Z. Wahrscheinlichkeitstheorie verw. Geb. 24, 159–166 (1972)

On Finite Tail σ-Algebras*

Marius Iosifescu

1. Introduction

Let $\xi=(\xi_t)_{t\in\Theta}$ be a stochastic process on a probability space $(\Omega,\mathscr{K},P)$ assuming values in an arbitrary measurable space $(X,\mathscr{X})$. Here Θ is either $N=\{0,1,2,\ldots\}$ or the non-negative half line $[0,\infty)$.

Let $\mathscr{K}_s^t$, $s\leqq t$, denote the σ-algebra generated by the family $(\xi_u)_{s\leqq u\leqq t}$, $\mathscr{K}_s=\mathscr{K}_s^s$, and $\mathscr{K}_s^\infty=$ the least σ-algebra containing all the σ-algebras $\mathscr{K}_s^t$, $t\geqq s$.

The σ-algebra $\mathscr{T}=\bigcap_{s\in\Theta}\mathscr{K}_s^\infty$ is called the *tail* σ-algebra of ξ.

A σ-algebra contained in $\mathscr{K}$ will be said to be δ-*trivial* (under P), $0\leqq\delta<1$, if the probability of any set belonging to it is either 0 or greater than or equal to $1-\delta$. Clearly, a δ-trivial σ-algebra is equivalent to a finite σ-algebra containing at most $[(1-\delta)^{-1}]$ atoms, and, conversely, any finite σ-algebra is δ-trivial for some $0\leqq\delta<1$. A 0-trivial σ-algebra will be simply called *trivial.* Notice that δ-trivialness with $\delta<\frac{1}{2}$ amounts to trivialness because $0<P(A)<1$ would imply that both $P(A)$ and $P(A^c)$ are greater than $\frac{1}{2}$.

It is well known that if ξ is a sequence of independent random variables, then its tail σ-algebra is trivial (Kolmogorov's 0–1 law). Blackwell and Freedman [2] have proved that the same conclusion holds for homogeneous irreducible, aperiodic, recurrent, countable state, discrete parameter Markov chains. Recently, Cohn [3] has proved that the tail σ-algebra of any finite state, discrete parameter Markov chain *(homogeneous or not)* is δ-trivial.

In this paper we characterize (Theorems 1 and 2) stochastic processes having a δ-trivial tail σ-algebra, then we exhibit (Theorem 3) the simplification arising in the Markovian case.

Several applications to discrete state space Markov processes are made. Theorem 4 gives necessary and sufficient conditions for the δ-trivialness of the tail σ-algebra in terms of the transition and absolute probabilities. Theorem 5 is a necessary and sufficient condition for the tail σ-algebra to be trivial under any initial distribution. Theorem 7 exhibits a class of Markov processes for which the strongly mixing property and the trivialness of the tail σ-algebra are equivalent. Finally, Theorems 6 and 8 extend the above quoted results in [2] and [3] to the continuous parameter case.

2. The General Case

The theorems below slightly extend some results by Bártfai and Révész [1] and Blackwell and Freedman [2].

* This paper was written while the author was an Overseas Fellow of Churchill College, Cambridge, Great Britain.

Theorem 1. *Assume that*

$$\lim_{t\to\infty} \sup_{B\in\mathscr{K}_t^\infty} \big(P(A\cap B)-P(A)\,P(B)\big) \leqq \delta P(A)$$

for any $A\in\bigcup_{s\in\Theta}\mathscr{K}_0^s$. *Then* $\mathscr{T}$ *is* δ*-trivial.*

