Z. Wahrscheinlichkeitstheorie verw. Gebiete
60, 143–150 (1982)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

On Wald's Identity for Dependent Variables

P. Franken and B. Lisek

Sektion Mathematik, Humboldt-Universität, PSF 1297, DDR-1086 Berlin

Summary. Let P be the distribution of a stationary real-valued random sequence $\Phi=\{X_i\}_{i=0}^{\infty}$, and $\tau(\Phi)$ a stopping time with $E_P\tau(\Phi)<\infty$. Then there exists a distribution Q of a stationary sequence $\Psi=\{[X_i, Y_i]\}_{i=0}^{\infty}$, $Y_i\in\{0,1\}$, with the properties

(I) $Q(Y_0=1)>0, \quad Q(\{X_i\}_{i=0}^{\infty}\in(.))=P(\Phi\in(.))$.

(II) $E_Q\left(\sum_{i=0}^{\tau(\Phi)-1} X_i \middle| Y_0=1\right)=E_Q(\tau(\Phi)|Y_0=1)\,E_P X_0$.

Equation (II) is a generalization of Wald's identity.

1. Introduction

We consider the problem of calculating the expectation of a sum constituted by a random number of random variables (r.v.'s). The classical result by Wald [11] is as follows: Let $X_0, X_1, \ldots$ be a sequence of i.i.d. r.v.'s, and let τ be a stopping time with respect to this sequence, i.e. the σ-field generated by $X_0, \ldots, X_{n-1}$ includes the event $\{\tau \leqq n\}$ for all n. The distribution of the given sequence will be denoted by P. Then under some integrability conditions of the usual kind we have

$$E_P \sum_{i=0}^{\tau-1} X_i=(E_P\tau)\,E_P X_0. \tag{1}$$

The proof by Wald [11] (see also Širjaev [10], Borovkov [1]) may not be extended to the case of dependent variables. However a very simple proof of a generalization of (1) was provided in a somewhat different model by Franken and Streller, cf. [4]. This approach will be summarized in Chap. 2. Then the problem arises as to whether this new model is equivalent in some sense to the initial model of a stopped sequence. This question will be answered positively in Chap. 4. In this way we get a formula that generalizes (1) for a stationary sequence of dependent r.v.'s and a stopping time τ with finite expectation.

0044-3719/82/0060/0143/$01.60

2. One Approach to Wald's Identity

Now we will introduce a model, in which Wald's identity (in a generalized form) is very easy to obtain. This approach is closely connected with the concept of random marked point processes and Palm distributions.

Let $\Psi=\{[X_i, Y_i]\}_{i=-\infty}^{+\infty}$ be a (strictly) stationary sequence taking values in $E \times K$. Here E and K are separable metric spaces. The distribution of Ψ will be denoted by Q. For a realization ψ of Ψ we define the shift operator T_j via

$$\psi=\{[x_i, y_i]\}_{i=-\infty}^{+\infty} \to \{[x_{i+j}, y_{i+j}]\}_{i=-\infty}^{+\infty}.$$

We consider a subset $L \subseteq K$ with

$$Q(\# \{i: Y_i \in L\}=\infty)=1. \tag{2}$$

It follows from (2) that $Q(Y_0 \in L)>0$ holds.

Furthermore, we define

$$\nu(\psi)=\inf\{i: i>0, y_i \in L\}.$$

Consider the conditional distribution $Q_L(.)=Q(.\,|Y_0 \in L)$ and let $\Psi_L =\{[X_i^L, Y_i^L]\}_{i=-\infty}^{+\infty}$ be a random sequence with the distribution Q_L. We regard the sequences Ψ and Ψ_L as random marked point processes on Γ (the set of integral numbers) with the mark space $E \times K$. Then Q_L is the Palm distribution of Q with respect to $E \times L$, cf. [3, 5]. Hence it has the following properties:

