THEORY OF PROBABILITY AND ITS APPLICATIONS

Volume II — Number 1

1957

SHORT COMMUNICATIONS

AN ERGODIC THEOREM FOR MARKOV PROCESSES AND ITS APPLICATION TO TELEPHONE SYSTEMS WITH REFUSALS

B. A. SEVAST'YANOV

Many problems of the telephone type are solved by considering Markov processes in appropriate phase spaces. This is greatly simplified if the telephone system operates under stationary conditions. Consideration of a stationary distribution is justified in the case of ergodicity, i.e. when the probability distribution of the process converges to a stationary distribution as $t \to \infty$ for all initial distributions.

In § 1 of this paper an ergodic theorem for Markov processes is proved in a form which in the author's opinion is convenient for problems of the telephone type. Theorems on the limiting behaviour of transition probabilities of Markov stationary processes were obtained by Doob [1]. In § 2 of the present paper, under certain conditions which are probably necessary, there are proved the existence and uniqueness of a stationary distribution which is ergodic. With more particular assumptions, the ergodicity and uniqueness of a stationary distribution are proved by A. M. Yaglom [2].

In the following paragraphs an ergodic theorem is applied to the study of a telephone system with n serving lines, a Poisson flow of incoming calls, and an arbitrary distribution law with finite mean value for the duration of conversation. A stationary distribution is obtained from which in particular Erlang's well known formulas in telephony are deduced. Up to the present time the strict derivation of Erlang's formulas has been achieved only in the case of an exponential distribution law for the duration of conversation (see [3]). The validity of Erlang's formulas for an absolutely continuous distribution law was shown by Fortet [4]. But the proof of the uniqueness and ergodicity of the stationary distribution is not given in [4] and only a brief indication is given concerning the possibility of such a proof. The method of proof of the existence, uniqueness and ergodicity of a stationary distribution based on the theorem in § 1 can be applied to various problems connected with mass service.

1. An Ergodic Theorem

Let Ω be an abstract space and F be a Borel field of measurable sets in this space containing Ω. The function $P(x, t, A)$, $x \in \Omega$, $A \in F$, $t > 0$, will define the transition probabilities of a homogeneous Markov process in time t from state n into the set A if

1) $0 \leqq P(x, t, A) \leqq P(x, t, \Omega) = 1$;

2) for fixed x and t the function $P(x, t, A)$ is a fully additive function of the set A; for fixed A and t the function $P(x, t, A)$ is measurable in x with respect to the field of sets F;

3) for $t > 0$, $s > 0$ the following equation is satisfied:

$$P(x, s+t, A) = \int_\Omega P(x, s, dy) P(y, t, A).$$

If P_0 is the initial probability distribution, the probability distribution P_t at time $t > 0$ is defined as follows:

$$P_t(A) = \int_\Omega P_0(dx) P(x, t, A);$$

P is called a stationary probability distribution if, for any $t > 0$ and $A \in F$,

$$P(A) = \int_\Omega P(dx) P(x, t, A).$$

We denote by $V(P, Q)$ the distance function defined in terms of the variation between the distributions P and Q

$$V(P, Q) = \operatorname{Var}(P-Q) = \int_\Omega |P(dx) - Q(dx)|.$$

The distribution P is called ergodic if, for every initial distribution, $V(P_t, P) \to 0$ as $t \to \infty$.

Theorem 1. *A Markov process homogeneous in time, has a unique stationary probability distribution which is ergodic if, for any $\varepsilon > 0$, there exists a measurable set C, a probability distribution R in Ω, and $t_1 > 0$, $k > 0$, $K > 0$ such that*

1) $kR(A) \leqq P(x, t_1, A)$ *for all points $x \in C$ and measurable sets $A \subseteq C$; for any initial distribution P_0 there exists a t_0 such that for any $t \geqq t_0$*
2) $P_t(C) \geqq 1-\varepsilon$,
3) $P_t(A) \leqq KR(A)+\varepsilon$ *for all measurable sets $A \subseteq C$.*

The proof of this theorem is deduced from the following five lemmas.

Lemma 1. *If $t' < t''$, then $V(P_{t'}, Q_{t'}) \geqq V(P_{t''}, Q_{t''})$.*

PROOF.

$$V(P_{t''}, Q_{t''}) = \int_\Omega |P_{t''}(dy)-Q_{t''}(dy)| = \int_{\Omega_y} \left| \int_{\Omega_x} [P_{t'}(dx)-Q_{t'}(dx)] \right| P(x, t''-t', dy)$$
$$\leqq \int_\Omega |P_{t'}(dx)-Q_{t'}(dx)| \int_\Omega P(x, t''-t', dy) = \int_\Omega |P_{t'}(dx)-Q_{t'}(dx)| = V(P_{t'}, Q_{t'}).$$