Proof. Let us first notice that the above limit does exist as the quantity under 'lim' does not increase when t increases. Moreover, it is non-negative because

$$P(A\cap B)-P(A)\,P(B) = -\big(P(A\cap B^c)-P(A)\,P(B^c)\big).$$

Passing to the proof of the theorem, it is easily seen that the assumption made implies that for any $T\in\mathscr{T}$ and $s\in\Theta$ we have

$$P(T\,|\,\mathscr{K}_0^s)-P(T)\leqq\delta \quad \text{a.s.}$$

Now, by a well-known property (see e.g. [6], p. 409) $P(T\,|\,\mathscr{K}_0^{s_n}) \to P(T\,|\,\mathscr{K}_0^\infty)$ a.s. for any increasing sequence $(s_n)_{n\in N}$ such that $\lim_{n\to\infty} s_n=\infty$. Therefore

$$P(T\,|\,\mathscr{K}_0^\infty)-P(T)\leqq\delta \quad \text{a.s.}$$

Taking into account that $T\in\mathscr{T}\subset\mathscr{K}_0^\infty$ and integrating the above inequality over T we obtain

$$P(T)-P^2(T)\leqq\delta P(T),$$

whence either $P(T)=0$ or $P(T)\geqq 1-\delta$, q.e.d.

The proof of the next theorem is based on the following simple

Proposition. *Let* $E_j\in\mathscr{K}$ *be such that* $P(E_j)>0$, $1\leqq j\leqq r$. *Then*

$$P(A\,|\,E_j)\leqq\frac{P(A)}{\min\limits_{1\leqq k\leqq r} P(E_k)}$$

for any $A\in\mathscr{K}$ *and* $1\leqq j\leqq r$.

Proof.[1] For any $1\leqq j\leqq r$ we can write

$$P(A)\geqq P(A\cap E_j)=P(A\,|\,E_j)\,P(E_j)\geqq P(A\,|\,E_j)\min_{1\leqq k\leqq r} P(E_k). \qquad \text{q.e.d.}$$

Theorem 2. *Assume that* $\mathscr{T}$ *is* δ*-trivial. Then*

$$\lim_{t\to\infty} \sup_{B\in\mathscr{K}_t^\infty} \big(P(A\cap B)-P(A)\,P(B)\big) \leqq \delta P(A)$$

for any $A\in\mathscr{K}$.

Proof. If $\mathscr{T}$ is δ-trivial the above Proposition implies that for any $A\in\mathscr{K}$

$$P(A\,|\,\mathscr{T})\leqq(1-\delta)^{-1}P(A) \quad \text{a.s.},$$

whence

$$P(A\,|\,\mathscr{T})-P(A)\leqq\delta P(A\,|\,\mathscr{T}) \quad \text{a.s.} \tag{1}$$

[1] I am indebted to Dr. D. Williams for this simple proof.

Then, for $A \in \mathscr{K}$ and $B \in \mathscr{K}_t^\infty$ we can write

$$P(A \cap B) - P(A)\,P(B)$$
$$= \int_B [P(A|\mathscr{K}_t^\infty)(\omega) - P(A|\mathscr{T})(\omega)]\,P(d\omega) + \int_B [P(A|\mathscr{T})(\omega) - P(A)]\,P(d\omega).$$

The first integral above is dominated by

$$\int_\Omega |P(A|\mathscr{K}_t^\infty)(\omega) - P(A|\mathscr{T})(\omega)|\,P(d\omega)$$

which by [6] loc. cit. tends to 0 as $t \to \infty$. By virtue of (1) the second integral is dominated by $\delta P(A)$, q.e.d.

3. The Markovian Case

3.1. When ξ is a Markov process a simplification occurs in Theorems 1 and 2. We have namely

Theorem 3. *Assume that ξ is a Markov process. Then in Theorems* 1 *and* 2 $\sup\limits_{B \in \mathscr{K}_t^\infty}$ *can be replaced by* $\sup\limits_{B \in \mathscr{K}_t}$.

Proof. We shall prove that

$$f(A) = \sup_{B \in \mathscr{K}_t^\infty} \big(P(A \cap B) - P(A)\,P(B)\big)$$
$$= \sup_{B \in \mathscr{K}_t} \big(P(A \cap B) - P(A)\,P(B)\big) = g(A)$$

for any $A \in \mathscr{K}_0^s$ with $s < t$.

