

On occurrence of subpattern and method of gambling teams

Vladimir Pozdnyakov

Received: 15 September 2005 / Revised: 23 March 2006 / Published online: 21 September 2006
© The Institute of Statistical Mathematics, Tokyo 2006

Abstract The expected value and generating function of the number of overlapping occurrences of a pattern P in a Markov chain until the first occurrence of a member from a finite collection of patterns that start with P is obtained. A martingale technique is employed to address the problem.

Keywords Pattern · Gambling teams · Waiting times · Markov chains · Martingales · Test of randomness

Mathematics Subject Classification (2000) Primary: 60C05 · 60G42;
Secondary: 60G40 · 62E17

1 Introduction

The occurrence of patterns in stochastic sequences is an important topic with a wide range of applications, and it has been studied by many different methods. The first systematic treatment of the problem can be found in Feller (1968) who used the recurrent event theory. The combinatorial methods were introduced by Guibas and Odlyzko (1981a,b). Various probabilistic techniques were employed in Biggins and Cannings (1987), Blom and Thorburn (1982), Breen et al. (1985), Chrysaphinou and Papastavridis (1990), Han and Hirano (2003), Robin and Daudin (2001), Stefanov (2003) and Uchida (1998). The modern probabilistic approach is the Markov chain embedding method. For this method and related techniques we refer readers to Antzoulakos (2001), Fu (1986, 2001), Fu and Chang (2002), Fu and Koutras (1994), Stefanov (2000), and Stefanov

V. Pozdnyakov (✉)
Department of Statistics, University of Connecticut 215 Glenbrook Road,
U-4120, Storrs, CT 06269-4120, USA
e-mail: Vladimir.Pozdnyakov@Uconn.edu

and Pakes (1997, 1999). Two recent books provide a useful review of these methods: Balakrishnan and Koutras (2002) and Fu and Lou (2003).

In this note we will use the martingale method that was introduced by Li (1980) and Li and Gerber (1981), and further developed in Pozdnyakov and Kulldorff (2006) and Pozdnyakov, Glaz, Kulldorff, Steele (2005). The idea of the method is that a valuable information on the occurrence of patterns can be obtained from a specially designed martingale. However, the common perception is that the martingale approach is not easy to be employed for Markov dependent trials. But, in fact, it has been shown by Glaz, Kulldorff, Pozdnyakov, Steele (2006) that the occurrence of patterns in a Markov chain also can be treated by the martingale technique.

Here we will consider the problem of the occurrence of subpattern in Markov chains. The simplest version of this problem in the case of the Bernoulli trials is as follows. What is the distribution of the number of (overlapping) occurrences of run of length l until run of length $L > l$ is observed? Closely related results can be found in Aki, Balakrishnan, Mohanty (1996), Hirano and Aki (2003) and Uchida (1998), somewhat related – in Aki and Hirano (1994), Chadjiconstantinidis and Koutras (2001) and Hirano et al. (1997). The progress made here is twofold. First, we demonstrate that the martingale approach is fruitful in this case also. Second, what we consider here is more general than that in the papers mentioned above. Here the stopping rule is determined by many patterns, not just by one. A potential application of the introduced technique is a test of randomness that can be designed in a way similar to one suggested by Rukhin (2001, 2002). A detailed exposition is provided in the last section of the note.

2 Formal statement of the problem

Let $\{Z_n, n \geq 1\}$ be a Markov chain of order m over a finite alphabet Ω . Let $P = a_1 a_2 \dots a_l$ be a pattern (a word from the alphabet Ω) of length $l \geq m$. Consider a collection of patterns $\mathcal{C} = \{PB_1, PB_2, \dots, PB_K\}$, where $PB_i = a_1, \dots, a_l b_1^{(i)}, \dots, b_l^{(i)}$. Assume that no pattern from \mathcal{C} contains another as a subpattern. Let τ_i be the waiting time until the first occurrence of the pattern PB_i , and

$$\tau = \min\{\tau_1, \dots, \tau_K\}.$$

Assume that $\mathbf{P}(\tau < \infty) = 1$, and $\mathbf{P}(\tau = \tau_i) > 0$ for all $1 \leq i \leq K$. The main object of interest is the number of overlapping occurrences of subpattern P until τ which is denoted by \mathcal{N} . The goal is to obtain the expected value and generating function of \mathcal{N} .