Lemma 2. *Under the conditions of Theorem 1, for all initial distributions P_0 and Q_0*

$$V(P_t, Q_t) \to 0 \text{ as } t \to \infty.$$

PROOF. By virtue of Lemma 1, $V(P_t, Q_t)$ does not increase as t increases. Hence there exists a limit $\lim_{t\to\infty} V(P_t, Q_t) = V$. We prove that $V = 0$. Let us suppose that it is not so. We choose $\varepsilon > 0$ such that $V' = V-4\varepsilon > 0$. We reach a contradiction after showing that $V(P_{t_0+nt_1}, Q_{t_0+nt_1}) \to 0$ as $n \to \infty$, where t_0 and t_1 are the same quantities as appear in the conditions of Theorem 1 corresponding to the chosen $\varepsilon > 0$ and the initial distributions P_0 and Q_0. We represent $V_t = V(P_t, Q_t)$ and $V_{t+t_1} = V(P_{t+t_1}, Q_{t+t_1})$ in the following form ($A_+ \cup A_- = \Omega$, $B_+ \cup B_- = \Omega$):

$$V_t = \int_\Omega |P_t(dx)-Q_t(dx)| = \int_{A_+} [P_t(dx)-Q_t(dx)] + \int_{A_-} [Q_t(dx)-P_t(dx)] = 2[P_t(A_+)-Q_t(A_+)],$$
$$V_{t+t_1} = \int_\Omega |P_{t+t_1}(dx)-Q_{t+t_1}(dx)| = \int_{B_+} [P_{t+t_1}(dx)-Q_{t+t_1}(dx)]$$
$$+ \int_{B_-} [Q_{t+t_1}(dx)-P_{t+t_1}(dx)] = 2[P_{t+t_1}(B_+)-Q_{t+t_1}(B_+)] = 2[Q_{t+t_1}(B_-)-P_{t+t_1}(B_-)].$$

We introduce the notation $P_{t, t+t_1}(A, B)$ and $Q_{t, t+t_1}(A, B)$ for the two-dimensional distributions which equal the probabilities that the system lies in the set A at time t and in the set B at time $t+t_1$, the initial distributions being P_0 and Q_0, respectively. Then we can write

$$V_t = 2[P_{t, t+t_1}(A_+, B_+)+P_{t, t+t_1}(A_+, B_-)-Q_{t, t+t_1}(A_+, B_+)-Q_{t, t+t_1}(A_+, B_-)],$$
$$V_{t+t_1} = 2[P_{t, t+t_1}(A_+, B_+)+P_{t, t+t_1}(A_-, B_+)-Q_{t, t+t_1}(A_+, B_+)-Q_{t, t+t_1}(A_-, B_+)],$$

whence

$$V_t - V_{t+t_1} = 2[P_{t, t+t_1}(A_+, B_-)-Q_{t, t+t_1}(A_+, B_-)]+2[Q_{t, t+t_1}(A_-, B_+)-P_{t, t+t_1}(A_-, B_+)]$$
$$= 2\int_{A_+} [P_t(dx)-Q_t(dx)]P(x, t_1, B_-)+2\int_{A_-} [Q_t(dx)-P_t(dx)]P(x, t_1, B_+)$$
$$\geqq 2\int_{A_+} [P_t(dx)-Q_t(dx)]\, P(x, t_1, B_-).$$

Let C be a measurable set, R a probability distribution, and $k > 0$, $K > 0$ correspond to the chosen $\varepsilon > 0$ under the conditions of Theorem 1. We put $A'_+ = A_+ \cap C$, $B'_- = B_- \cap C$. Then

$$V_t - V_{t+t_1} \geqq 2\int_{A'_+} [P_t(dx)-Q_t(dx)]\, P(x, t_1, B'_-) \geqq 2\int_{A'_+} [P_t(dx)-Q_t(dx)]\, kR(B'_-)$$
$$\geqq 2[P_t(A'_+)-Q_t(A'_+)]\, \frac{k}{K}\, [Q_{t+t_1}(B'_-)-\varepsilon] \tag{1}$$
$$\geqq \frac{k}{2K}\, 2[P_t(A'_+)-Q_t(A'_+)] \cdot 2[Q_{t+t_1}(B'_-)-P_{t+t_1}(B'_-)-\varepsilon].$$

Further, let $t \geqq t_0$. Since $P_t(C) \geqq 1-\varepsilon$ and $Q_t(C) \geqq 1-\varepsilon$,

$$P_t(A'_+) \geqq P_t(A_+)-\varepsilon, \qquad Q_{t+t_1}(B'_-) \geqq Q_{t+t_1}(B_-)-\varepsilon.$$

Since $Q_t(A'_+) \leqq Q_t(A_+)$ and $P_{t+t_1}(B'_-) \leqq P_{t+t_1}(B_-)$, it follows from (1) that

$$V_t - V_{t+t_1} \geqq \frac{k}{2K}(V_t - 4\varepsilon)(V_{t+t_1} - 4\varepsilon).$$

We put $V'_t = V_t - 4\varepsilon$. Then we can write the above inequality in the form

$$V'_t - V'_{t+t_1} \geqq \frac{k}{2K} V'_t V'_{t+t_1},$$

whence it follows that

$$\frac{1}{V'_{t+t_1}} - \frac{1}{V'_t} \geqq \frac{k}{2K}.$$

Putting $t = t_0,\ t_0+t_1,\ t_0+2t_1$, etc., and adding the inequalities which are obtained, we have

$$\frac{1}{V_{t_0+nt_1}} \geqq \frac{kn}{2K} + \frac{1}{V'_{t_0}} > \frac{kn}{2K},$$

whence it follows that, as $n \to \infty$,

$$V'_{t_0+nt_1} < \frac{2K}{kn} \to 0, \tag{2}$$

and this contradicts the supposition that $V' > 0$. Hence $V = 0$.

