13. *Ergodic Theorems and the Markoff Process with a Stable Distribution.*

By Shizuo KAKUTANI.

Mathematical Institute, Osaka Imperial University.

(Comm. by T. TAKAGI, M.I.A., March 12, 1940.)

1. *Introduction.* A Markoff process $P(t, E)$[1] is called to have a *stable distribution* $\varphi(E)$, if there exists a completely additive non-negative set function $\varphi(E)$ (with $\varphi(\Omega)=1$) defined for all Borel set $E \subset \Omega$ such that

$$(1) \qquad \int_\Omega \varphi(dt) P(t, E) = \varphi(E) \quad \text{for any Borel set } E \subset \Omega .$$

For example, the Markoff process defined by a φ-measure preserving transformation $S(t)$: $P(t, E)=1$ if $S(t) \in E$, $=0$ if $S(t) \bar{\in} E$, has $\varphi(E)$ as its stable distribution; and another example thereof is given by the Markoff process with symmetric φ-density $p(t, s)$: $P(t, E) = \int_E p(t, s)\,\varphi(ds)$, $p(t, s)=p(s, t)$.

It is the purpose of the present paper to discuss such a class of Markoff processes. The same problem is also discussed by K. Yosida[2] in the preceding paper. He has proved that a mean ergodic theorem (for exact formulation see Theorem 2 below) holds for any Markoff process with stable distributions. Since this class of Markoff processes contains the case of measure preserving transformations, his result is a generalization of the mean ergodic theorem of J. v. Neumann.[3] In the present paper, we shall first prove (Theorem 1) that even an ergodic theorem of G. D. Birkhoff's type[4] is valid for such a class of Markoff processes. Indeed, we shall prove that for any *bounded* Borel measurable function $x(t)$ defined on Ω the sequence $\left\{\frac{1}{N}\sum_{n=1}^{N} x_n(t)\right\}$ $(N=1, 2, \dots)$, where

$$(2) \qquad x_n(t) = \int_\Omega P^{(n)}(t, ds)\, x(s) , \qquad n=1, 2, \cdots ,$$

converges φ-almost everywhere on Ω. This result is, in essential, due to J. L. Doob.[5] We shall next show that the mean ergodic theorem

---

1) As for the notions concerning Markoff process, see:

S. Kakutani: Some results in the operator-theoretical treatment of the Markoff process, Proc. **15** (1939), 260–264. K. Yosida: Operator-theoretical treatment of the Markoff process, Proc. **14** (1938), 363–367, Proc. **15** (1939), 127–130.

2) K. Yosida: Markoff process with stable distribution, Proc. **16** (1940), 43–48.

3) J. v. Neumann: Proof of the quasi-ergodic hypothesis, Proc. Nat. Acad. U.S.A., **18** (1932), 70–82.

4) G. D. Birkhoff: Proof of the ergodic theorem, Proc. Nat. Acad. U.S.A., **18** (1932), 650–655.

5) J. L. Doob: Stochastic processes with an integral valued parameter, Trans. Amer. Math. Soc., **44** (1938), 87–150.

follows at once from this result (see the proof of Theorem 2). Consequently, Theorem 1 is more precise than the result of K. Yosida. (But the proof of K. Yosida is simpler and elegant). It is, however, to be remarked that the complete analogy of G. D. Birkhoff's ergodic theorem is not yet established in the indeterministic case. Indeed, the φ-almost everywhere convergence of the sequence $\left\{\frac{1}{N}\sum_{n=1}^{N} x_n(t)\right\}$ $(N=1, 2, \ldots)$ for any φ-integrable function $x(t)$, which is expected by analogy, is not yet proved for general Markoff processes with stable distributions.

In § 3, the inverse probability and inverse Markoff process are discussed, and it will be seen that these notions are useful in many problems concerning the Markoff process with stable distributions.