Clearly, $f(A) \geqq g(A)$. To prove the converse inequality let us notice that by the Markov property we can write

$$P(A \cap B) - P(A)\,P(B) = P(A)\Big[\int_\Omega P(B|\mathscr{K}_t)(\omega)\,P_{t,A}(d\omega) - \int_\Omega P(B|\mathscr{K}_t)(\omega)\,P_t(d\omega)\Big]$$
$$= P(A) \int_\Omega P(B|\mathscr{K}_t)(\omega)(P_{t,A} - P_t)(d\omega),$$

for any $A \in \mathscr{K}_0^s$ such that $P(A) \neq 0$, and any $B \in \mathscr{K}_t^\infty$, $t > s$, where P_t is the restriction of P at $\mathscr{K}_t$ (i.e. $P_t(E) = P(E)$ for any $E \in \mathscr{K}_t$) and $P_{t,A}(.) = P_t(.|A)$. The inequality $f(A) \leqq g(A)$ is an immediate consequence of the above equation on account of the following

Lemma. *Let $(\Omega, \mathscr{L}, \lambda)$ be a measure space and assume that the signed measure λ is finite and σ-additive and such that $\lambda(\Omega) = 0$. If h is any bounded real valued and $\mathscr{L}$-measurable function defined on Ω, then*

$$\int_\Omega h(\omega)\,\lambda(d\omega) \leqq (\operatorname{ess\,sup} h - \operatorname{ess\,inf} h) \sup_{L \in \mathscr{L}} \lambda(L).$$

The proof of this Lemma can be found in [5], p. 40, q.e.d.

3.2. Theorem 3 allows us to give a necessary and sufficient condition for δ-trivialness of the tail σ-algebra of a Markov process in terms of its transition probabilities. For the sake of simplicity assume that the state space X is at most a countable set with elements $i, j, \dots$ Put $p_i(t) = P(\xi(t) = i)$, $i \in X$, and

$$X(t) = \{i: p_i(t) > 0\}, \qquad t \in \Theta.$$

Denote by $p_{ij}(s, s+u)$, $i, j \in X$, $s, u \in \Theta$ the transition probabilities of the process, that is

$$p_{ij}(s, s+u) = P(\xi(s+u) = j \mid \xi(s) = i).$$

On account of the well known fact that

$$\sup_{E \subset X} \sum_{i \in E} (u_i - v_i) = \tfrac{1}{2} \sum_{i \in X} |u_i - v_i|$$

for any two probability distributions $(u_i)_{i \in X}$ and $(v_i)_{i \in X}$ on X, Theorems 1, 2 and 3 yield

Theorem 4. *Assume that ξ is a Markov process with at most a countable set of states. Then*

1. *A necessary and sufficient condition for its tail σ-algebra to be δ-trivial ($\delta \geqq \frac{1}{2}$) is that*

$$\lim_{u \to \infty} \sum_{j \in X} |p_{ij}(s, s+u) - p_j(s+u)| \leqq 2\delta \tag{2}$$

for any $s \in \Theta$ and $i \in X(s)$.

2. *A necessary and sufficient condition for its tail σ-algebra to be trivial is that*

$$\lim_{u \to \infty} \sum_{j \in X} |p_{ij}(s, s+u) - p_j(s+u)| = 0 \tag{3}$$

for any $s \in \Theta$ and $i \in X(s)$.

(For any fixed $s \in \Theta$ and $i \in X(s)$ the quantity under 'lim' *is a nonincreasing function of u.)*

Notice that (2) with $\delta < \frac{1}{2}$ must imply (3) (on account of Theorems 1, 2, and 3, too). I do not know a direct proof of this fact.

3.3. Until now the initial distribution of our Markov process, $p_i(0) = P(\xi(0) = i)$, $i \in X$, has been kept fixed. Theorem 4 allows us to characterize Markov processes for which the tail σ-algebra is trivial under any initial distribution (the transition probabilities remaining the same). To this end put $\overline{X}(0) = X$ and $\overline{X}(t) = \{j: p_{ij}(0, t) > 0$ for some $i \in X\}$ for $t \neq 0$.