Corollary. *Under the conditions of Theorem* 1, *there can exist not more than one stationary distribution P, and, moreover,* $V(P_t, P) \to 0$ *always as* $t \to \infty$ *for every initial distribution* P_0.

This corollary is an obvious consequence of Lemma 2.

Lemma 3. *The metric space* P_Ω *of distributions of P over* Ω *with metric* $V(P, Q)$ *is complete.*

Proof. Let P_n be a fundamental sequence of distributions. This means that for any $\varepsilon > 0$ there exists n_0 such that for $n \geqq n_0$ and any integer $m > 0$

$$V(P_n, P_{n+m}) \leqq \varepsilon.$$

Since $V(P, Q) = 2\max_{A\in F} |P(A) - Q(A)|$,

$$|P_n(A) - P_{n+m}(A)| \leqq \frac{\varepsilon}{2} \tag{3}$$

for $n \geqq n_0$, i.e., for any measurable A the sequence of numbers $P_n(A)$ is fundamental. Hence there exists the limit

$$\lim_{n\to\infty} P_n(A) = P(A).$$

The function $P(A)$ of the sets forms a probability distribution. Only the continuity property requires proof. Let the sequence of measurable sets $A_1 \supseteq A_2 \supseteq \cdots \supseteq A_k \supseteq \cdots$ be such that their intersection is empty. Then, for any n, $P_n(A_k) \to 0$ as $k \to \infty$. We show that $P(A_k) \to 0$ also as $k \to \infty$. For this purpose we substitute A_k for A in the inequality (3), so that for $n \geqq n_0$

$$|P_n(A_k) - P_{n+m}(A_k)| \leqq \frac{\varepsilon}{2}.$$

Since the inequality holds for any m, it holds also in the limit as $m \to \infty$:

$$|P_{n_0}(A_k) - P(A_k)| \leqq \frac{\varepsilon}{2}.$$

Let $P_{n_0}(A_k) \leqq \varepsilon/2$ for $k \geqq k_0$. Then for $k \geqq k_0$ the inequality $P(A_k) \leqq \varepsilon$ will hold. The lemma is proved.

Lemma 4. *If the conditions of Theorem* 1 *are satisfied, for every initial probability distribution* P_0 *and the sequence* $t_n \to \infty$, P_{t_n} *is a fundamental sequence in the space* P_Ω.

Proof. It follows immediately from Lemma 2 that, for any fixed τ as $t \to \infty$,

$$V(P_t, P_{t+\tau}) \to 0. \tag{4}$$

To prove that convergence to the limit in (4) is uniform with respect to τ, one must refer back to the proof of Lemma 2. The evaluation of (2) depends on the arbitrary $\varepsilon > 0$ and on the original probability distributions P_0 and Q_0. In this case the initial distributions are P_0 and P_τ. Since these distributions lie on one path, the t_0 which appears in the conditions of Theorem 1 does not depend on τ. Hence we obtain uniform convergence to zero in (4).

Lemma 5. *If for a certain initial distribution Q_0 and for $t \to \infty$, $V(Q_t, P) \to 0$, then P is a stationary probability distribution.*

PROOF. Let $P_0 = P$. We show that for any $t > 0$, $P_t \equiv P$. In fact, for any $\varepsilon > 0$ there exists a τ_0 such that for $\tau \geqq \tau_0$,

$$V(P_t, P) \leqq V(P_t, Q_{t+\tau}) + V(Q_{t+\tau}, P) \leqq V(Q_\tau, P) + V(Q_{t+\tau}, P) \leqq \varepsilon.$$

Hence $V(P_t, P) \equiv 0$ and $P_t \equiv P$.

PROOF OF THEOREM 1. It follows from Lemmas 3, 4, 5, that there exists a stationary distribution, and from Lemma 2 follow its uniqueness and ergodicity.

2. A Stationary Distribution for a Telephone System with Refusals

Let there be n telephone lines subjected to a Poisson flow of calls with parameter λ. In the case of the exponential distribution law for the duration of conversation with mean μ, Erlang's well known classical formulas for the probabilities p_k of k lines being occupied when operating under stationary conditions are

$$p_k = \frac{\dfrac{(\lambda\mu)^k}{k!}}{\displaystyle\sum_{i=0}^{n} \frac{(\lambda\mu)^i}{i!}}, \qquad k = 0, 1, 2, \cdots, n. \tag{5}$$

In particular, the probability of the loss of a call is equal to p_n. This case is examined using a Markov process with a finite number of states (3). It will be proved below that Erlang's formulas (5) remain valid for an arbitrary distribution law with a finite mean μ. In this case also one can make a study of the telephone system with the aid of a homogeneous Markov process, but here it is necessary to consider a Markov process in a more complex phase space. Most convenient will be a phase space consisting of an isolated point ω_0, a half-line ω_1 $(0 \leqq x_1 < \infty)$, part of a plane ω_2 $(0 \leqq x_1, x_2 < \infty)$, $\cdots$, and part of an n-dimensional space ω_n $(0 \leqq x_1, x_2, \cdots, x_n < \infty)$.