2. *Operator-theoretical considerations.* We have often[1)] observed that

$$x \to T(x) = y: \quad y(E) = \int_\Omega x(dt)\, P(t, E) \tag{3}$$

is a positive bounded linear operation of norm 1 which maps the Banach space $(\mathfrak{M})$ of all the completely additive real valued set functions $x(E)$ ($\|x\|$ = total variation of $x(E)$ on Ω) into itself, and that

$$x \to \bar{T}(x) = y: \quad y(t) = \int_\Omega P(t, ds)\, x(s) \tag{4}$$

is a positive bounded linear operation of norm 1 which maps the Banach space (M^*) of all the bounded Borel measurable real valued functions $x(t)$ ($\|x\|$ = l. u. b. of $|x(t)|$ on Ω) into itself. From the standpoint of this operator theory, the existence of a stable distribution $\varphi(E)$ means that $\varphi(E)$ is a positive proper element of T belonging to the proper value 1.

If we now consider the Banach space $L(\varphi)$ of all the φ-integrable Borel measurable real valued functions $x(t)$ $\left(\|x\| = \int_\Omega |x(t)|\, \varphi(dt)\right)$, then $\bar{T}$ defines also a bounded linear operation of norm 1 which maps $L(\varphi)$ into itself. Moreover, $\bar{T}$ may also be considered as a bounded linear operation of norm 1 which maps the Banach space $M(\varphi)$ of all the bounded Borel measurable real valued functions $x(t)$ ($\|x\|$ = φ-essential maximum of $|x(t)|$ on Ω) into itself.

On the other hand, the operation T maps a φ-absolutely continuous function $x(E) \in (\mathfrak{M})$ into a φ-absolutely continuous function $y(E) \in (\mathfrak{M})$. Since the subspace of all the φ-absolutely continuous functions of $(\mathfrak{M})$ is isometric with $L(\varphi)$, T may also be considered as a bounded linear operation of norm 1 which maps $L(\varphi)$ into itself:

$$x \to T(x) = y: \quad \int_E y(t)\, \varphi(dt) = \int_\Omega x(t)\, \varphi(dt)\, P(t, E) \tag{5}$$

for any Borel set $E \subset \Omega$. It will again be clear that the same operation may be considered as a bounded linear operation of norm 1 which maps $M(\varphi)$ into itself.

Thus the Markoff process with a stable distribution $\varphi(E)$ defines two classes of bounded linear operations T and $\bar{T}$, which map the Banach spaces $(\mathfrak{M})$, $L(\varphi)$, $M(\varphi)$ and (M^*), $L(\varphi)$, $M(\varphi)$ respectively into themselves.

3. *Inverse probability and the inverse Markoff process.* Since $Q(F, E) \equiv \int_F \varphi(dt) P(t, E) \leqq \varphi(E)$ and $Q(\Omega, E) = \varphi(E)$ for any Borel sets F and E, there exists for any F a bounded Borel measurable real valued function $R(F, s)$ such that $0 \leqq R(F, s) \leqq 1$, $R(\Omega, s) = 1$ and $Q(F, E) = \int_E R(F, s) \varphi(ds)$ for any Borel sets F and E. $R(F, s)$ is determined up to a set of φ-measure zero for each F. If we assume that $R(F, s)$ is completely additive in F for any fixed s, then $R(F, s)$ defines a Markoff process which is inverse to the given Markoff process $P(t, E)$. Indeed, $R(F, s)$ is the inverse probability that a point $s \in \Omega$ is transferred from the point of the Borel set F by the Markoff process $P(t, E)$. It is clear that $\varphi(E)$ is also a stable distribution for $R(F, s)$, and that $P(t, E)$ is again the inverse of $R(F, s)$. Further, it will be almost clear that in this case (5) becomes:

$$(6) \qquad x \rightarrow T(x) = y: \quad y(s) = \int_\Omega x(t) R(dt, s).$$

(It is to be noted that, without assuming the complete additivity of $R(F, s)$ in F, $y(s)$ is determined only up to a set of φ-measure zero). In this way, the complete duality between $P(t, E)$ and $R(F, s)$ will be established. But we shall not discuss this problem in detail here.

