A LAW OF SMALL NUMBERS IN MARKOFF CHAINS

BY

B. O. KOOPMAN

1. Introduction. The object of this paper is to establish for repeated trials in a Markoff chain a limiting distribution similar to the Poisson distribution in the case of independent trials. Each trial has two outcomes ("success" and "failure") and their sequence has a one-step dependence. The case of stationary transitions has already been treated by elementary methods; it leads to a Laguerre distribution[1]. In this paper the general nonstationary case is studied; a recent theorem in analysis is required[2].

Let $\{U_{n,k}\}$ $(n=1, 2, \cdots ; k=1, \cdots, n)$ be an infinite triangular array of variates (chance variables). The row $(U_{n,1}, \cdots, U_{n,n})$ represents a sequence of n trials in a Markoff chain, $U_{n,k}$ having the value 1 or 0 according as the kth trial succeeds or fails. Denote by $P_n(s)$ $(0 \le s \le n)$ the probability of exactly s successes in this sequence of trials. Our object is to find $P(s)$ $= \lim_{n \to \infty} P_n(s)$ under suitable hypotheses regarding $\{U_{n,k}\}$.

We introduce the absolute probabilities[3] $(1 \le k \le n)$

$$(1.1) \qquad p_{n,k} = p(U_{n,k} = 1), \qquad q_{n,k} = 1 - p_{n,k} \qquad (1 \le k \le n),$$

and the transition probabilities $(2 \le k \le n)$

$$(1.2) \qquad \begin{aligned} a_{n,k} &= p(U_{n,k} = 1 \mid U_{n,k-1} = 1), & a'_{n,k} &= 1 - a_{n,k}, \\ b_{n,k} &= p(U_{n,k} = 1 \mid U_{n,k-1} = 0), & b'_{n,k} &= 1 - b_{n,k}; \end{aligned}$$

and also, the expected total number of successes

$$(1.3) \qquad m(n) = \sum_{k=1}^{n} p_{n,k} = \sum_{s=0}^{n} s P_n(s).$$

In the case of independent trials $(a_{n,k} = b_{n,k} = p_{n,k})$, it is shown[4] that Poisson's distribution $P_n(s) \to e^{-m} m^s / s!$ is obtained if and only if $m(n) \to m$ and $p_{n,k} \to 0$ uniformly over k, as $n \to \infty$.

The $a_{n,k}$, $b_{n,k}$, $p_{n,k}$, are related by the difference equation $(2 \le k \le n)$

(1.4) $p_{n,k} = r_{n,k}p_{n,k-1} + b_{n,k}, \qquad r_{n,k} = a_{n,k} - b_{n,k},$

which in the case of Markoffian dependence determines every $p_{n,k}$ in terms of the initial probability $p_{n,1}$. Accordingly, any unconditional statement concerning $p_{n,k}$ (such as that $\lim_{n\to\infty} p_{n,k} = 0$) largely expresses a property of $p_{n,1}$, an "accidental" feature of the situation. On the other hand, a property of the *transition probabilities* reflects a more "intrinsic" property of the mechanism by means of which the trials are made. Such properties are the following: *The transition probabilities* (thought of as fixed, independently of $p_{n,k}$) *are such that*

(1.5) $\lim_{n\to\infty} p_{n,1} = 0$ *implies that* $\lim_{n\to\infty} \max_{1\le k\le n} p_{n,k} = 0;$

(1.6) $\lim_{n\to\infty} p_{n,1} = p_1$ *implies that* $\lim_{n\to\infty} m(n) = m.$

And we shall say that a limit $P_n(s)\to P(s)$, obtained under any hypothesis implying (1.5) and (1.6), is a *law of small numbers*. This evidently includes the Poisson case when $a_{n,k} = b_{n,k}$.

We shall establish such a law when, in addition to (1.5), relation (1.6) is amplified as follows:

(1.7) $\lim_{n\to\infty} p_{n,1} = p_1$ *implies that* $\lim_{n\to\infty} \sum_{k=0}^{n-v} E(U_{n,k}U_{n,k+v}) = R_v,$

for all $v = 0, 1, 2, \cdots$ (evidently $R_0 = m$); and when, furthermore, infinite sequences $\{a_i\}$, $\{c_i\} \neq \{$all $0\}$, exist such that

(1.8) $a_{n,k} \to a_k \le a < 1$ *uniformly over k as $n \to \infty$*;

and for all n large enough so that $\{c_2, \cdots, c_n\} \neq \{0, \cdots, 0\}$, *and all $2 \le k \le n$,*

$$b_{n,k}\left[\sum_{i=2}^{n} b_{n,i}\right]^{-1} = c_k\left[\sum_{i=2}^{n} c_i\right]^{-1} \cdot (1 + \epsilon_{n,k}),$$

(1.9)
$$\epsilon_{n,k} \to 0 \quad \textit{uniformly over } k \textit{ as } n \to \infty.$$

As we shall soon see, $b_{n,k}\to 0$; then (1.9) expresses a requirement of uniformly limiting proportionality, $b_{n,2}:b_{n,3}:b_{n,4}: \cdots \to c_2:c_3:c_4: \cdots$.

Hypotheses (1.8) and (1.9) may be regarded as expressing the heuristic conception that the mechanism by means of which the trials are made "approaches a limiting mechanism," so that the quantitative aspects of the former approach those of the latter "in detail," as $n\to\infty$.