We establish the following correspondence between the points of the phase space and the states of the system. The point ω_0 corresponds to that state of the system when all the lines are free. The point $(x_1, x_2, \cdots, x_k)$ in ω_k, $k = 1, 2, \cdots, n$, corresponds to k lines being occupied and conducting k conversations which have already lasted for times $x_1, x_2, \cdots, x_k$. In what follows we shall denote the point in ω_k with co-ordinates $(x_1, x_2, \cdots, x_k)$ by $\omega_k(x_1, x_2, \cdots, x_k)$.

We assume that the lengths of conversations are independent of one another and of the flow of calls. In this case the corresponding random process will be a Markov process. This process is determined by the following transition probabilities for a small interval of time Δt:

$$P(\omega_0, \Delta t, \omega_0) = 1 - \lambda\Delta t + o(\Delta t) \quad \text{(no call made)},$$

$$P(\omega_0, \Delta t, \omega_1(0)) = \lambda\Delta t + o(\Delta t) \quad \text{(one call made)},$$

$$P(\omega_k(x_1, x_2, \cdots, x_k), \Delta t, \omega_k(x_1+\Delta t, x_2+\Delta t, \cdots, x_k+\Delta t))$$
$$= \prod_{i=1}^{k} \frac{1-F(x_i+\Delta t)}{1-F(x_i)} \left(1-\lambda\Delta t + o(\Delta t)\right) \quad \text{(no call made and no conversation completed)},$$

$$P(\omega_n(x_1, x_2, \cdots, x_n), \Delta t, \omega_n(x_1+\Delta t, x_2+\Delta t, \cdots, x_n+\Delta t))$$
$$= \prod_{i=1}^{n} \frac{1-F(x_i+\Delta t)}{1-F(x_i)} \quad \text{(no conversation completed)},$$

$$P(\omega_k(x_1, x_2, \cdots, x_k), \Delta t, \omega_{k-1}(x_1+\Delta t, \cdots, x_{i-1}+\Delta t, x_{i+1}+\Delta t, \cdots, x_k+\Delta t))$$
$$= \frac{F(x_i+\Delta t)-F(x_i)}{1-F(x_i)} \prod_{\substack{j\neq i \\ 1\leqq j\leqq k}} \frac{1-F(x_j+\Delta t)}{1-F(x_j)} \left(1-\lambda\Delta t+o(\Delta t)\right)$$

(conversation completed having lasted a time x_i but no call made·

$$P(\omega_k(x_1, x_2, \cdots, x_k), \Delta t, \omega_{k+1}(x_1+\Delta t, \cdots, x_{j-1}+\Delta t, 0, x_j+\Delta t, \cdots, x_k+\Delta t))$$
$$= \prod_{i=1}^{k} \frac{1-F(x_i+\Delta t)}{1-F(x_i)} \left(\frac{\lambda}{k+1} \Delta t + o(\Delta t)\right)$$

(one call made, occupying one of the free lines, but no conversation completed)

The remaining transition probabilities either have a higher order of smallness or do not enter into the integro-differential equation (9), since after averaging they give probabilities of the order $o(\Delta t)$. Perhaps the last transition probability requires an explanation. We have not so far fixed the order of occupation of the free lines by the incoming calls. Although it is not essential for our purposes, it is convenient to consider the whole system as being symmetrical so that there are identical probabilities of an incoming call occupying any one of the free lines. In setting up the correspondence between the point $\omega_k(x_1, x_2, \cdots, x_k)$ and the state of the telephone system we did not relate in any way the order of $x_1, x_2, \cdots, x_k$ and the number of engaged lines and the order of the incoming calls. Therefore all the points of the phase space $\omega_k(x_{i_1}, x_{i_2}, \cdots, x_{i_k})$, where $x_{i_1}, x_{i_2}, \cdots, x_{i_k}$ are arranged in some order $x_1, x_2, \cdots, x_k$, correspond essentially to one state of the system. But in considering the Markov process it is convenient for us to regard all these points as different. For symmetry of the Markov process in the phase space it is convenient to assume in the case of an incoming call that the point $\omega_k(x_1, x_2, \cdots, x_k)$ can change with the same probability $1/(k+1)$ into any one of the points

$$\omega_{k+1}(x_1, x_2, \cdots, x_{i-1}, 0, x_i, \cdots, x_k), \qquad i = 1, 2, \cdots, k+1.$$

We denote the probability distribution in the phase space by P_t (for $t = 0$ we shall have the initial distribution P_0) and the stationary distribution by P. In what follows we shall always assume that the distribution of P_0 over ω_k is symmetrical with respect to the variables $x_1, x_2, \cdots, x_k$. Therefore P_t will also possess this symmetry. We also introduce the notation $p_k(t) = P_t(\omega_k)$ and $p_k = P(\omega_k)$, $k = 0, 1, 2, \cdots, n$. The probability distribution over the paths of the Markov process (for any initial distribution) will be denoted by $\mathscr{P}$.