Theorem 5. *Assume that ξ is a Markov process with at most a countable set of states. Then a necessary and sufficient condition for its tail σ-algebra to be trivial under any initial distribution is that*

$$\lim_{u \to \infty} \sum_{j \in X} |p_{ij}(s, s+u) - p_{kj}(s, s+u)| = 0 \tag{4}$$

for any $s \in \Theta$ and $i, k \in \overline{X}(s)$.

Proof. Suppose first that the tail σ-algebra of ξ is trivial under any initial distribution. Taking as initial distribution that which assigns to each of the states i and k probability $\frac{1}{2}$, on account of (3) we obtain that

$$\lim_{u\to\infty} \sum_{j\in X} |p_{ij}(0,u)-p_{kj}(0,u)|=0 \tag{5}$$

for any $i, k\in X$. Further, for any $s\neq 0$ and $i\in \bar{X}(s)$ for which $p_{li}(0,s)>0$, Eq. (3) implies that

$$\lim_{u\to\infty} \sum_{j\in X} |p_{ij}(s,s+u)-p_{lj}(0,s+u)|=0. \tag{6}$$

Now it is easily seen that (5) and (6) imply (4).

Conversely, suppose that (4) holds and consider any initial distribution. As

$$p_{ij}(s,s+u)-p_j(s+u)=\sum_{l\in\bar{X}(s)} p_l(s)\,[p_{ij}(s,s+u)-p_{lj}(s,s+u)]$$

for any $s\in\Theta$, $i,j\in X$, by dominated convergence and on account of (4) we have

$$\lim_{u\to\infty} \sum_{j\in X} |p_{ij}(s,s+u)-p_j(s+u)|$$

$$\leqq \sum_{l\in\bar{X}(s)} p_l(s) \lim_{u\to\infty} \sum_{j\in X} |p_{ij}(s,s+u)-p_{lj}(s,s+u)|=0$$

for any $s\in\Theta$ and $i\in\bar{X}(s)\supset X(s)$.

But this means that (3) holds for the initial distribution we considered, q.e.d.

As an application of Theorem 5 we shall prove

Theorem 6. *Assume that ξ is a homogeneous continuous parameter Markov process with at most a countable set of states (i.e. $p_{ij}(s,s+u)=p_{ij}(u)$ for any $i,j\in X$, $s,u\in\Theta$) that form a recurrent class. Then the tail σ-algebra of ξ is trivial under any initial distribution.*

Proof. By Theorem 5 it is sufficient to prove that

$$\lim_{u\to\infty} \sum_{j\in X} |p_{ij}(u)-p_{kj}(u)|=0 \tag{6}$$

for any $i,k\in X$. But this is an immediate consequence of a theorem by Orey [7] establishing (6) for irreducible, recurrent, aperiodic, discrete parameter Markov chains, so that $\sum_{j\in X} |p_{ij}(nh)-p_{kj}(nh)|\to 0$ as $n\to\infty$ (through integral values) for any $h>0$, and of the fact that the quantity under 'lim' in (6) is a nonincreasing function of u, q.e.d.

3.4. We close the study of the Markovian case with a remark about strongly mixing Markov processes.

An arbitrary stochastic process ξ is said to be *strongly mixing* if

$$\alpha(\tau)=\sup_{s\in\Theta} \sup_{A\in\mathscr{X}_0^s,\, B\in\mathscr{X}_{s+\tau}^\infty} (P(A\cap B)-P(A)\,P(B))$$

tends to zero as $\tau\to\infty$. On account of Theorem 1, the strong mixing property implies the trivialness of the tail σ-algebra. The theorem below exhibits a case for which the converse implication is also true.