Under these hypotheses, we show that (when $p_{n,1}\to p_1$) the generating function

(1.10) $\phi_n(t) = \sum_{s=0}^{n} P_n(s)t^s$

approaches a limit $\phi(t)$ uniformly, and that $\phi(t)$ is given explicitly as an exponential expression involving only the generating function of $\{R_v\}$ when $p_1 = 0$, and involving in addition that of $\{a_i\}$ when $p_1 \neq 0$ (cf. (8.1), (8.2)).

We may note that when the proportionality condition (1.9) is replaced by the requirement that for all sufficiently great n, $b_{n,k} = 0$ $(k = 1, \cdots, n)$, our theorem is a trivial consequence of §5.

A natural generalization suggests itself: Let $\{U_{n,k}\}$ be replaced by $\{X_{n,k}\}$, where $(X_{n,1}, \cdots, X_{n,n})$ is a one-step Markoff chain, but where each variate $X_{n,k}$ can have any values in $(-\infty, +\infty)$ (or is a vector variate, confined only to an Euclidean N-space, or, more generally, belongs to an abstract set). If Σ is a subset of the possible values (or positions), and $U_{n,k}$ is 1 or 0 according as $X_{n,k} \in \Sigma$ or $X_{n,k} \notin \Sigma$, $(U_{n,1}, \cdots, U_{n,n})$ will not in general be a Markoff chain. But it is possible to extend the results of this paper to $\{U_{n,k}\}$: scalar products are replaced by matrix or operator products; but the formulas are similar and equally explicit. This will form the subject of a later communication.

2. **The generating function as a matrix product; elementary formulas.** The generating function $\phi_n(t)$, which is a product of the generating functions $t p_{n,k} + q_{n,k}$ of the individual trials in the independent case, expresses itself as a product of the corresponding conditional matrix generating functions

$$(2.1) \qquad W_{n,k} = \left\| \begin{array}{cc} t a_{n,k} & a'_{n,k} \\ t b_{n,k} & b'_{n,k} \end{array} \right\|$$

in the present one.

To see this, consider the segment $(U_{n,1}, \cdots, U_{n,k})$ $(1 \leq k \leq n)$ of the nth chain; and let $u'_{n,k,s}$ and $u''_{n,k,s}$ be the probabilities of its producing exactly s successes with a success in the kth trial, and with a failure in the kth trial, respectively. They satisfy the well known equations $(0 \leq s \leq k)$

$$u'_{n,k+1,s+1} = a_{n,k+1} u'_{n,k,s} + b_{n,k+1} u''_{n,k,s},$$
$$u''_{n,k+1,s} = a'_{n,k+1} u'_{n,k,s} + b'_{n,k+1} u''_{n,k,s},$$

and clearly $u'_{n,k,0} = u''_{n,k,k} = 0$. Introducing the generating functions

$$\phi'_{n,k}(t) = \sum_{s=0}^{k} u'_{n,k,s} t^s, \qquad \phi''_{n,k}(t) = \sum_{s=0}^{k} u''_{n,k,s} t^s,$$

we find at once that they satisfy a recurrence formula expressible in the matrix form

$$\left\| \phi'_{n,k+1}(t) \quad \phi''_{n,k+1}(t) \right\| = \left\| \phi'_{n,k}(t) \quad \phi''_{n,k}(t) \right\| W_{n,k+1},$$

and that $\phi'_{n,1}(t) = t p_{n,1}$, $\phi''_{n,1}(t) = q_{n,1}$. On solving this matrix difference equation and observing that $\phi_n(t) = \phi'_{n,n}(t) + \phi''_{n,n}(t)$, we obtain the desired

expression

$$(2.2) \qquad \phi_n(t) = \left\| t p_{n,1} \, q_{n,1} \right\| W_{n,2} W_{n,3} \cdots W_{n,n} J, \qquad J = \left\| \begin{matrix} 1 \\ 1 \end{matrix} \right\|.$$

Some further known results will be needed.

Let $\sigma_{n,k}$ be the standard deviation of $U_{n,k}$, and $\rho_{n,j,k}$ the correlation coefficient of $U_{n,j}$ with $U_{n,k}$ ($\rho_{n,k-1, \, k} = \rho_{n,k}$ being the one-step correlation). Then

$$\sigma_{n,j}\sigma_{n,k}\rho_{n,j,k} = E(U'_{n,j}U'_{n,k}) = E(U_{n,j}U_{n,k}) - p_{n,j}p_{n,k}$$
$$= p_{n,j}a_{n,j,k} - p_{n,j}p_{n,k},$$

where $U'_{n,k} = U_{n,k} - p_{n,k}$, and $a_{n,j,k} = p(U_{n,k}=1 \,|\, U_{n,j}=1)$. These $j \to k$ transition probabilities satisfy the equation (cf. (1.4))

$$(2.3) \qquad a_{n,j,k} = r_{n,k}a_{n,j,k-1} + b_{n,k} \qquad (1 \leqq j < k \leqq n).$$

Substracting (1.4) from (2.3), a recurrence formula for $a_{n,j,k} - p_{n,k}$ is obtained whose solution (with $a_{n,j,j} - p_{n,j} = q_{n,j}$), when substituted into the previous equation, gives

$$(2.4) \qquad \rho_{n,j,k} = r_{n,j+1}r_{n,j+2} \cdots r_{n,k}\sigma_{n,j}/\sigma_{n,k} = r_{n,j+1}^{(k-j)}\sigma_{n,j}/\sigma_{n,k}.$$

From this we draw the conclusion (since $\rho_{n,j,j}=1$) that

$$(2.5) \qquad \rho_{n,j,k} = \rho_{n,j+1}\rho_{n,j+2} \cdots \rho_{n,k} = \rho_{n,j+1}^{(k-j)};$$

and also, that *the coefficient of regression of $U_{n,k}$ on $U_{n,j}$* (being $\rho_{n,j,k}\sigma_{n,k}/\sigma_{n,j}$) *is* $r_{n,j+1}^{(k-j)}$.