Theorem 2. *For every distribution P_0, the distribution P_t has at the point $\omega_k(x_1, x_2, \cdots, x_k)$ a k-dimensional probability density $p_k(x_1, x_2, \cdots, x_k; t)$, if $t > \max\{x_1, x_2, \cdots, x_n\}$, where*

$$p_k(x_1, x_2, \cdots, x_k; t) \leqq \frac{\lambda^k}{k!}[1-F(x_1)] \cdots [1-F(x_k)], \qquad k = 1, 2, \cdots, n. \tag{6}$$

Proof. We denote by A the event that the system is situated at time t in ω_k in the set of points whose co-ordinates $y_1, y_2, \cdots, y_k$ satisfy the inequalities $x_i < y_i < x_i+\Delta_i$, $i = 1, 2, \cdots, k$. We denote by B the event that calls are placed in the time intervals $(t-(x_i+\Delta_i), t-x_i)$, $i = 1, 2, \cdots, k$, but the conversations started thereby do not finish by the time t. Clearly $A \subseteq B$ for $t > \max_{1 \leqq i \leqq k} x_i$. Therefore

$$P_t\{\omega_k(y_1, y_2, \cdots, y_k) : x_i < y_i < x_i+\Delta_i;\ i = 1, 2, \cdots, k\} = \mathscr{P}(A) \leqq \mathscr{P}(B)$$
$$\leqq \frac{\lambda^k}{k!}[1-F(x_1)] \cdots [1-F(x_k)]\Delta_1\Delta_2 \cdots \Delta_k,$$

whence the existence of the probability density $p_k(x_1, x_2, \cdots, x_k; t)$ and inequality (6) follow.

Corollary. *The stationary distribution P has probability densities $p_k(x_1, x_2, \cdots, x_k)$, $k = 1, 2, \cdots, n$, satisfying inequality* (6).

Proof. If we put $P_0 = P$, then $P_t \equiv P$ for any t. Since t may be taken as large as we please, there exists, by Theorem 2, a probability density $p_k(x_1, x_2, \cdots, x_k)$ in the whole set ω_k.

Theorem 3. *If $F(x)$ has a finite mean μ, the Markov process defined above has the following stationary distribution*:

$$p_k(x_1, x_2, \cdots, x_k) = \frac{\lambda^k}{k!}p_0[1-F(x_1)] \cdots [1-F(x_k)], \qquad k = 1, 2, \cdots, n, \tag{7}$$

$$p_0 = \frac{1}{\sum_{k=0}^{n} \frac{(\lambda\mu)^k}{k!}}.$$

Proof. First of all we derive the equations which are satisfied by the functions

$$p_k^*(x_1, x_2, \cdots, x_k; t) = \frac{p_k(x_1, x_2, \cdots, x_k; t)}{[1-F(x_1)][1-F(x_2)] \cdots [1-F(x_k)]}, \qquad k = 0, 1, \cdots, n. \tag{8}$$

For this purpose we express the probability distribution at time $t+\Delta t$ in terms of the probability distribution at time t using the total probability formula. Making use of the Markov properties of the process and the transition probabilities for time Δt, we obtain

$$p_0(t+\Delta t) = p_0(t)(1-\lambda\Delta t)+\int_0^\infty p_1(x_1;\, t)\,\frac{F(x_1+\Delta t)-F(x_1)}{1-F(x_1)}\,dx_1+o(\Delta t),$$

$$p_1(x_1,\, t+\Delta t) = p_1(x_1-\Delta t,\, t)\,\frac{1-F(x_1)}{1-F(x_1-\Delta t)}\,(1-\lambda\Delta t)$$

$$+\,2\int_0^\infty p_2(x_1-\Delta t,\, x_2;\, t)\,\frac{1-F(x_1)}{1-F(x_1-\Delta t)}\cdot\frac{F(x_2+\Delta t)-F(x_2)}{1-F(x_2)}\,dx_2+o(\Delta t),$$

$$p_k(x_1, x_2, \cdots, x_k;\, t+\Delta t) = p_k(x_1-\Delta t, x_2-\Delta t, \cdots, x_k-\Delta t;\, t)\prod_{i=1}^{k}\frac{1-F(x_i)}{1-F(x_i-\Delta t)}\,(1-\lambda\Delta t)$$

$$+\,(k+1)\int_0^\infty p_{k+1}(x_1-\Delta t, x_2-\Delta t, \cdots, x_k-\Delta t, x_{k+1};\, t)\prod_{j=1}^{k}\frac{1-F(x_j)}{1-F(x_j-\Delta t)}$$

$$\times\frac{F(x_{k+1}+\Delta t)-F(x_{k+1})}{1-F(x_{k+1})}\,dx_{k+1}+o(\Delta t),$$

$$p_n(x_1, x_2, \cdots, x_n;\, t+\Delta t) = p_n(x_1-\Delta t, x_2-\Delta t, \cdots, x_n-\Delta t,\, t)\prod_{i=1}^{n}\frac{1-F(x_i)}{1-F(x_i-\Delta t)}\,.$$