Theorem 7.[2] *Assume that ξ is a stationary homogeneous Markov process with at most a countable set of states (i.e. $p_i(s)=\pi_i$, $p_{ij}(s,s+u)=p_{ij}(u)$ for any $i,j\in X$, $s,u\in\Theta$). Then the strong mixing property is equivalent to the trivialness of the tail σ-algebra. Moreover,*

$$\alpha(\tau)\leqq\tfrac{1}{2}\sum_{i\in X}\pi_i\sum_{j\in X}|p_{ij}(\tau)-\pi_j|.$$

Proof. For any $A\in\mathscr{K}_0^s$, $B\in\mathscr{K}_{s+\tau}$ we have

$$P(A\cap B)-P(A)\,P(B)\leqq\sum_{i\in X}\pi_i\,\Big|\sum_{j\in B}(p_{ij}(\tau)-\pi_j)\Big|$$

and the proof follows on account of Theorems 3 and 4, q.e.d.

The above theorem (together with Theorem 6) extends a result by Davydov [4].

4. An Application to Finite State Markov Processes

By making use of Theorem 1 we shall now deduce a result first proved by Cohn [3] for the discrete parameter case.

Theorem 8. *Assume that $\xi=(\xi_t)_{t\in\Theta}$ is a finite state Markov process. Then its tail σ-algebra is δ-trivial for some $0\leqq\delta<1$.*

Proof. Let $E_i(t)=\{\xi_t=i\}$ and put as above $p_i(t)=P(E_i(t))$, $i\in X$, and $X(t)=\{i\colon p_i(t)>0\}$.

We shall prove that *there is an increasing sequence $(t_n)_{n\in N}\subset\Theta$ with $\lim_{n\to\infty}t_n=\infty$ and there are two disjoint sets Y and Z, $Y\neq\varnothing$, $Y\cup Z=X$ such that*

$$\liminf_{n\to\infty}\min_{i\in X(t_n)\cap Y}p_i(t_n)=\gamma>0$$

and $\lim_{n\to\infty}\sum_{i\in Z}p_i(t_n)=0$.[3]

Admitting that this is true, let us show that δ-trivialness follows with $\delta=1-\gamma$. For given $0<\varepsilon<\gamma$ and $A\in\bigcup_{t\in\Theta}\mathscr{K}_0^t$ choose an n large enough in order that $A\in\mathscr{K}_0^{t_n}$, $\sum_{i\in Z}p_i(t_n)\leqq\varepsilon$, and $\min_{i\in X(t_n)\cap Y}p_i(t_n)\geqq\gamma-\varepsilon$. Notice that our Proposition implies that whatever $B\in\mathscr{K}$ we have

$$P(B)\geqq(\gamma-\varepsilon)\,P(B\,|\,\mathscr{K}_{t_n})(\omega) \tag{7}$$

for almost all ω not belonging to the event $E'(t_n)=\bigcup_{i\in Z}E_i(t_n)$, the probability of which is $\leqq\varepsilon$. By the Markov property

$$P(A\cap B)-P(A)\,P(B)=\int_{A\cap E'(t_n)}+\int_{A\cap(E'(t_n))^c}\big(P(B\,|\,\mathscr{K}_{t_n})(\omega)-P(B)\big)\,P(d\omega)$$

for any $B\in\mathscr{K}_s^\infty$ with $s>t_n$. The first integral above is dominated by ε and the second one, because of the estimate (7) for $P(B)$, is dominated by $(1-\gamma+\varepsilon)\,P(A)$, so that

$$P(A\cap B)-P(A)\,P(B)\leqq\varepsilon+(1-\gamma+\varepsilon)\,P(A)$$

---

[2] Compare with Lemma 2 in [8] p. 207.

[3] This is one of the main points in Cohn's proof, too.

for any $0<\varepsilon<\gamma$ and any $A\in\bigcup_{t\in\Theta}\mathscr{K}_0^t$ provided that $B\in\mathscr{K}_s^\infty$ with $s>t_n$, where t_n depends upon both A and ε. But δ-trivialness then follows at once from Theorem 1, with $\delta=1-\gamma$, because of the $\lim_{t\to\infty}$ which occurs in the sufficient condition given there.