4. *Measures in the product space.* Consider a space Ω^* whose element t^* is a sequence of points $t^* = \{\overset{(i)}{t_i}\}$ $(i = 0, \pm 1, \pm 2, \ldots)$, where t_i is an arbitrary point of Ω, and the notation (i) written over t_i denotes that t_i is the i-th coordinate of t^*. We may write symbolically: $\Omega^* = \cdots \overset{(-2)}{\Omega} \times \overset{(-1)}{\Omega} \times \overset{(0)}{\Omega} \times \overset{(1)}{\Omega} \times \overset{(2)}{\Omega} \times \cdots$. We shall introduce a completely additive measure $\varphi^*(E^*)$ on Ω^*. Let us first consider the subset E^* of Ω^* of the form: $E^* = \cdots \times \overset{(m-2)}{\Omega} \times \overset{(m-1)}{\Omega} \times \overset{(m)}{E_m} \times \overset{(m+1)}{E_{m+1}} \times \cdots \times \overset{(n)}{E_n} \times \overset{(n+1)}{\Omega} \times \overset{(n+2)}{\Omega} \times \cdots$ or more symbolically: $E^* = \overset{(m)}{E_m} \times \overset{(m+1)}{E_{m+1}} \times \cdots \times \overset{(n)}{E_n}$, where $-\infty < m \leqq n < \infty$ and $E_i (m \leqq i \leqq n)$ is a Borel set of Ω. Strictly speaking, E^* is the set of all the sequences $t^* = \{\overset{(i)}{t_i}\}$ $(i = 0, \pm 1, \pm 2, \ldots)$ such that $t_i \in E_i$ for $m \leqq i \leqq n$, and $t_i \in \Omega$ for $i \leqq m-1$ and $i \geqq n+1$. Let us denote by $\mathfrak{E}^*$ the family of all such sets E^*, and by $\mathfrak{F}^*$ the family of all the subsets F^* of Ω^* of the form: $F^* = \sum_{i=1}^{n} E_i^*$, $E_i^* \cdot E_j^* = 0$ $(i \neq j)$, $E_i^* \in \mathfrak{E}^*$ $(i = 1, 2, \ldots, n)$. If we put

$$\varphi^*(E^*) = \int_{E_m} \int_{E_{m+1}} \cdots \int_{E_n} \varphi(dt_m) P(t_m, dt_{m+1}) P(t_{m+1}, dt_{m+2}) \ldots P(t_{n-1}, dt_n)$$

$$\equiv \int_{E_m} \int_{E_{m+1}} \cdots \int_{E_n} R(dt_m, t_{m+1}) R(dt_{m+1}, t_{m+2}) \ldots R(dt_{n-1}, t_n) \varphi(dt_n)$$

and $\varphi^*(F^*) = \sum_{i=1}^{n} \varphi^*(E_i^*)$, then $\mathfrak{F}^*(F^*)$ is clearly a finitely additive measure on $\mathfrak{F}^*$ (with $\varphi^*(\Omega^*)=1$). By a theorem of A. Kolmogoroff-J. L. Doob, $\varphi^*(F^*)$ is even completely additive on $\mathfrak{F}^*$. Consequently, by a well-known result, $\varphi^*(F^*)$ can be extended to a completely additive measure $\varphi^*(F^*)$ on the least Borel family $B(\mathfrak{F}^*)$ which contains $\mathfrak{F}^*$. It will be almost clear that this $\varphi^*(F^*)$ is invariant under the one-to-one mapping (translation) S of Ω^* into itself: $S(t^*)=t^{*\prime}$, where $t^*=\{\overset{(i)}{t_i}\}$ and $t^{*\prime}=\{\overset{(i)}{t_{i+1}}\}$ ($i=0, \pm 1, \pm 2, \ldots$) (i. e., the i-th coordinate of t^* is the $(i+1)$-th coordinate of t^*).