We assume the existence of the derivatives $\partial p_k^*/\partial t$, $\partial p_k^*/\partial x_i$, $i = 1, 2, \cdots, k$; $k = 0, 1, 2, \cdots, n$. From these equalities we then obtain the following system of integro-differential equations [1]

$$\frac{\partial p_0^*}{\partial t} + \lambda p_0^* = \int_0^\infty p_1^*(x_1,\, t)\,dF(x_1),$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\frac{\partial p_k^*}{\partial t}+\frac{\partial p_k^*}{\partial x_1}+\cdots+\frac{\partial p_k^*}{\partial x_k}+\lambda p_k^* = (k+1)\int_0^\infty p_{k+1}^*(x_1, x_2, \cdots, x_k, x_{k+1};\, t)\,dF(x_{k+1}), \tag{9}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\frac{\partial p_n^*}{\partial t}+\frac{\partial p_n^*}{\partial x_1}+\cdots+\frac{\partial p_n^*}{\partial x_n} = 0.$$

From the relations

$$p_k(x_1, x_2, \cdots, x_k;\, t)\prod_{i=1}^{k}\frac{1-F(x_i+\Delta t)}{1-F(x_i)}\left(\frac{\lambda\Delta t}{k+1}+o(\Delta t)\right) = p_{k+1}(x_1+\Delta t, \cdots, x_k+\Delta t, 0;\, t+\Delta t)$$

we obtain the boundary conditions

$$\lambda p_k^*(x_1, x_2, \cdots, x_k;\, t) = (k+1)p_{k+1}^*(x_1, x_2, \cdots, x_k, 0;\, t), \qquad k = 0, 1, 2, \cdots, n-1. \tag{10}$$

If the functions (8) correspond to a stationary distribution, then all derivatives with respect to t vanish and system (9) is written in the form

$$\lambda p_0^* = \int_0^\infty p_1^*(x_1)\,dF(x_1)$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\frac{\partial p_k^*}{\partial x_1}+\cdots+\frac{\partial p_k^*}{\partial x_n}+\lambda p_k^* = (k+1)\int_0^\infty p_{k+1}^*(x_1, x_2, \cdots, x_k, x_{k+1})\,dF(x_{k+1}) \tag{11}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\frac{\partial p_n^*}{\partial x_1}+\cdots+\frac{\partial p_n^*}{\partial x_n} = 0.$$

Conversely any probability distribution satisfying system (11) and the boundary conditions (10) is stationary. The solution of system (11) subject to conditions (10) is

$$p_k^* = p_0\,\frac{\lambda^k}{k!}\,, \qquad k = 0, 1, 2, \cdots, n. \tag{12}$$

---

[1] In deriving equations (9) we make use of the fact that, for a continuous bounded function $g(x)$ and a distribution function $F(x)$, the following equality holds:

$$\lim_{t\to 0}\int_{-\infty}^{\infty} g(x)\,\frac{F(x+t)-F(x)}{t}\,dx = \int_{-\infty}^{\infty} g(x)dF(x).$$

This is easily verified by substitution. Therefore distribution (7) is stationary. Here p_0 is determined by the normality conditions

$$p_0+p_1+\cdots+p_n=p_0+\int_0^\infty p_1(x)dx+\cdots+\int_0^\infty\cdots\int_0^\infty p_n(x_1,x_2,\cdots,x_n)\,dx_1dx_2\cdots dx_n$$

$$=p_0\left[1+\lambda\int_0^\infty[1-F(x_1)]\,dx_1+\cdots+\frac{\lambda^n}{n!}\int_0^\infty\cdots\int_0^\infty[1-F(x_1)]\cdots[1-F(x_n)]dx_1dx_2\cdots dx_n\right.$$

$$=p_0\sum_{k=0}^n\frac{(\lambda\mu)^k}{k!}=1.$$

The proof of uniqueness and ergodicity of this stationary distribution will be a consequence of Theorem 1.

3. Proof of Uniqueness and Ergodicity for a Stationary Process

We now show that the Markov process under consideration satisfies the condition of Theorem 1. Below we shall keep throughout to the following notation. We denote by C_l the set consisting of the points ω_0 and

$$\{\omega_k(x_1,x_2,\cdots,x_k);\ 0\leqq x_i\leqq l;\ i=1,2,\cdots,k\},\qquad k=1,2,\cdots,n.$$

For the probability distribution R we shall assume the stationary distribution defined in formula (7).