3 Key martingale related to pattern PB

Let us imagine a table in a casino that generates Markov chain $\{Z_n, n \geq 1\}$. Consider a gambler that arrives in the casino right after n -th trial. First, the gambler

watches l consecutive trials. If he does not see pattern P then he does not bet. If he sees P then he starts to place bets on $B = b_1, \dots, b_N$. More specifically, he bets a dollar that the

$$Z_{n+l+1} = b_1.$$

If it is not, he loses his bet and leaves the game. If he wins, he gets

$$1/p_{a_{l-m+1}, \dots, a_l|b_1},$$

where $p_{a_{l-m+1}, \dots, a_l|b_1} = \mathbf{P}(Z_{m+1} = b_1 | Z_1 = a_{l-m+1}, \dots, Z_m = a_l)$, the transition probability. Then he bets his entire capital on the event that

$$Z_{n+l+2} = b_2.$$

If he loses, he leaves. If he wins his total capital is now

$$1/p_{a_{l-m+1}, \dots, a_l|b_1} / p_{a_{l-m+2}, \dots, a_l|b_1|b_2}.$$

He continues the betting through the pattern B . If he finishes the pattern, he leaves the game with his winning.

Now, if $\{X_n, n \geq 1\}$ is a total amount of money that the casino has from the gambler at moments n , then it is obvious that this stochastic sequence forms a martingale with respect to filtration generated by $\{Z_n, n \geq 1\}$. The initial value of this martingale is 0.

4 Expected value of \mathcal{N}

To compute the expected number of occurrences till τ we introduce K teams of gamblers. A gambler from the j -th team that arrives at moment n bets y_j dollars (the amount of money that we specify a bit later) on pattern A_j in the way described in the previous section. Let X_n be a net gain of the casino at moment n . Since it is a sum of finite number of martingales, X_n is a martingale itself.

Next step is to obtain the expression of the stopped martingale X_τ . To do that one needs to introduce a new notion of “profit matrix”. Let $y_j W_{ij}$ will be total winnings of the j -th team whenever $\tau = \tau_i$. The matrix $W_{ij}, 1 \leq i, j \leq K$ is called the *profit matrix*. Only gamblers who observe P enter the game, and most of them are losers. Also note only those gamblers that enter the game right before the stopping time τ (more precisely, those who enter the game after $\tau - L$, where L is length of the longest pattern from \mathcal{C}) can have some money in their pockets at time τ . Moreover, when $\tau = \tau_i$, the amount of winnings of the j -th team is completely determined by overlapping of PB_i and PB_j , and this quantity is a constant not a random variable. Later an explicit formula for W_{ij} will be provided. Also one needs to distinguish the patterns from \mathcal{C} that ends by P from those that do not. If the game is stopped because of the first occurrence

of a pattern that also ends with P the total amount of money paid by the j -th team is $y_j(\mathcal{N} - 1)$. The gambler who observes the last occurrence of P does not place his bet before τ . If the game is stopped because of the first occurrence of a pattern that does not end with P the total amount of money paid by the j -th team is $y_j\mathcal{N}$.

Without loss of generality let us assume that PB_1, PB_2, \dots, PB_Q ends with P , and $PB_{Q+1}, PB_{Q+2}, \dots, PB_K$ do not. Then the stopped martingale is given by

$$X_\tau = \begin{cases} (y_1 + \dots + y_K)(\mathcal{N} - 1) - (W_{11}y_1 + \dots + W_{1K}y_K), & \text{if } \tau = \tau_1, \\ \dots & \dots \\ (y_1 + \dots + y_K)(\mathcal{N} - 1) - (W_{Q1}y_1 + \dots + W_{QK}y_K), & \text{if } \tau = \tau_Q, \\ (y_1 + \dots + y_K)\mathcal{N} - (W_{Q+11}y_1 + \dots + W_{Q+1K}y_K), & \text{if } \tau = \tau_{Q+1}, \\ \dots & \dots \\ (y_1 + \dots + y_K)\mathcal{N} - (W_{K1}y_1 + \dots + W_{KK}y_K), & \text{if } \tau = \tau_K. \end{cases}$$