Here, as henceforth, when $\{A_i\}$ is a sequence of symbols, each element of which is specified by the (last) subscript i, we write

$$A_i^{(0)} = 1 \text{ (or the identity)}, \qquad A_i^{(\lambda)} = A_i A_{i+1} \cdots A_{i+\lambda-1}.$$

Thus in terms of this notation (2.2) takes the form

$$\phi_n(t) = \left\| t p_{n,1} \quad q_{n,1} \right\| [W_{n,i}^{(n-1)}]_{i=2} J$$
$$= \left\| t p_{n,1} \quad q_{n,1} \right\| W_{n,2}^{(n-1)} J.$$

Equations (1.4) and (2.3) are of the form $y_v = r_v y_{v-1} + b_v$. The solution (found by standard methods of the calculus of finite differences, and verified by substitution) is

$$(2.6) \qquad y_v = y_1 r_2^{(v-1)} + \sum_{i=2}^{v} b_i r_{i+1}^{(v-i)} \qquad (2 \leqq v).$$

3. Asymptotic form of $W_{n,k}$. From the hypothesis (1.5) we conclude that

(3.1) $$\lim_{n\to\infty} \max_{2\leq k\leq n} b_{n,k} = 0.$$

This is seen immediately from (1.4) when $p_{n,1}\to 0$ so that $\lim \max p_{n,k}=0$. And since $b_{n,k}$ is independent of $p_{n,1}$, (3.1) holds in all cases.

Let c_ν be the first nonzero element in $\{c_i\}$. Then hypothesis (1.9) shows that for all $n\geq\nu$, $b_{n,2}= \cdots =b_{n,\nu-1}=0$. It also shows that *no* c_i is negative. And (3.1), (1.9) show that *either* the series $c_2+c_3+ \cdots$ is divergent and that $c_n/(c_2+ \cdots +c_n)\to 0$, *or else* that $\lim_{n\to\infty}\sum_{k=2}^n b_{n,k}=0$ (or both). In the former (the general) case we write

(3.2) $$\omega_1 = \cdots = \omega_\nu = c_\nu, \qquad \omega_n = c_2 + \cdots + c_n \qquad (n \geq \nu).$$

Then the sequence $\{\omega_i\}$ has the following three properties:

(3.3) $$0 < \omega_1 \leq \omega_2 \leq \cdots \lim_{n\to\infty} \omega_n = \infty, \qquad \lim_{n\to\infty} \frac{\omega_{n+1}}{\omega_n} = 1.$$

In other words, it is what we have recently called an *averaging* sequence[2][4]. In the special case noted above, the second property in (3.3) may be lacking.

Summing over k in (1.4), we obtain

$$\sum_{k=2}^n b_{n,k} = \sum_{k=1}^n p_{n,k} - \sum_{k=1}^{n-1} p_{n,k} r_{n,k+1} - p_{n,1}$$

$$= \sum_{k=1}^n p_{n,k} - \sum_{k=1}^{n-1} p_{n,k} a_{n,k+1} - p_{n,1} + \sum_{k=1}^{n-1} p_{n,k} b_{n,k+1}$$

$$= \sum_{k=1}^n E(U_{n,k}) - \sum_{k=1}^{n-1} E(U_{n,k}U_{n,k+1}) - p_{n,1} + \sum_{k=1}^{n-1} p_{n,k} b_{n,k+1}.$$

Now suppose that $p_{n,1}\to p_1$ as $n\to\infty$. The first two sums approach the limits $R_0=m$ and R_1, by the hypothesis (1.7). The last sum approaches zero, being between 0 and $(\max_{2\leq k\leq n} b_{n,k}) \sum E(U_{n,k})$ (cf. (3.1)). Therefore

$$\lim_{n\to\infty} \sum_{k=2}^n b_{n,k} = m - R_1 - p_1,$$

which is zero in the special case noted above, and is *never* negative.

Applying this to (1.9), (3.2), we obtain the asymptotic expression

(3.4)
$$b_{n,k} = \frac{h_k + \eta_{n,k}}{\omega_n}, \qquad h_k = (m - R_1 - p_1)c_k,$$

$$\lim_{n\to\infty} \frac{1}{\omega_n} \sum_{k=1}^n |\eta_{n,k}| = 0,$$

valid for all $n=1, 2, \cdots ; 2\leq k\leq n$. It will be convenient to introduce the

symbols $b_{n,1} = c_1 = 0$: (3.4) remains valid when $k = 1$ ($\eta_{n,1} = 0$). In the special case, $h_k = 0$.

In all cases, $h_k \geqq 0$.