We shall next introduce another kind of measures on Ω^*. Let $\overline{\mathfrak{E}}^*(\overline{\overline{\mathfrak{E}}}^*)$ be the family of all the sets E^* of the form: $E^*=\overset{(m)}{E_m}\times \overset{(m+1)}{E_{m+1}}\times \cdots \times \overset{(n)}{E_n}$ $(E^*=\overset{(-n)}{E_{-n}}\times \overset{(-n+1)}{E_{-n+1}}\times \cdots \times \overset{(-m)}{E_{-m}})$ with $1\leqq m\leqq n$, and consider the family $\overline{\mathfrak{F}}^*(\overline{\overline{\mathfrak{F}}}^*)$ of all the sets F^* of the form: $E^*=\sum_{i=1}^{n}E_i^*$, $E_i^*\cdot E_j^*=0$ $(i\neq j)$, $E_i^*\in\overline{\mathfrak{E}}^*(\overline{\overline{\mathfrak{E}}}^*)$ $(i=1, 2, \ldots, n)$. If we put

$$\bar\varphi_{t_0}^*(E^*)=\int_{E_m}\int_{E_{m+1}}\cdots\int_{E_n}P^{(m)}(t_0, dt_m)P(t_m, dt_{m+1})P(t_{m+1}, dt_{m+2})\cdots P(t_{n-1}, dt_n)$$

$$\Big(\bar{\bar\varphi}_{t_0}^*(E^*)=\int_{E_{-n}}\int_{E_{-n+1}}\cdots\int_{E_{-m}}R(dt_{-n}, t_{-n+1})R(dt_{-n+1}, t_{-n+2})\cdots$$
$$R(dt_{-m-1}, t_{-m})R^{(m)}(dt_{-m}, t_0)\Big)$$

and $\bar\varphi_{t_0}^*(F^*)=\sum_{i=1}^{n}\bar\varphi_{t_0}^*(E_i^*)$ $\big(\bar{\bar\varphi}_{t_0}^*(F^*)=\sum_{i=1}^{n}\bar{\bar\varphi}_{t_0}^*(E_i^*)\big)$, then $\bar\varphi_{t_0}^*(F^*)$ $\big(\bar{\bar\varphi}_{t_0}^*(F^*)\big)$ is finitely additive on $\overline{\mathfrak{F}}^*(\overline{\overline{\mathfrak{F}}}^*)$, and it may be easily seen that $\bar\varphi_{t_0}^*(F^*)$ $\big(\bar{\bar\varphi}_{t_0}^*(F^*)\big)$ is even completely additive on $\overline{\mathfrak{F}}^*(\overline{\overline{\mathfrak{F}}}^*)$. Hence $\bar\varphi_{t_0}^*(F^*)$ $\big(\bar{\bar\varphi}_{t_0}^*(F^*)\big)$ can be extended to $B(\overline{\mathfrak{F}}^*)$ $\big(B(\overline{\overline{\mathfrak{F}}}^*)\big)$, and it will be easily seen that we have

$$(7)\qquad \int_{\Omega}\varphi(dt_0)\bar\varphi_{t_0}^*(E^*)=\varphi^*(E^*)\quad \Big(\int_{\Omega}\varphi(dt_0)\bar{\bar\varphi}_{t_0}(E^*)=\varphi^*(E^*)\Big)$$

for any $E^*\in B(\overline{\mathfrak{F}}^*)$ $\big(E^*\in B(\overline{\overline{\mathfrak{F}}}^*)\big)$.

5. *Individual ergodic theorem.*

Theorem 1. Let $P(t, E)$ be a Markoff process with a stable distribution $\varphi(E)$. Then for any $x(t)\in M(\varphi)$ there exists an $\bar x(t)\in M(\varphi)$ such that

$$(8)\qquad \frac{1}{N}\sum_{n=1}^{N}x_n(t)\to\bar x(t)\qquad \varphi\text{-almost everywhere on } \Omega,$$

where $x_n(t)$ $(n=1, 2, \ldots)$ is defined by (2) $\big($or equivalently, by (4), $x_n=\bar T^n(x)\big)$.