(a) $Q_L(Y_0^L \in L)=1$,

(b) $E_{Q_L} \nu(\Psi_L)<\infty$,

(c) $Q_L(T_{\nu(\Psi_L)} \Psi_L \in (.))=Q_L(\Psi_L \in (.))$,

and

(d) $$Q(\Psi \in (.))=\frac{1}{E_{Q_L} \nu(\Psi_L)} E_{Q_L} \sum_{i=0}^{\nu(\Psi_L)-1} 1_{\{T_i \Psi_L \in (.)\}}.$$

Equation (d) is the well-known inversion formula for point processes. Using standard techniques of integration theory, we get from (d) the validity of

$$E_Q f(\Psi)=\frac{1}{E_{Q_L} \nu(\Psi_L)} E_{Q_L} \sum_{i=0}^{\nu(\Psi_L)-1} f(T_i \Psi_L) \tag{3}$$

for all measurable functionals f with defined $E_Q f(\Psi)$.

Example. Let μ be the invariant probability measure of an irreducible ergodic Markov chain $\{X_i\}$ with the state space E. Define $Y_i=1_A(X_i)$, where A is a measurable subset of E with $\mu(A)>0$. Using the notations $\Psi=\{[X_i, Y_i]\}$,

$$\mu_A(.)=\frac{\mu((.) \cap A)}{\mu(A)}, \quad K=\{0,1\}, \; L=\{1\},$$

$\tau_A=\nu(\Psi_L)$ - the hitting time of A, we get from (3) for every $f \in L_1(\mu)$

$$E_{\mu_A} \sum_{i=0}^{\tau_A-1} f(X_i)=(E_{\mu_A} \tau_A) E_\mu f(X_0). \tag{4}$$

Formula (4) is well-known in the theory of Markov chains, cf. Chung [2] for discrete E.

Now we choose $f(\psi)=x_0$ in (3) and obtain

$$E_{Q_L} \sum_{i=0}^{\nu(\Psi_L)-1} X_i^L = (E_{Q_L} \nu(\Psi_L))\, E_Q X_0. \tag{5}$$

Comparing (5) with (1) we notice a formal analogy. To clarify the relations between (5) and (1), we consider a sequence $\Phi=\{X_i\}_{i=-\infty}^{+\infty}$ of i.i.d. r.v.'s and a stopping time τ with respect to $\{X_i\}_{i=0}^{\infty}$. The distribution of Φ will be denoted by P. The well-known property that the sequence $\{X_{\tau+i}\}_{i=0}^{\infty}$ has the same distribution as $\{X_i\}_{i=0}^{\infty}$, cf. e.g. Borovkov [1], provides the existence of a random sequence $\Psi_L=\{[X_i^L, Y_i^L]\}_{i=-\infty}^{+\infty}$, $Y_i^L \in \{0,1\}$, with the distribution Q_L, $L=\{1\}$, satisfying the conditions (c),

$$Q_L(\{X_i^L\}_{i=-\infty}^{+\infty} \in (\cdot)) = P(\Phi \in (\cdot)), \tag{6}$$

and

$$\begin{aligned} &Q_L((X_0^L, \dots, X_{\nu(\Psi_L)-1}^L, \nu(\Psi_L)) \in (\cdot)) \\ &\quad = P((X_0, \dots, X_{\tau-1}, \tau) \in (\cdot)). \end{aligned} \tag{7}$$

That means, the stopping rule τ defines the first occurrence of a 1 in the second component of Ψ_L and then the stopping mechanism starts again, terminates at the next 1 and so on. Condition (a) is fulfilled in view of (c) and condition (b) holds in the case of $E_P \tau < \infty$. Thus Q_L is the Palm distribution with respect to $R \times L$ of a stationary distribution Q given by (d). Since the events $\{\nu(\Psi_L) \leqq i\}$ and $\{X_i^L < x\}$ are independent for arbitrary i and x,

$$\begin{aligned} Q(X_0<x) &= \frac{1}{E_{Q_L} \nu(\Psi_L)} \sum_{i=0}^{\infty} Q_L(\nu(\Psi_L)>i, X_i^L<x) \\ &= \frac{1}{E_{Q_L} \nu(\Psi_L)} \sum_{i=0}^{\infty} (Q_L(X_i^L<x) - Q_L(X_i^L<x, \nu(\Psi_L) \leqq i)) \\ &= \frac{1}{E_{Q_L} \nu(\Psi_L)} \sum_{i=0}^{\infty} Q_L(X_0^L<x)(1-Q_L(\nu(\Psi_L) \leqq i)) \\ &= Q_L(X_0^L<x) \end{aligned}$$

holds for all $x \in R$. Then (1) immediately follows from (5), (6), and (7). Thus Wald's identity (1) is a special case of (5) if $E_P \tau < \infty$ holds.