Theorem 4. *Given any $\varepsilon>0$, there exists an $l>0$ such that for every initial distribution P_0 it is possible to find a t_0 such that for $t\geqq t_0$*

$$P_t(C_l)\geqq 1-\varepsilon.$$

Proof [2]. We show that if we choose $l>0$ according to $\varepsilon>0$ and t_0 according to P_0, it is always possible to ensure that for $t\geqq t_0$ the probability of an event arising in the complement to the set C_l does not exceed ε, i.e.

$$P_t(\bar{C}_l)\leqq\varepsilon.$$

We denote by ξ_{it} the random variable which equals zero if at time t the line with number i is free, and equals the time from the beginning of the conversation to the instant t if the i-th line is occupied. We denote the distribution function of ξ_{it} by $G_t(x)$. One can express the distribution function $G_0(x)$ in terms of the distribution P_0. All the ξ_{it}, $i=1,2,\cdots,n$, have the same distribution function $G_t(x)$. It is not difficult to see that

$$P_t(\bar{C}_l)\leqq\sum_{i=1}^n\mathbf{P}\{\xi_{it}>l\}.$$

We shall show how one must select $l>0$ according to $\varepsilon>0$ and a value t_0 according to P_0, or rather $G_0(x)$, so that for $t>t_0$

$$\mathbf{P}\{\xi_{it}>l\}\leqq\frac{\varepsilon}{n}\cdot\tag{13}$$

Let $t>l$. We evaluate the three integrals on the right-hand side of the equation

$$\mathbf{P}\{\xi_{it}>l\}=\int_l^t dG_t(x)+\int_t^{t+l_1}dG_t(x)+\int_{t+l_1}^\infty dG_t(x)$$

By a method similar to that used in Theorem 2, one can show that

$$\int_l^t dG_t(x)\leqq\lambda\int_l^t[1-F(x)]dx\leqq\lambda\int_l^\infty[1-F(x)]dx.$$

For $x>t$ the function $G_t(x)$ can be expressed in terms of $G_0(x)$ as follows

$$dG_t(x)=\frac{1-F(x)}{1-F(x-t)}\,dG_0(x-t).$$

[2] In the proofs of Theorems 4 and 5 it is assumed that $F(x)<1$ for all $0\leqq x<\infty$. Modifications of the proof in the case $F(x)=1$ for $x\geqq x_0$ are almost obvious.

Therefore

$$\int_t^{t+l_1} dG_t(x) = \int_t^{t+l_1} \frac{1-F(x)}{1-F(x-t)}\, dG_0(x-t)$$

and

$$\int_{t+l_1}^{\infty} dG_t(x) = \int_{t+l_1}^{\infty} \frac{1-F(x)}{1-F(x-t)}\, dG_0(x-t).$$

Let there be given an $\varepsilon > 0$. We choose l so that

$$\lambda \int_l^{\infty} [1-F(x)]\, dx < \frac{\varepsilon}{3n}. \tag{14}$$

By virtue of the convergence of the integral $\int_0^\infty [1-F(x)]dx = \mu$ this is always possible. We choose l_1 so that $\int_{l_1}^\infty dG_0(x) < \varepsilon/3n$, and t_0 so that $(1-F(t_0))/(1-F(l_1)) < \varepsilon/3n$. Then for $t < t_0$

$$\begin{aligned} &\int_l^{t+l_1} \frac{1-F(x)}{1-F(x-t)}\, dG_0(x-t) \leqq \frac{1-F(t)}{1-F(l_1)} \leqq \frac{1-F(t_0)}{1-F(l_1)} < \frac{\varepsilon}{3n}, \\ &\int_{t+l_1}^{\infty} \frac{1-F(x)}{1-F(x-t)}\, dG_0(x-t) \leqq \int_{t+l_1}^{\infty} dG_0(x-t) = \int_{l_1}^{\infty} dG_0(x) < \frac{\varepsilon}{3n}. \end{aligned} \tag{15}$$

Inequality (13) follows from (14) and (15) and hence the theorem is true.

Theorem 5. *There exists a $t_1 > 0$ and a $k > 0$ such that for all points $x \in C_l$, $l > 0$, and measurable sets $A \subseteq C_l$ the following inequality holds*:

$$P(x, t_1, A) \geqq kR(A). \tag{16}$$

Proof. We choose t_2 so that $F(t_2) - F(l) = \eta > 0$, $t_3 > l$, and we put $t_1 = t_2 + t_3$. We show that inequality (16) is satisfied for $k = \eta^n e^{-\lambda t_1}$.

Since

$$P(x, t_1, A) \geqq P(x, t_2, \omega_0) P(\omega_0, t_3, A),$$

it is sufficient for us to show that for $x \in C_l$ and for measurable $A \subseteq C_l$

$$P(x, t_2, \omega_0) \geqq \eta^n e^{-\lambda t_2} \tag{17}$$

and

$$P(\omega_0, t_3, A) \geqq e^{-\lambda t_3} R(A). \tag{18}$$

Let $x = \omega_k(x_1, x_2, \cdots, x_k) \in C_l$. Then

$$P(x, t_2, \omega_0) \geqq e^{-\lambda t_2} \prod_{i=1}^{k} \frac{F(t_2+x_i) - F(x_i)}{1-F(x_i)} \geqq e^{-\lambda t_2} \eta^k \geqq e^{-\lambda t_2} \eta^n,$$

since the event, whose probability is stated on the left-hand side, includes the event in which no call occurs in the time t_2 and all conversations in progress are completed.