Proof. By the ergodic theorem of G. D. Birkhoff in $L(\varphi^*)$ on Ω^*, there exists for any $x^*(t^*)\in L(\varphi^*)$ an $\bar x^*(t^*)\in L(\varphi^*)$ such that

$$(9)\qquad \frac{1}{N}\sum_{n=1}^{N}x^*\big(S^n(t^*)\big)\to\bar x^*(t^*)\qquad \varphi^*\text{-almost everywhere on } \Omega^*.$$

As a special case put $x^*(t^*)=x(t_0)$, where $t^*=\{\overset{(i)}{t_i}\}$ $(i=0, \pm 1, \pm 2, \ldots)$ and $x(t) \in M(\varphi)$. Then $x^*(t^*) \in M(\varphi^*) \subset L(\varphi^*)$, and (9) becomes:

$$\frac{1}{N}\sum_{n=1}^{N} x(t_n) \to \bar{x}^*(t^*), \quad \varphi^*\text{-almost everywhere on } \Omega^*. \tag{10}$$

Let E_0^* be the set of φ^*-measure zero where the convergence (10) might not hold. We may assume that $E_0^* \in B(\overline{\mathfrak{F}}^*)$. If we have $\bar{\varphi}_{t_0}^*(E_0^*)=0$ for some $t_0 \in \Omega$, then we have, by integrating (10) with respect to $\bar{\varphi}_{t_0}^*(E_0^*)$ all over Ω^*,

$$\frac{1}{N}\sum_{n=1}^{N} x_n(t_0) \equiv \frac{1}{N}\sum_{n=1}^{N} P^{(n)}(t_0, dt_n)x(t_n) \to \int_{\Omega^*} \bar{\varphi}_{t_0}^*(dt^*)\,\bar{x}^*(t^*) \equiv \bar{x}(t_0).$$

Hence, all what we have to prove is that there exists a set E_0 of φ-measure zero in Ω such that $t_0 \bar{\in} E_0$ implies $\bar{\varphi}_{t_0}^*(E_0^*)=0$. This is, however, clear since we have, by (7), $\int_\Omega \varphi(dt_0)\bar{\varphi}_{t_0}^*(E_0^*)=\varphi^*(E_0^*)=0$.

Remark. If we assume the complete additivity of $R(F, s)$, then by considering S^{-1}, $R(F, s)$ and $\bar{\bar{\varphi}}_{t_0}^*(E^*)$ instead of S, $P(t, E)$ and $\bar{\varphi}_{t_0}^*(E^*)$ respectively, there exists for any $x(t) \in M(\varphi)$ an $\bar{\bar{x}}(t) \in M(\varphi)$ such that

$$\frac{1}{N}\sum_{n=1}^{N} x_{-n}(t) \to \bar{\bar{x}}(t) \qquad \varphi\text{-almost everywhere on } \Omega, \tag{11}$$

where $x_{-n}(s)=\int_\Omega x(t)R^{(n)}(dt, s)$ $\big($or equivalently, by (6), $x_{-n}=T^n(x)\big)$.

6. *Deduction of the mean ergodic theorem from the individual ergodic theorem.*

Theorem 2. Let $P(t, E)$ be a Markoff process with a stable distribution $\varphi(E)$. Then there exists for any $x(t) \in L(\varphi)$ an $\bar{x}(t) \in L(\varphi)$ such that

$$\int_\Omega \left|\frac{1}{N}\sum_{n=1}^{N} x_n(t)-\bar{x}(t)\right|\varphi(dt) \equiv \left\|\frac{1}{N}\sum_{n=1}^{N} \bar{T}^n(x)-\bar{x}\right\| \to 0, \tag{12}$$

where $x_n=\bar{T}^n(x)$ $(n=1, 2, \ldots)$ is defined by (2), and $\|\cdots\|$ means the norm in $L(\varphi)$.