In the case of dependent r.v.'s, however, the Eq. (5) cannot be reduced to the simple form (1) in general, as the following example shows:

Let $\{X_i\}_{i=0}^{\infty}$ be a homogeneous Markov chain taking values 1, 2, and 3. Its distribution P may be defined by the transition matrix

$$\begin{Vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{Vmatrix}$$

and by the stationary initial probabilities $q_1 = q_2 = q_3 = 1/3$. Consider the stopping time

$$\tau = 1 + \inf\{i\colon i \geqq 0,\ X_i \in \{1, 2\}\}.$$

Then $E_P \tau = 4/3$, $E_P X_0 = 2$, but

$$E_P \sum_{i=0}^{\tau-1} X_i = 7/3 \neq 2 \cdot (4/3).$$

A formula for the difference

$$E_P \sum_{i=0}^{\tau-1} X_i - (E_P \tau) E_P X_0$$

was given in the case of a stopped irreducible Markov chain with finite state space by Küchler and Semenov [6]. (In [6] a somewhat general model was investigated.) We mention that this formula is quite complicated and hard to use.

Example. We apply (5) to the analysis of a queueing model $G/G/1/\infty$, where the sequence $\{[\alpha_n, \beta_n]\}_{n=-\infty}^{+\infty}$ of interarrival and service times is only assumed to be strictly stationary and ergodic, and to satisfy $\rho = E_P \beta_0 / E_P \alpha_0 < 1$. Then there exists a uniquely determined stationary and ergodic sequence $\Psi = \{[\alpha_n, \beta_n, w_n]\}_{n=-\infty}^{+\infty}$ with a distribution Q satisfying

$$Q(w_{n+1} = (w_n - \alpha_n + \beta_n)_+ \text{ for all } n) = 1,$$

cf. e.g. Loynes [8], and [4]. The variables w_n are the waiting times of customers. The distribution Q has the property

$$Q(\# \{n\colon w_n = 0\} = \infty) = 1.$$

For the analysis of a generic busy cycle we have to investigate the sequence $\Psi_L = \{[\alpha_n^L, \beta_n^L, w_n^L]\}$ with the distribution $Q_L(.) = Q(. \mid w_0 = 0)$. (Here, in terms of our previous considerations, $K = R_+^2$, $L = R_+ \times \{0\}$.) Now we can use (5):

$$E_{Q_L} \eta = (E_{Q_L} \nu) E_Q \beta_0 = (E_{Q_L} \nu) E_P \beta_0,$$

where $\nu = \nu(\Psi_L)$ is the number of customers served during the busy cycle, and $\eta = \beta_0^L + \ldots + \beta_{\nu-1}^L$ is the length of the busy period. For further results in this direction cf. [4].

3. The General Problem

We are interested in results of the form (5). Thus our aim is to clarify whether there is a Palm distribution corresponding to a given stopped sequence. A similar problem occurs in the investigations of Mogulski and Trofimov [9]. We will briefly summarize them.

