are Borel spaces.

12. Stationary transition functions.

12.1. We suppose now that a Markov transition function $p(s, x; t, \Gamma)$ is *stationary* which means that: (i) S is a subgroup of the additive group of real numbers; (ii) all spaces $(E_s, \mathscr{B}_s) = (E, \mathscr{B})$ are identical; (iii) $p(s, x; t, \Gamma) = p(t - s, x, \Gamma)$ depends only on the difference $t - s$. We shall consider only two possibilities: S is the group of all integers (the discrete case) and S is the group of all real numbers (the continuous case). In the second case we assume that $p(t, x, \Gamma)$ is measurable with respect to the pair t, x.

We denote by θ_t a shift in $(E^S, \mathscr{B}^S)$ which corresponds to the transformation $s \to s + t$ of S. A Markov process (X_s, P) is called *stationary* if P is invariant with respect to the group θ_t, $t \in S$.

THEOREM 12.1. *Let* $\mathscr{F}^0 = \bigcap \mathscr{F}_{\leqq s}$ *be the tail σ-algebra and let* $\mathscr{F}_\theta{}^0$ *be a collection of all* $A \in \mathscr{F}^0$ *which are invariant with respect to the group* θ_t. *Then* $\mathscr{F}_\theta{}^0$ *is an H-sufficient σ-algebra for the class* $M_\theta(P)$ *of all stationary Markov processes with a transition function p.*

PROOF. In the discrete case we can apply the corollary to Lemma 6.1 to $T = \theta_1$, the class $M = M(p)$, and σ-algebra $\mathscr{F}^0$ (which are invariant with respect to T). In the continuous case $(E^S, \mathscr{B}^S)$ is not a B-space. This obstacle can be overcome in the same way as in the proof of Theorem 9.1 but we will not go into details.

12.2. A probability measure ν is called a *stationary distribution* for p if $\nu P_t = \nu$ for all $t \in S$. (Here $P_t = P_t^0$ are the Markov operators associated with the transition function p.) Each stationary distribution defines an entrance law $\nu_t = \nu$, $t \in S$. The corresponding Markov process P_ν^1 belongs to the class $M_\theta(p)$ investigated in Theorem 12.1. In this way we establish an isomorphism between the class N of all stationary distributions and $M_\theta(p)$. Hence N is a simplex.

12.3. An excessive measure ν_t is called *stationary* if ν_t does not depend on t. Obviously a measure ν on $(E, \mathscr{B})$ is a stationary excessive measure if and only if $\nu P_t \leqq \nu$ for all $t \in S$ and $\nu P_t \uparrow \nu$ as $t \downarrow 0$. In a similar way, we introduce the concept of a stationary excessive function.

THEOREM 12.2. *Let l be a strictly positive measurable function on $(E, \mathscr{B})$. A class of all stationary excessive measures ν normed by the condition $\nu(l) = 1$ is a simplex.*

PROOF. Since p is stationary, the formula $(T_t\nu)_s = \nu_{s+t}$ defines for each t a transformation of the set of all p-excessive measures. Obviously, ν is stationary if and only if it is invariant with respect to the group T_t.

Consider a p-excessive function

$$h^s(x) = \tfrac{1}{2} \int_0^\infty e^{-|s+u|} P_u l(x)\, du .$$

A simple calculation shows that, for each excessive measure ν,

$$\{\nu, h\} = \tfrac{1}{2} \int_{-\infty}^{+\infty} e^{-|u|} \nu_u(l)\, du \tag{12.1}$$

which implies that

$$\{\nu, h\} = \nu(l) \qquad \text{if} \quad \nu \quad \text{is stationary,} \tag{12.2}$$

and

$$\{T_s\nu, h\} \leqq e^{|s|}\{\nu, h\} . \tag{12.3}$$

Denote by M^* the set of all excessive measures ν satisfying the condition $\{\nu, h\} < \infty$ and by M the set of $\nu \in M^*$ for which $\{\nu, h\} = 1$. According to Section 11, M is a Borel space and a simplex. It follows from (12.3) that M^* is invariant with respect to T_t, and Theorem 12.2 follows from Theorem 6.3.

12.4.