Finally, we write $a_{n,1} = 1$ and define $W_{n,1}$ by equation (2.1). Evidently $\|tp_{n,1} \, q_{n,1}\| = \|p_{n,1} \, q_{n,1}\| W_{n,1}$. Consequently (2.2) becomes

$$\phi_n(t) = \|p_{n,1} \quad q_{n,1}\| W_{n,1}^{(n)} J,$$

which, in view of (3.4), may be written in the form

$$(3.5) \qquad \phi_n(t) = \|p_{n,1} \quad q_{n,1}\| \prod_{k=1}^{n} \left[B_{n,k} + \frac{G_k + C_{n,k}}{\omega_n} \right] J,$$

$$(3.6) \qquad B_{n,k} = \left\| \begin{matrix} ta_{n,k} & a'_{n,k} \\ 0 & 1 \end{matrix} \right\|, \qquad G_k = h_k \left\| \begin{matrix} 0 & 0 \\ t & -1 \end{matrix} \right\|,$$

$$C_{n,k} = \eta_{n,k} \left\| \begin{matrix} 0 & 0 \\ t & -1 \end{matrix} \right\|,$$

with $\{\omega_n\}$ satisfying (3.3). And in the special case

$$(3.7) \qquad \lim_{n \to \infty} \sum_{k=1}^{n} b_{n,k} = 0,$$

we have, instead of (3.3), (3.5), (3.6), the expression

$$(3.8) \qquad \phi_n(t) = \|p_{n,1} \quad q_{n,1}\| \prod_{k=1}^{n} [B_{n,k} + V_{n,k}] J,$$

$$(3.9) \qquad V_{n,k} = b_{n,k} \left\| \begin{matrix} 0 & 0 \\ t & -1 \end{matrix} \right\|,$$

with $B_{n,k}$ still given by (3.6).

4. The limit of the generating function. Let τ be a real number fixed once for all so that $\tau > 1$, $a\tau < 1$ (cf. (1.8)); and denote by T the circular region $|t| \leqq \tau$ in the plane of the complex variable t. If we can show that $\phi_n(t) \to \phi(t)$ uniformly on T as $n \to \infty$, it will follow at once from the expressions, as contour integrals around $|t| = \tau$, of the coefficients of t^s in the expansions of these functions, that $P_n(s) \to P(s)$ generated by $\phi(t)$. And by the same reasoning applied to the coefficients of $(t-1)^s$ in the expansions about $t = 1$, it will follow that the moments $\mu_{n,s} \to \mu_s$ corresponding to $P(s)$.

We introduce, with E. Hille[5], the following norm N for matrices

$$D = \|d_{ij}(t)\|, \qquad N(D) = \sup_{t \in T; \, i=1,2,\dots} \sum_j |d_{ij}(t)|.$$

---

[5] E. Hille, *Functional analysis and semi-groups*, Amer. Math. Soc. Colloquium Publications, vol. 31, §21.7.

In terms of this norm, the class of all square two-order matrices whose elements are functions of t, each one being defined and bounded on T, is a Banach algebra with identity[6]. Moreover, convergence with respect to N is equivalent to uniform convergence of the elements on T. The task that remains for us, therefore, is to find the limit with respect to N of the product of n square matrices in (3.5), (3.8)—on the assumption that $p_{n,1} \to p_1$.

According to the Exponential Limiting Product Theorem (Theorem VI of footnote 2), we shall have

$$(4.1) \qquad \lim_{n \to \infty} \prod_{k=1}^{n} \left[B_{n,k} + \frac{G_k + C_{n,k}}{\omega_n} \right] = B_1^{(\infty)} e^{M}$$

whenever the following conditions are all satisfied (in addition to the fact that $B_{n,k}$, G_k are elements of a Banach algebra with identity and that $\{\omega_i\}$ is an averaging sequence):

$$(4.2) \qquad N(B_{n,k}) \leqq \beta < \infty \qquad\qquad (n = 1, 2, \cdots ; 1 \leqq k \leqq n);$$

$$(4.3) \qquad B_{n,k}^{(\lambda)} = B_{n,k} B_{n,k+1} \cdots B_{n,k+\lambda-1} \to B_k^{(\infty)}$$

uniformly over k as λ and n become infinite independently, but with $k+\lambda-1 \leqq n$ (k remaining fixed);

$$(4.4) \qquad \frac{1}{\omega_n} \sum_{k=1}^{n} N(G_k) \leqq \gamma < \infty ;$$

$$(4.5) \qquad \lim_{n \to \infty} G_n/\omega_n = 0;$$

$$(4.6) \qquad \lim_{n \to \infty} \frac{1}{\omega_n} \sum_{k=1}^{n} G_k B_{k+1}^{(\infty)} = M ;$$

$$(4.7) \qquad \lim_{n \to \infty} \frac{1}{\omega_n} \sum_{k=1}^{n} N(C_{n,k}) = 0.$$

We shall verify the conditions (4.2)–(4.7) seriatim, on the assumption that $p_{n,1} \to p_1$ and that hypotheses (1.5)–(1.9) are made, using (3.3)–(3.6). The special case (3.7)–(3.9) will be treated by elementary methods in §7.

Calculating the norm of $B_{n,k}$ in (3.6), we have

$$N(B_{n,k}) = \sup_{t \in T} \max \left\{ 1, |t| a_{n,k} + a'_{n,k} \right\} = \max \left\{ 1, \tau a_{n,k} + a'_{n,k} \right\}$$

$$\leqq \max \left\{ 1, \tau + 1 \right\} = \tau + 1.$$

This establishes (4.2) with $\beta = \tau + 1$.

From (3.6) and (3.4) we obtain

---

[6] Cf. E. Hille, loc. cit., §1.14. For the definition of the exponential, cf. §5.7. For the purpose of this paper, the Banach algebra may be thought of simply as the algebra of 2-order square matrices with convergence defined by N.

$$N(G_k) = h_k N \left\| \begin{matrix} 0 & 0 \\ t & -1 \end{matrix} \right\| = h_k(\tau + 1) = (m - R_1 - p_1)(\tau + 1)c_k;$$

so that, in view of (3.2), condition (4.4) applies with $\gamma = (m - R_1 - p_1)(\tau + 1)$. And condition (4.5) is an evident consequence of this calculation and (3.3).