Proof. We shall first treat the case when $x(t) \in M(\varphi)$. By Theorem 1, there exists an $\bar{x}(t) \in M(\varphi)$ such that (8) is true. Since $\{x_n(t)\}$ $(n=1, 2, \ldots)$ is uniformly bounded, we have (12) by integrating (8) with respect to $\varphi(E)$ all over Ω. In order to prove the general case, we have only to notice that the sequence $\left\{\frac{1}{N}\sum_{n=1}^{N} \bar{T}^n\right\}$ $(N=1, 2, \ldots)$ of uniformly bounded linear operations which map $L(\varphi)$ into itself converges strongly at any $x(t) \in M(\varphi) \subset L(\varphi)$. Since $M(\varphi)$ is dense in $L(\varphi)$, the same sequence converges strongly at any $x(t) \in L(\varphi)$.

Remark 1. In the special case of a φ-measure preserving transformation, $\bar{T}$ may also be considered as a bounded linear operation of norm 1 which maps $L^p(\varphi)$ $(p \geqq 1)$ into itself. Hence, by the same arguments as in above (since $M(\varphi)$ is dense in $L^p(\varphi)$), the mean ergodic theorem holds in $L^p(\varphi)$ $(p \geqq 1)$. This result is also obtained by M. Fukamiya[6] by appealing to the dominated ergodic theorem of N. Wiener.[7]

Remark 2. Theorem 2 may be deduced from the mean ergodic theorem in $L(\varphi^*)$ on Ω^*. Indeed, since S is φ^*-measure preserving on Ω^*, there exists for any $x^*(t^*) \in L(\varphi^*)$ an $\bar{x}^*(t^*) \in L(\varphi^*)$ such that

$$\int_{\Omega^*} \left| \frac{1}{N} \sum_{n=1}^{N} x^*(S^n(t^*)) - \bar{x}^*(t^*) \right| \varphi^*(dt^*) \to 0 . \tag{13}$$

In the special case when $x^*(t^*) = x(t_0)$, where $t^* = \{\overset{(i)}{t_i}\}$ $(i = 0, \pm 1, \pm 2, \dots)$ and $x(t) \in L(\varphi)$, (13) implies, by (7),

$$\int_{\Omega} \left| \int_{\Omega^*} \left(\frac{1}{N} \sum_{n=1}^{N} x(t_n) - \bar{x}^*(t^*) \right) \bar{\varphi}^*_{t_0}(dt^*) \right| \varphi(dt_0) \to 0 ;$$

and this relation is equivalent to (12) if we put $\bar{x}(t_0) \equiv \int_{\Omega^*} \bar{x}^*(t^*) \bar{\varphi}^*_{t_0}(dt^*)$.

Furthermore, if we consider S^{-1}, $R(F, s)$ and $\bar{\bar{\varphi}}^*_{t_0}(E^*)$ instead of S, $P(t, E)$ and $\bar{\varphi}^*_{t_0}(E^*)$ respectively, then we have, in the same way as in above,

$$\int_{\Omega} \left| \frac{1}{N} \sum_{n=1}^{N} x_{-n}(t) - \bar{\bar{x}}(t) \right| \varphi(dt) \equiv \left\| \frac{1}{N} \sum_{n=1}^{N} T^n(x) - \bar{\bar{x}} \right\| \to 0 , \tag{14}$$

where $x_{-n} = T^n(x)$ $(n = 1, 2, \dots)$ is defined as in (11). It is to be remarked that, as is shown by K. Yosida in the preceding paper, the existence of the inverse probability (i. e., the complete additivity of $R(F, s)$) is not needed for the validity of (14).

Remark (Added by the proof). As is kindly pointed out by K. Yosida, we may assume that the inverse probability $R(F, s)$ is completely additive in F if Ω is separable with respect to the measure $\varphi(E)$. This follows by the same arguments as in J. L. Doob (see footnote (5)).

---

6) M. Fukamiya: On dominated ergodic theorem in $L^p (p > 1)$, Tôhoku Math. Journ., **46** (1939), 150–153.

7) N. Wiener: The homogeneous chaos, Amer. Journ. of Math., **60** (1938), 897–936. N. Wiener: The ergodic theorem, Duke Math. Journ., **5** (1939), 1–18.