Let $\{\kappa_i\}_{i=0}^{\infty}$ be an irreducible and aperiodic homogeneous Markov chain with the state space $\{1, 2, \ldots, k\}$, and P its stationary distribution. Furthermore, let $[\![a_{ij}]\!]_{i,j\in\{1,\ldots,k\}}$ be a matrix of positive real numbers. Consider the sums

$$S_n=\sum_{i=1}^{n} a_{\kappa_i\kappa_{i+1}},\qquad n=1,2,\ldots.$$

(In the special case of $a_{ij}=i$ we get

$$S_n=\sum_{i=1}^{n}\kappa_i.\Big)$$

Let τ be a stopping time of the chain $\{\kappa_i\}_{i=0}^{\infty}$. Consider a new transition matrix defined by

$$[\![\tilde{P}_{ij}]\!]=[\![P(\kappa_\tau=j\,|\,\kappa_0=i)]\!].$$

This matrix describes the transition to the time of the next stopping. We remark that there are cases in which $[\![\tilde{P}_{ij}]\!]$ is not irreducible and aperiodic. Mogulski and Trofimov [9] assume that there is exactly one stationary initial distribution corresponding to $[\![\tilde{P}_{ij}]\!]$. Then

$$E_{\tilde{P}}S_\tau=(E_{\tilde{P}}\tau)\,E_P S_1$$

holds, where $\tilde{P}$ is the stationary distribution of the Markov chain with the transition matrix $[\![\tilde{P}_{ij}]\!]$. The proof is very simple by use of the individual ergodic theorem.

Comparing this result with (5) we see that $\tilde{P}$ plays the role of Q_L in our context, where the existence of the distribution $\tilde{P}$ had to be assumed. Now we want to formulate the existence problem in our general model. We consider a stationary real-valued sequence $\Phi=\{X_i\}_{i=-\infty}^{+\infty}$ with the distribution P. Let τ be a stopping time with respect to $\{X_i\}_{i=0}^{\infty}$. This is equivalent to the existence of a sequence of measurable functions $f_n\colon R^n\to\{0,1\}$, $n=1,2,\ldots$, with

$$1_{\{\tau\leqq n\}}=f_n(X_0,\ldots,X_{n-1})\qquad P\text{-a.s.},$$

i.e.

$$\tau=\inf\{n\colon n\geqq 1, f_n(X_0,\ldots,X_{n-1})=1\}\qquad P\text{-a.s.}.$$

We will use the following notations:

$$\varphi=\{x_i\}_{i=-\infty}^{+\infty}\ \text{– a realization of } \Phi,$$

$$\theta_j\varphi=\{x_{i+j}\}_{i=-\infty}^{+\infty},$$

$$\tau(\varphi)=\inf\{n\colon n\geqq 1, f_n(x_0,\ldots,x_{n-1})=1\}.$$

We want to investigate the existence of a distribution Q of a random sequence $\Psi=\{[X_i,Y_i]\}_{i=-\infty}^{+\infty}$, $Y_i\in\{0,1\}$, with the properties:

(A) Q is stationary and (2) holds for $L=\{1\}$.

(B) For the Palm distribution Q_1 of Q with respect to $R\times\{1\}$ (cf. Chap. 2)

$$Q_1(\nu(\Psi_1)=\tau(\{X_i^1\}_{i=-\infty}^{+\infty}))=1$$

holds, where $\Psi_1 = \{[X_i^1, Y_i^1]\}_{i=-\infty}^{+\infty}$ is a random sequence distributed according to Q_1, in particular

$$E_{Q_1} \nu(\Psi_1) = E_Q(\tau(\Phi) \mid Y_0 = 1).$$

(C) $Q(\{X_i\}_{i=-\infty}^{+\infty} \in (.)) = P(\Phi \in (.))$.

4. Solution of the Problem

Let $i \in \Gamma$ and φ be fixed. Consider the number

$$n(i, \varphi) = \sum_{j=-\infty}^{i} 1_{\{\tau(\theta_j \varphi) > i-j\}}$$

of all indices $j \leqq i$ with $\tau(\theta_j \varphi) > i - j$. We define

$$M' = \{\varphi : n(i, \varphi) < \infty \text{ for all } i \in \Gamma\}.$$

Now we can formulate our main result:

Theorem. *Let Φ be a stationary sequence with the distribution P, and τ a stopping time defined by a sequence of measurable functions $\{f_n\}_{n=1}^{\infty}$. We assume that*

$$P(\Phi \in M') = 1 \tag{8}$$

holds. Then there is a distribution Q of a random sequence Ψ with the properties (A), (B), *and* (C).