Again using (3.6), we find $N(C_{n,k}) = |\eta_{n,k}|(\tau + 1)$; then condition (4.7) is an immediate consequence of (3.4).

It remains to study conditions (4.3) and (4.6). This will be done in the following two sections.

5. Proof of condition (4.3). As a first step toward (4.3), we derive from (3.6) by mathematical induction

$$(5.1) \qquad B_{n,k}^{(\lambda)} = \left\| \begin{matrix} t^\lambda a_{n,k}^{(\lambda)} & 1 + (t-1)\alpha_{n,k,\lambda}(t) - t^{\lambda-1} a_{n,k}^{(\lambda)} \\ 0 & 1 \end{matrix} \right\|,$$

where $\alpha_{n,k,0}(t) = \alpha_{n,k,1}(t) = 0$; and for $\lambda > 1$,

$$(5.2) \qquad \alpha_{n,k,\lambda}(t) = \sum_{v=1}^{\lambda-1} t^{v-1} a_{n,k}^{(v)} .$$

We must examine the limiting behavior of the two elements in the first row of this matrix. On account of the hypothesis (1.8) and the relation $a\tau < 1$, we can chose a constant θ with $a\tau < \theta < 1$, and then have, for all sufficiently large n and all $2 \geq k \geq n$, the relation $|ta_{n,k}| < \theta$ (and we recall that $a_{n,1} = 1$). Then we shall have for $v \geq 1$ and $\lambda \geq 1$ and all k between 1 and $n - v + 1$ or 1 and $n - \lambda + 1$ inclusive:

$$(5.3) \qquad |t^{v-1} a_{n,k}^{(v)}| \leq \theta^{v-1}, \qquad |t^\lambda a_{n,k}^{(\lambda)}| \leq \tau\theta^{\lambda-1}.$$

We conclude at once that, as $n \to \infty$,

$$(5.4) \qquad t^\lambda a_{n,k}^{(\lambda)} \to 0, \qquad t^{\lambda-1} a_{n,k}^{(\lambda)} \to 0,$$

uniformly over $\{1 \leq k \leq n, t \in T\}$.

Relation (5.3) shows that, for each sufficiently great n, the series in (5.2) is majorized by the convergent geometric series $1 + \theta + \theta^2 + \cdots$ for each $t \in T$. The latter, in view of (1.8) and $a\tau < \theta$, also majorizes the following series for all $t \in T$ and every $k = 1, 2, \cdots$:

$$(5.5) \qquad \alpha_k(t) = \sum_{v=1}^{\infty} t^{v-1} a_k^{(v)} .$$

Therefore (5.5) is uniformly convergent on T for each fixed k; and since, by (1.8), each term of (5.5) is the limit as $n \to \infty$ of the corresponding term in (5.2), a familiar argument shows that, as n and λ both approach ∞ independently but always with $k + \lambda - 1 \leq n$ (k remaining fixed), $\alpha_{n,k}(t) \to \alpha_k(t)$

uniformly over $\{1 \leq k \leq n, t \in T\}$. This, in connection with (5.4), establishes the condition (4.3), and shows, moreover, that

(5.6)
$$B_k^{(\infty)} = \left\| \begin{array}{cc} 0 & 1 + (t-1)\alpha_k(t) \\ 0 & 1 \end{array} \right\|.$$

6. **Proof of condition** (4.6). From equations (5.6), (3.6), (3.4), (3.2), we readily obtain

(6.1)
$$\frac{1}{\omega_n} \sum_{k=1}^{n} G_k B_{k+1}^{(\infty)} = \left\| \begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array} \right\| \left[m - R_1 - p_1 + \frac{1}{\omega_n} \sum_{k=1}^{n} h_k \alpha_{k+1}(t) t \right] (t-1).$$

We have left the task of examining the summation on the right. We may write it as follows

(6.2)
$$\frac{1}{\omega_n} \sum_{k=1}^{n} h_k \alpha_{k+1}(t) t = \sum_{v=1}^{\infty} t^v \frac{1}{\omega_n} \sum_{k=1}^{n} h_k a_{k+1}^{(v)}.$$

Now for each coefficient of t^v we have for every n,

$$0 \leq \frac{1}{\omega_n} \sum_{k=1}^{n} h_k a_{k+1}^{(v)} \leq \frac{a^v}{\omega_n} \sum_{k=1}^{n} h_k = a^v (m - R_1 - p_1).$$

Hence, throughout T, the power series on the right in (6.2) is majorized by the convergent series $(m - R_1 - p_1)(\theta + \theta^2 + \theta^3 + \cdots)$. If we can show that each one of its coefficients approaches a limit, the series itself will approach a limit (uniformly on T); namely, the power series having these limiting coefficients as its respective coefficients: this will establish (4.6).