Coming back to the proof of the above statement in italics, let

$$\alpha=\liminf_{t\to\infty}\min_{i\in X(t)}p_i(t).$$

If $\alpha>0$, clearly, we can take $Y=X$, $Z=\varnothing$, $\gamma=\alpha$ and as $(t_n)_{n\in N}$ any sequence through which the above 'lim inf' is reached. If $\alpha=0$, let $(a_n)_{n\in N}\subset\Theta$ be a sequence such that

$$\lim_{n\to\infty}\min_{i\in X(a_n)}p_i(a_n)=0$$

and let i_n be a state such that

$$\min_{i\in X(a_n)}p_i(a_n)=p_{i_n}(a_n),\qquad n\in N.$$

In the sequence $(i_n)_{n\in N}$ at least one state, say 1, appears infinitely often. Therefore, for a subsequence $(b_n)_{n\in N}\subset(a_n)_{n\in N}$ we have $\lim_{n\to\infty}p_1(b_n)=0$, so that state 1 will be the first element of Z. Next let

$$\beta=\liminf_{n\to\infty}\min_{i\in X(b_n)-\{1\}}p_i(b_n).$$

If $\beta>0$, we take $Y=X-\{1\}$, $Z=\{1\}$, $\gamma=\beta$, and $t_n=b_n$, $n\in N$. If $\beta=0$, we repeat the above procedure to include a new state, say 2, in Z, that is, we find a subsequence $(c_n)_{n\in N}\subset(b_n)_{n\in N}$ such that $\lim_{n\to\infty}p_2(c_n)=0$. We must stop before including all the states in Z as $\sum_{i\in X}p_i(t)=1$ for all $t\in\Theta$, q.e.d.

Corollary. *Let ξ be a finite state Markov process. Then, either*

$$\lim_{u\to 0}\sum_{j\in X}|p_{ij}(s,s+u)-p_j(s+u)|=0$$

for any $s\in\Theta$ and $i\in X(s)$, or

$$1\leqq\sup_{s\in\Theta}\sup_{i\in X(s)}\lim_{u\to 0}\sum_{j\in X}|p_{ij}(s,s+u)-p_j(s+u)|<2.$$

Proof. Combine Theorems 4 and 8, q.e.d.

Acknowledgement. The author wishes to express his gratitude to Professor D.G. Kendall and to Dr. David Williams for their useful comments on this paper

References

1. Bártfai, P., Révész, P.: On a zero-one law. Z. Wahrscheinlichkeitstheorie verw. Gebiete **7**, 43–47 (1967).
2. Blackwell, D., Freedman, D.: The tail σ-field of a Markov chain and a theorem of Orey. Ann. Math. Statistics **35**, 1291–1295 (1964).
3. Cohn, H.: On the tail σ-algebra of the finite nonhomogeneous Markov chains. Ann. Math. Statistics **41**, 2175–2176 (1970).

4. Davydov, Ju. A.: The strong mixing property for Markov chains with a countable number of states. Soviet. Math., Doklady **10**, 825–827 (1969).
5. Iosifescu, M., Theodorescu, R.: Random processes and learning. Berlin-Heidelberg-New York: Springer 1969.
6. Loève, M.: Probability theory. 3d ed. Princeton: Van Nostrand 1963.
7. Orey, S.: An ergodic theorem for Markov chains. Z. Wahrscheinlichkeitstheorie verw. Gebiete **1**, 174–176 (1962).
8. Rosenblatt, M.: Markov processes. Structure and asymptotic behaviour. Berlin-Heidelberg-New York: Springer 1971.

Prof. Dr. M. Iosifescu
Centre of Mathematical Statistics
of the Academy of the
Socialist Republic of Romania
Calea Griviţei 21
Bucharest 12, Romania

(Received March 20, 1972)