The proof of the theorem is given in appendix. From Campbell's Theorem it follows that

$$E_P \tau < \infty \tag{9}$$

is sufficient for (8), cf. [3, 4]. Thus we obtain the following

Corollary. *If $E_P \tau < \infty$, then there is a distribution Q_1 with*

$$E_{Q_1} \sum_{i=0}^{\nu(\Psi_1)-1} X_i^1 = (E_{Q_1} \nu(\Psi_1)) E_P X_0. \tag{10}$$

Using the notation Φ for the first component of Ψ, we can rewrite (10) into the form (cf. (B))

$$E_Q \left(\sum_{i=0}^{\tau(\Phi)-1} X_i \mid Y_0 = 1 \right) = E_Q(\tau(\Phi) \mid Y_0 = 1) E_P X_0.$$

This is the desired generalization of Wald's identity for dependent variables.

5. Some Remarks

The assumption (9) seems not to be natural, since the expectation $E_P \tau$ does not occur in any form in (10). On the other hand, the statement of the theorem does not hold without any assumptions, as the following example shows: Let Φ be a sequence of i.i.d. r.v.'s with $P(X_i=k)=6/(\pi^2 k^2)$, $k=1,2,\ldots$, and $\tau=X_0$. Then

$$E_P \tau=\infty \tag{11}$$

and $P(\Phi\in M')=0$ hold. (The latter may be proved by direct calculating.) Now suppose that there were distributions Q and Q_1 with (A), (B), and (C). Then the points i with $Y_i=1$ would form a stationary renewal process. In view of (B) this contradicts (11). Thus a solution of (A), (B), (C) does not exist for our example. Of course Wald's identity (1) holds in the sense that both sides are infinite.

A last remark concerns the possibility to generalize Wald's equations of higher order. It would be very important for statistical applications to have equations for higher moments, too. However, there is no direct generalization of Wald's fundamental identity

$$E_P \frac{\exp\left(u\sum_{i=0}^{\tau-1} X_i\right)}{(E_P \exp(uX_0))^\tau}=1 \qquad \text{for some } u\neq 0$$

(which is true for a stopped sequence of i.i.d. r.v.'s) in our model from Chap. 2. For example, consider the sequence $\Psi_1=\{[X_i^1, Y_i^1]\}_{i=-\infty}^{+\infty}$ with the Palm distribution Q_1 defined by (c) and $Q_1(Y_0^1=1,\ Y_1^1=0,\ Y_2^1=1,\ X_0^1=0,\ X_1^1=1)=1$. Then

$$Q_1\left(\nu(\Psi_1)=2, \sum_{i=0}^{\nu(\Psi_1)-1} X_i^1=1\right)=1$$

and $Q(X_0=0)=Q(X_0=1)=1/2$ hold. (The latter follows from the inversion formula (d).) Hence

$$E_{Q_1} \frac{\exp\left(u\sum_{i=0}^{\nu(\Psi_1)-1} X_i^1\right)}{(E_Q \exp(uX_0))^{\nu(\Psi_1)}}=\frac{4\exp(u)}{(1+\exp(u))^2}\neq 1$$

follows for all $u\neq 0$. Notice that

$$\tau(\{X_i\}_{i=-\infty}^{+\infty})=\inf\left\{k\colon k>0, \sum_{i=0}^{k-1} X_i=1\right\}$$

is a stopping time with respect to $\{X_i\}_{i=0}^{\infty}$, and (A), (B) hold.

Appendix: The Proof of the Theorem

A one-one mapping between the stationary sequence $\Psi=\{[X_i, Y_i]\}_{i=-\infty}^{+\infty}$, for which (A), (B), and (C) hold, and a stationary sequence $\Omega=\{[X_i, Z_i]\}_{i=-\infty}^{+\infty}$, $Z_i\in\{0,1,2,\ldots\}$, with the distribution H that fulfills

$$H(Z_{j+1}=(Z_j+1)1_{\{\tau(\{X_{i+j-Z_j}\}_{i=-\infty}^{+\infty})\neq Z_j+1\}} \quad \text{for all } j\in\Gamma)=1 \tag{12}$$

may be defined by $Z_i = \inf\{i-j\colon j \leqq i,\ Y_j = 1\}$ or, conversely, $Y_i = 1_{\{Z_i = 0\}}$. The number Z_i may be interpreted as the time from the preceding point with mark 1 to the point i. Then it becomes clear that (12) is equivalent to

$$Q(\nu(T_j \Psi) = \tau(\{X_{i+j}\}_{i=-\infty}^{+\infty}) \text{ for all } j \text{ with } Y_j = 1) = 1. \tag{13}$$