As a matter of fact, we will show that

(6.3)
$$\lim_{n \to \infty} \frac{1}{\omega_n} \sum_{k=1}^{n} h_k a_{k+1}^{(v)} = R_v - R_{v+1} - p_1 a_2^{(v)}.$$

Multiply (1.4) by $r_{n,k+1}^{(v)}$ and sum over k; we obtain

(6.4)
$$\sum_{k=2}^{n-v} b_{n,k} r_{n,k+1}^{(v)} = \sum_{k=1}^{n-v} p_{n,k} r_{n,k+1}^{(v)} - \sum_{k=1}^{n-v-1} p_{n,k} r_{n,k+1}^{(v+1)} - p_{n,1} r_{n,2}^{(v)}.$$

Now under our assumptions, $p_{n,1} r_{n,2}^{(v)} \to p_1 a_2^{(v)}$; consequently (6.3) will be established when we show, first, that for every v

(6.5)
$$\lim_{n \to \infty} \sum_{k=1}^{n-v} p_{n,k} r_{n,k+1}^{(v)} = R_v,$$

and, second, that

(6.6)
$$\lim_{n \to \infty} \frac{1}{\omega_n} \sum_{k=1}^{n} h_k a_{k+1}^{(v)} = \lim_{n \to \infty} \sum_{k=2}^{n-v} b_{n,k} r_{n,k+1}^{(v)}.$$

Applying the solution (2.6) to the equation (2.3) in which j and k are replaced by k and $k+v$, respectively, and $y_v = a_{n,k,k+v}$, $r_v = r_{n,k+v}$, $b_v = b_{n,k+v}$, we obtain (using (2.3) again, and the fact that $a_{n,k,k} = 1$):

$$a_{n,k,k+v} = \overset{(v)}{r}_{n,k+1} + \sum_{i=1}^{v} b_{n,k+i} \overset{(v-i)}{r}_{n,k+i+1}.$$

For the θ of §5, $a < \theta < 1$; therefore, in view of (1.8) and (3.1), n can be taken so large that for all $k \leq n$,

$$\left| a_{n,k} - a_k \right| < (\theta - a)/2, \qquad \left| b_{n,k} \right| < (\theta - a)/2,$$

and consequently $\left| r_{n,k} \right| < \theta$. For all such n,

$$\left| \sum_{i=1}^{v} b_{n,k+i} \overset{(v-i)}{r}_{n,k+i+1} \right| < \max_{1 \leq k \leq n} b_{n,k} \cdot (1 - \theta)^{-1} = \delta_n.$$

Therefore we may write

$$(6.7) \qquad a_{n,k,k+v} = \overset{(v)}{r}_{n,k+1} + \delta_{n,k,v}, \qquad \left| \delta_{n,k,v} \right| < \delta_n, \qquad \lim_{n \to \infty} \delta_n = 0.$$

Now we have

$$\sum_{k=1}^{n-v} E(U_{n,k} U_{n,k+v}) = \sum_{k=1}^{n-v} p_{n,k} a_{n,k,k+v}$$

$$= \sum_{k=1}^{n-v} p_{n,k} \overset{(v)}{r}_{n,k+1} + \sum_{k=1}^{n-v} p_{n,k} \delta_{n,k,v};$$

$$\left| \sum_{k=1}^{n-v} p_{n,k} \delta_{n,k,v} \right| \leq \sum_{k=1}^{n} p_{n,k} \delta_n.$$

If we let $n \to \infty$ and apply (6.7) and (1.7), (6.5) is established.

To prove equation (6.6), we set $r_{n,k} = a_k + (a_{n,k} - a_k - b_{n,k}) = a_k + z_{n,k}$; then, in virtue of (1.8) and (3.1), $z_{n,k} \to 0$ uniformly over k as $n \to \infty$. We have then

$$(6.8) \qquad \overset{(v)}{r}_{n,k} = (a_k + z_{n,k})^{(v)} = \overset{(v)}{a}_k + \Delta_{n,k,v},$$

where $\Delta_{n,k,v}$ is a sum of $2^v - 1$ products of v factors chosen from the quantities a_i, $z_{n,i}$ ($k \leq i \leq k+v-1$), each containing at least one factor $z_{n,i}$. Since every such product is less in absolute value than $\max_{1 \leq k \leq n} z_{n,k}$, we see that $\Delta_{n,k,v} \to 0$ uniformly over k as $n \to \infty$.

Combining (6.8) with (3.4), we have

$$\sum_{k=2}^{n-v} b_{n,k} \overset{(v)}{r}_{n,k+1} = \sum_{k=2}^{n-v} b_{n,k} \overset{(v)}{a}_{k+1} + \sum_{k=2}^{n-v} b_{n,k} \Delta_{n,k+1,v},$$

$$\left| \sum_{k=2}^{n-v} b_{n,k} \Delta_{n,k+1,v} \right| \leqq \left(\max_{k} \Delta_{n,k+1,v} \right) \sum_{k=2}^{n-v} b_{n,k} \to 0;$$

$$\sum_{k=2}^{n-v} b_{n,k} a_{k+1}^{(v)} = \frac{1}{\omega_n} \sum_{v=2}^{n-v} h_k a_{k+1}^{(v)} + \frac{1}{\omega_n} \sum_{v=2}^{n-v} \eta_{n,k} a_{k+1}^{(v)},$$

$$\left| \frac{1}{\omega_n} \sum_{v=2}^{n-v} \eta_{n,k} a_{k+1}^{(v)} \right| \leqq \frac{a^v}{\omega_n} \sum_{k=1}^{n-v} | \eta_{n,k} | \to 0.$$

These relations establish (6.6); and therewith, prove (6.3).