THEOREM 12.3. *Suppose a stationary transition function $p(t, x, \Gamma)$ is absolutely continuous with respect to a measure γ for each t and x. Then the set of all stationary excessive functions h normed by the condition $\gamma(h) = 1$ is a simplex.*

PROOF. We consider transformations $(T_t h)^s = h^{t+s}$ of the set of all excessive functions. The formula

$$\nu_t(\Gamma) = \tfrac{1}{2} \int_0^\infty e^{-|t-u|} (\gamma P_u)(\Gamma)\, du$$

defines an excessive measure, and we have

$$\{\nu, h\} = \tfrac{1}{2} \int_{-\infty}^{+\infty} e^{-|u|} \gamma(h^u)\, du$$

for every excessive function h. In particular $\{\nu, h\} = \gamma(h)$ for a stationary h. To complete the proof, we apply Theorem 6.3 in the same way as in the proof of Theorem 12.2.

REMARK. Put

$$g_\lambda(x, \Gamma) = \int_0^\infty e^{-\lambda t} p(t, x, \Gamma)\, dt\,, \quad \lambda \geqq 0\,.$$

If a measure

$$\eta_\lambda(\Gamma) = \int_E \gamma(dx) g_\lambda(x, \Gamma)$$

is σ-finite for some λ, then Theorem 12.2 remains true if $p(t, x, \Gamma)$ is absolutely continuous with respect to η_λ (Kuznecov [14], Theorem 3).

REFERENCES

[1] BURKHOLDER, D. L. (1961). Sufficiency in the undominated case. *Ann. Math. Statist.* **32** 1191–1200.

[2] DOBRUSHIN, R. L. (1968). Description of a random field by means of conditional probabilities and the conditions governing its regularity. *Theor. Probability Appl.* **13** 197–224.

[3] DOOB, J. L. (1959). Discrete potential theory and boundaries. *J. Math. and Mech.* **8** 433–458.

[4] DYNKIN, E. B. (1971). Initial and final behavior of trajectories of Markov processes. *Uspehi Mat. Nauk* **26** 4 153–172. (English translation: *Russian Math. Surveys* **26** 4 165–185.)

[5] DYNKIN, E. B. (1972). Integral representation of excessive measures and excessive functions. *Uspehi Mat. Nauk* **27** 1 43–80. (English translation: *Russian Math. Surveys* **27** 1 43–84.)

[6] DYNKIN, E. B. (1975). Markov representations of stochastic systems. *Uspehi Mat. Nauk* **30** 1 61–99. (English translation: *Russian Math. Surveys* **30** 1 65–104.)

[7] DYNKIN, E. B. and JUSKEVIČ, A. A. (1975). *Controlled Markov Processes and Their Applications.* Nauka, Moscow. (English translation will be published by Springer-Verlag.)

[8] FARRELL, R. H. (1962). Representation of invariant measures. *Illinois J. Math.* **6** 3 447–467.

[9] DE FINETTI, B. (1931). Fuzione caratteristica di un fenomeno aleatorio. *Atti Accad. Naz. Lincei, Mem. Cl. Sci. Fis. Mat. Natur.* **4** 251–299.

[10] DE FINETTI, B. (1937). La prévision: ses lois logiques, ses sources subjectives. *Ann. Inst. H. Poincaré* **7** 1–68.

[11] HEWITT, E. and SAVAGE, L. J. (1955). Symmetric measures on cartesian products. *Trans. Amer. Math. Soc.* **80** 2 470–501.

[12] HUNT, G. A. (1960). Markov chains and Martin boundaries. *Illinois J. Math.* **4** 313–340.

[13] KRYLOV, N. and BOGOLIOUBOV, N. (1937). La théorie générale de la mesure dans son application à l'étude des systèmes de la mécanique non linéaires. *Ann. of Math.* **38** 65–113.

[14] KUZNECOV, S. E. (1974). On decomposition of excessive functions. *Dokl. Akad. Nauk SSSR* **214** 276–278.

[15] MARTIN, R. S. (1941). Minimal positive harmonic functions. *Trans. Amer. Math. Soc.* **49** 137–172.

[16] MEYER, P. A. (1966). *Probability and Potential.* Blaisdell, Waltham, Mass.

[17] NAIMARK, M. A. (1964). *Normed Rings.* P. Noordhoff N. V., Groningen.

[18] NEVEU, J. (1964). Bases mathématiques du calcul des probabilités. Masson et C^{ie}, Paris.

[19] PARTHASARATHY, K. R. (1967). *Probability Measures on Metric Spaces.* New York, London.

[20] PRESTON, C. I. (1974). Gibbs states on countable sets. Cambridge Univ. Press.

[21] VARADARAJAN, V. S. (1963). Groups of automorphisms of Borel spaces. *Trans. Amer. Math. Soc.* **109** 2 191-220.

DEPARTMENT OF MATHEMATICS
CORNELL UNIVERSITY
ITHACA, N.Y. 14853