For $t > l$ there exist, according to Theorem 2, in C_l transition probability densities $p(\omega_0, t, x_1, x_2, \cdots, x_k)$ from ω_0 into $\omega_k(x_1, x_2, \cdots, x_k)$. We shall show that

$$p(\omega_0, t_3, x_1, x_2, \cdots, x_k) \geqq e^{-\lambda t_3} \frac{\lambda^k}{k!} [1-F(x_1)] \cdots [1-F(x_k)], \tag{19}$$

from which (18) will follow.

Suppose that at time $t = 0$ the system is in the state ω_0. We denote by A the event in which k conversations are in progress at time t_3, the conversations having begun in the intervals $(t_3 - x_i - dx_i,\ t_3 - x_i)$, $i = 1, 2, \cdots, k$. By B we denote the event in which k conversations were begun in the intervals $(t_3 - x_i - dx_i,\ t_3 - x_i)$, $i = 1, 2, \cdots, k$, none of these being concluded at time t_3 and there were no other calls. Since $A \supseteq B$,

$$\begin{aligned} p(\omega_0, t_3, x_1, \cdots, x_k) dx_1 \cdots dx_k &= \mathscr{P}(A) \geqq \mathscr{P}(B) \\ &= e^{-\lambda t_3} \frac{\lambda^k}{k!} [1-F(x_1)] \cdots [1-F(x_k)]\, dx_1 \cdots dx_k, \end{aligned}$$

from which (19) and (18) follow. The theorem is proved.

Theorem 6. *There exists such a $K > 0$ that for every initial distribution P_0, for $t > l$ and for any measurable set $A \subseteq C_l$*

$$P_t(A) \leqq KR(A).$$

PROOF. This inequality is true for

$$K = \frac{1}{p_0} = \sum_{k=0}^{n} \frac{(\lambda\mu)^k}{k!}.$$

This follows at once from inequality (6), Theorem 2, and equalities (7), which determine the stationary distribution R.

Theorems 4, 5, and 6 show that the ergodic theorem in § 1 is applicable to the study of a Markov process. Consequently the stationary distribution (7) is unique and ergodic. Hence it follows that

$$\lim_{t\to\infty} p_k(t) = p_k = \frac{\dfrac{(\lambda\mu)^k}{k!}}{\displaystyle\sum_{i=0}^{n} \frac{(\lambda\mu)^i}{i!}}, \qquad k = 0, 1, 2, \cdots, n.$$

Received October 30, 1956.

REFERENCES

[1] J. L. DOOB, *Asymptotic properties of Markoff transition probabilities*, Trans. Amer. Math. Soc. 63, 1948, pp. 393–421.

[2] A. M. YAGLOM, *On the statistical reversibility of Brownian motion*, Matem. sb., 24 (66), No. 3, 1949, pp. 457–492.

[3] A. YA. KHINCHIN, *Mathematical methods of the theory of mass service*, Trudy Mat. Instituta im. V. A. Steklova, 49, 1955.

[4] R. FORTET, *Calcul des probabilités*, Centre national de la recherche scientifique, 1950.

AN ERGODIC THEOREM FOR MARKOV PROCESSES AND ITS APPLICATION TO TELEPHONE SYSTEMS WITH REFUSALS

B. A. SEVAST'YANOV (*MOSCOW*)

(*Summary*)

Let $P(x, t, A)$ be the probability of transition from state $x \in \Omega$ to a measurable set A in time t of a homogeneous Markov process in the space Ω. Let $P_t(A)$ be the probability that the system is in the measurable set A at time t if at $t = 0$ the distribution P_0 is valid. The stationary distribution P is called ergodic if for each initial distribution P_0 the variation $V(P_t - P) \to 0$ for $t \to \infty$.

The following ergodic theorem is proved:

Theorem 1. *A Markov process homogeneous in time has a unique stationary probability distribution which is ergodic if for any $\varepsilon > 0$ there exists a measurable set C, a probability distribution R in Ω, and values $t_1 > 0$, $k > 0$, $K > 0$ such that*

1) $kR(A) \leqq P(x, t_1, A)$ *for all points $x \in C$ and measurable sets $A \subseteq C$; for any initial distribution P_0 there exists a t_0 such that for any $t \geqq t_0$*

2) $P_t(C) \geqq 1-\varepsilon$;

3) $P_t(A) \leqq KR(A)+\varepsilon$ *for any measurable set $A \subseteq C$.*

This theorem is used to substantiate Erlang's formulas (5) for the case of an arbitrary distribution law with a finite mean value for the duration of conversation.

SOME REMARKS ON EVALUATING THE ACCURACY OF LINEAR EXTRAPOLATION AND FILTRATION

A. I. KORONKEVICH

(*Translated by R. A. Silverman*)

Usually in linear extrapolation and filtration of a stationary random function $\xi(t)$, the accuracy of solution is evaluated as the size of the mean square error $\sigma_T{}^2$, which has the units of the