Combining (6.1), (6.2), (6.3), we obtain, uniformly on T,

$$M = \lim_{n \to \infty} \sum_{k=1}^{n} G_k B_{k+1}^{(\infty)},$$

where

$$(6.9) \quad M = \left\| \begin{matrix} 0 & 0 \\ 0 & 1 \end{matrix} \right\| (t-1) \left[m + (t-1) \sum_{v=1}^{\infty} R_v t^{v-1} - \sum_{v=0}^{\infty} p_1 a_2^{(v)} t^v \right].$$

7. **The special case.** In the special case in which (3.7), (3.8), (3.9) replace (3.3)–(3.6), the derivation of the limit (uniform over T)

$$(7.1) \qquad \lim_{n \to \infty} \prod_{k=1}^{n} [B_{n,k} + V_{n,k}] = B_1^{(\infty)} e^{M},$$

which replaces (4.1), and of the equations (5.5), (5.6), (6.9), is done by elementary methods, not requiring the Exponential Limiting Product Theorem: one simply computes each side of (7.1).

To compute the right-hand side of (7.1), the first step is to show that M, *defined* as the expression on the right in (6.9), is zero. The results (6.4)–(6.5) being valid, we have (when $p_{n,1} \to p_1$)

$$\lim_{n \to \infty} \sum_{k=2}^{n-v} b_{n,k} r_{n,k+1}^{(v)} = R_v - R_{v+1} - p_1 a_2^{(v)}.$$

And since $\left| r_{n,k+1}^{(v)} \right| \leqq 1$, assumption (3.7) shows that the left member vanishes. Hence, for all $v = 0, 1, 2, \cdots$,

$$(7.2) \qquad R_v - R_{v+1} - p_1 a_2^{(v)} = 0.$$

Obvious computation in (6.9) shows that $M = 0$. Therefore the right-hand side of (7.1) is $B_1^{(\infty)}$ which we now *define* by equations (5.5), (5.6).

To compute the left-hand side of (7.1), direct multiplication (preserving order) shows that

$$\prod_{k=1}^{n} [B_{n,k} + V_{n,k}] = B_1^{(n)} + \Omega_n,$$

where Ω_n is a sum of products each containing at least one factor $V_{n,k}$. Now the considerations of §5 apply equally well to the present case; and we have, uniformly over T, $\lim_{n\to\infty} B_1^{(n)} = B_1^{(\infty)}$, given by (5.5), (5.6). Accordingly, the desired (7.1), and so forth, will be established once it is shown that $\lim_{n\to\infty}\Omega_n = 0$ uniformly over T.

Classifying the terms of Ω_n according to the numbers of $V_{n,k}$ factors, we write $\Omega_n = \Gamma_{n,1} + \Gamma_{n,2} + \cdots + \Gamma_{n,n}$ where $\Gamma_{n,m}$ $(m=1, \cdots, n)$ is the sum of all C_m^n products

$$* V_{n,k_1} * V_{n,k_2} * \cdots * V_{n,k_m} *,$$

each $*$ denoting a product $B_{n,k}^{(\lambda)}$ (including $B_{n,k}^{(0)}=I$). By §5 (particularly (5.1)–(5.3)), $N(B_{n,k}^{(\lambda)}) \leqq B = \max\{1, a/(1-\theta)\}$; and using (3.9), clearly $N(V_{n,k}) \leqq (\tau+1)b_{n,k}$. Furthermore, $N(\Gamma_{n,m})$ does not exceed the sum of the norms of its C_m^n terms, which latter norms do not exceed the product of the norms of their factors. Consequently

$$N(\Gamma_{n,m}) \leqq B[B(\tau + 1)]^m \sum_{(k)} b_{n,k_1} b_{n,k_2} \cdots b_{n,k_m},$$

the summation being extended over all $1 \leqq k_1 < k_2 < \cdots < k_m \leqq n$. Evidently, therefore,

$$N(\Gamma_{n,m}) \leqq B\left[B(\tau + 1)\sum_{k=1}^{n} b_{n,k}\right]^m = B(u_n)^m$$

so that, by the geometric formula,

$$N(\Omega_n) \leqq Bu_n \frac{1 - (u_n)^n}{1 - u_n}.$$

From this and (3.7) the required uniform $\Omega_n\to 0$ results.

8. The law of small numbers. The results of the preceding sections, and particularly formulas (3.5), (4.1), (5.5), (5.6), (6.9), furnish the proof of the following theorem, which embodies the law of small numbers under the conditions of this paper. The same theorem is established by the results of §7 in the special case.

THEOREM. *Any two-outcome one-step Markoff chain, for which hypotheses (1.5)–(1.9) together with $p_{n,1}\to p_1$ are assumed, will lead as $n\to\infty$ to the limiting distribution $P_n(s)\to P(s)$ generated by the function*

(8.1) $$\phi(t) = [1 - f(t)] \exp\{-(1 - t)[m - (1 - t)\psi(t)] + f(t)\},$$

(8.2) $$\psi(t) = \sum_{v=1}^{\infty} R_v t^{v-1}, \qquad f(t) = (1 - t)p_1 \sum_{v=0}^{\infty} a_2^{(v)} t^v.$$

The power series in (8.2) are convergent on T: $|t| < \tau > 1$, and the convergence $\phi_n(t)\to\phi(t)$ is uniform on T.

In the calculations leading to these formulas, we must recall that $a_1 = 1$. And the series for $\psi(t)$ can be majorized directly with the use of (6.5) and (6.8).

In view of (6.5), R_v can be described as the v-step auto-regression, averaged over the trials (weighted according to their absolute probabilities of success), as $n \to \infty$: $\psi(t)$ is a generating function of such quantities. And since $r_{n,2}^0 \to a_2^{(v)}$, $f(x)$ is $(1-t)$ times the generating function of the *initial* auto-regressions, weighted according to the initial probability p_1, for $n \to \infty$.