Under the condition (A) the statement (13) is equivalent to (B) by means of the definition of Q_1. Condition (2) is equivalent to $H(\#\{i\colon Z_i = 0\} = \infty) = 1$, and this equation follows from (8) and (12). Thus it suffices to show the existence of a stationary sequence Ω with (12). To do this, we will apply a general theorem by Lisek [7] concerning the existence of stationary solutions of recursive stochastic equations.

Let $i \in \Gamma$ and φ be fixed. The points $j \leqq i$ for which $\tau(\theta_j \varphi) > i-j$ holds will be ordered and denoted by $v_1(i,\varphi), \ldots, v_{n(i,\varphi)}(i,\varphi)$. The sets

$$A(i,\varphi) = \{i - v_1(i,\varphi), i - v_2(i,\varphi), \ldots, i - v_{n(i,\varphi)}(i,\varphi)\}$$

are finite for P-almost-all φ (because of (8)) for all $i \in \Gamma$. All these sets are not empty since $0 \in A(i,\varphi)$. The condition $f(\theta_i \varphi, A(i,\varphi)) \subseteq A(i+1,\varphi)$ with $f(\varphi, z) = (z+1)\, 1_{\{\tau(\theta_{-z}\varphi) \neq z+1\}}$ is valid for all i and almost all φ. That means, the assumptions of Corollary 1 in [7] are fulfilled. From there the existence of a stationary sequence Ω with (12) follows. The proof is completed.

References

1. Borowkow, A.A.: Wahrscheinlichkeitstheorie. Berlin: Akademie-Verlag 1976
2. Chung, K.L.: Markov Chains with Stationary Transition Probabilities. Berlin-Göttingen-Heidelberg: Springer 1960
3. Franken, P.: Einige Anwendungen der Theorie zufälliger Punktprozesse in der Bedienungstheorie I. Math. Nachr. **70**, 303–319 (1976)
4. Franken, P., König, D., Arndt, U., Schmidt, V.: Queues and Point Processes. Berlin: Akademie-Verlag. New York: Wiley 1981
5. König, D., Matthes, K., Nawrotzki, K.: Unempfindlichkeitseigenschaften von Bedienungsprozessen. Appendix to: Gnedenko, B.W., Kowalenko, I.N.: Einführung in die Bedienungstheorie, Berlin: Akademie-Verlag 1971
6. Küchler, I., Semenov, A.: Die Waldsche Fundamentalidentität und ein Sequentieller Quotiententest für eine zufällige Irrfahrt über einer homogenen irreduziblen Markovschen Kette mit endlichem Zustandsraum. Math. Operationsforsch. Statist., Ser. Statistics, **10**, no. 2, 319–331 (1979)
7. Lisek, B.: A Method for Solving a Class of Recursive Stochastic Equations. [To appear in: Z. Wahrscheinlichkeitstheorie verw. Gebiete]
8. Loynes, R.M.: The stability of a queue with non-independent inter-arrival and service times. Proc. Cambridge Philos. Soc. **58**, 494–520 (1962)
9. Mogulski, A.A., Trofimov, V.K.: Wald's Identity and Coding Costs for Markov Chains. (In Russian) Proc. 7th All-Union Symposium on Redundancy Problem in Information Systems. Leningrad 1977
10. Širjaev, A.N.: Statistical Sequential Analysis. (In Russian). Moscow: Nauka 1968
11. Wald, A.: Sequential Analysis. New York: Wiley 1947

Received January 22, 1981; in revised form December 12, 1981