A notable simplification of the result occurs in the important case of vanishing initial probability, $p_1 = 0$: then the distribution is determined by the sequence $\{R_0, R_1, R_2, \cdots\}$ alone; that is, by $\psi(t)$.

The most important special case is that of stationary limiting transitions; or, equivalently, of fixed limiting auto-regression:

$$(8.3) \qquad\qquad a_2 = a_3 = \cdots = \lim_{n \to \infty} r_{n,k} = r.$$

The equations (6.5) and (6.8) show that

$$R_v = \lim_{n \to \infty} \sum_{k=1}^{n-v} p_{n,k} a_k^{(v)} + \lim_{n \to \infty} \sum_{k=1}^{n-v} p_{n,k} \Delta_{n,k,v}.$$

The second limit is zero since $\Delta_{n,k,v} \to 0$ uniformly, and the sum of the non-negative $p_{n,k}$ is bounded. The first limit can be written

$$\lim_{n \to \infty} \sum_{k=1}^{n-v} p_{n,k} a_p^{(v)} = r^v \lim_{n \to \infty} \sum_{k=1}^{n-v} p_{n,k} = r^v m,$$

since each $p_{n,n-j} \to 0$ when i is fixed, as is seen on applying the solution (2.6) to (1.4). Therefore $R_v = mr^v$. Thus we obtain from (8.3), (8.2),

$$\psi(t) = \frac{mr}{1 - rt}, \qquad f(t) = p_1 \frac{1 - t}{1 - rt};$$

and (8.1) reduces to the form (obtained previously[1])

$$(8.4) \quad \phi(t) = \left[1 - p_1 \frac{1-t}{1-rt}\right] \exp\left\{-(1-t)[m(1-r) - p_1]/(1-rt)\right\}.$$

Explicit expressions for $P(s)$ are easily found in this case in terms of the Laguerre polynomials, defined by the equations
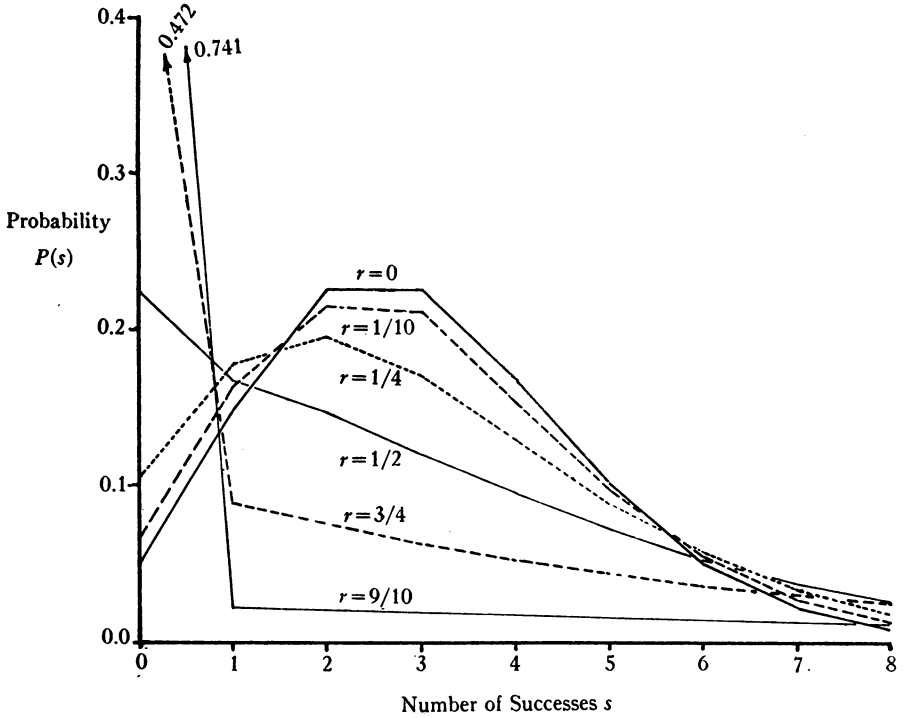
$$\frac{1}{1-x} \exp \frac{-wx}{1-x} = \sum_{s=0}^{\infty} L_s(w) x^s.$$

On setting $x = rt$, $w = -[m(1-r) - p_1](1-r)/r$, we obtain

$$(8.5) \quad P(s) = \exp\left\{-m(1-r) + p_1\right\}[q_1 L_s(w) r^s + (p_1 - r) L_{s-1}(w) r^{s-1}].$$

The simplest and most interesting case of all is that in which $p_1 = 0$. Then we have

(8.6)
$$\phi(t) = \exp -\frac{m(1-r)(1-t)}{1-rt},$$

$$P(s) = \exp\{-m(1-r)\}r^s[L_s(w) - L_{s-1}(w)].$$

THE EFFECT OF DEPENDENCE $(m=3)$

The accompanying figure illustrates the effect of the dependence of the trials in this case when $m=3$: The values of $P(s)$ given by (8.6) are plotted for $s=0, 1, 2, \cdots, 8$, and are then connected by straight line segments to aid the eye. This is done for $r=0$ (the Poisson case), $r=1/10$, $r=1/4$, $r=1/2$, $r=3/4$, $r=9/10$. Naturally, no negative value of r must be considered: a law of small numbers and a negative limiting correlation are inconsistent in principle (for (1.4), (3.1) lead to $\lim \inf_{n\to\infty} r_{n,k} = \lim \inf_{n\to\infty} a_{n,k} \geqq 0$).

COLUMBIA UNIVERSITY,
    NEW YORK, N. Y.