Or one can write this in a more compact form

$$X_\tau = \sum_{j=1}^K y_j\mathcal{N} - \sum_{i=1}^Q \sum_{j=1}^K y_j(W_{ij} + 1)I_{\tau=\tau_i} - \sum_{i=Q+1}^K \sum_{j=1}^K y_jW_{ij}I_{\tau=\tau_i}.$$

Now, let us assume that there exists a solution, (y_1^*, \dots, y_K^*) , of the following linear system:

$$\begin{aligned} (W_{11} + 1)y_1 + \dots + (W_{1K} + 1)y_K &= 1 \\ \dots & \dots \\ (W_{Q1} + 1)y_1 + \dots + (W_{QK} + 1)y_K &= 1 \\ W_{(Q+1)1}y_1 + \dots + W_{(Q+1)K}y_K &= 1 \\ \dots & \dots \\ W_{K1}y_1 + \dots + W_{KK}y_K &= 1 \end{aligned} \tag{1}$$

For this choice of the initial bets (y_1^*, \dots, y_K^*) , the stopped martingale is given by a very simple formula

$$X_\tau = (y_1^* + \dots + y_K^*)\mathcal{N} - 1.$$

Now one can show that the increments of the martingale X_n are almost sure bounded, and the expected value of τ is finite. Therefore, by Optional-Stopping Theorem (e.g., Williams 1991, p. 100) we find that

$$\mathbf{E}(X_\tau) = 0.$$

Since the weights are not random, we get the following result.

Theorem 1 *If (y_1^*, \dots, y_K^*) solves system (1), then*

$$\mathbf{E}(\mathcal{N}) = \frac{1}{y_1^* + \dots + y_K^*}.$$

Remark 1 An interesting observation [which is consistent, for instance, with results of Uchida (1998)] is that the expected value of \mathcal{N} does not depend on the initial distribution of the Markov chain.

5 Computation of the profit matrix W_{ij}

One can say that the definition of W_{ij} is too vague, but, in fact, it is not. The explicit expression for that matrix can be presented. First, one needs to introduce the following measure of overlapping of $B = b_1b_2, \dots, b_N$ and $C = c_1c_2, \dots, c_M$:

$$W(B, C) = \sum_{k=1}^{\min(N,M)-l} \delta_k(B, C),$$

where

$$\delta_k(B, C) = \begin{cases} \frac{1}{P_{c_{l-m+1} \dots c_l | c_{l+1} \dots c_l} P_{c_{l-m+k} \dots c_{l+k-1} | c_{l+k}}}, & \text{if } c_1 = b_{N-k-l+1}, \\ & c_2 = b_{N-k-l+2}, \\ & \dots \\ & c_{l+k} = b_N, \\ 0, & \text{otherwise,} \end{cases}$$

and l is the length of the pattern P . Now, having this notation at hand one can show that

$$W_{ij} = W(PB_i, PB_j).$$

For instance, if $\{Z_n, n \geq 1\}$ is a second-order Markov chain with two states H and T , $C = \{HHTHH, HHHHT\}$, and $P = HH$, then

$$W_{21} = W(HHHHT, HHTHH) = \frac{1}{p_{HH|T}} + 0 + 0,$$

and

$$W_{22} = W(HHHHT, HHHHT) = 0 + 0 + \frac{1}{p_{HH|HPHH|HPHH|T}}.$$

6 Examples: expected number of occurrences

Let $\{Z_n, n \geq 1\}$ be a sequence of Bernoulli trials with $\mathbf{P}(Z_n = H) = p > 0$ and $\mathbf{P}(Z_n = T) = q > 0$. Let $C = \{HHTHH, HHHHT\}$, and $P = HH$. Then the matrix $W_{ij}, 1 \leq i, j \leq 2$ is equal to

$$\begin{pmatrix} \frac{1}{p^2q} & 0 \\ \frac{1}{q} & \frac{1}{p^2q} \end{pmatrix}.$$

Solving (1) we find that

$$y_1^* = \frac{p^2q - p^4q^2}{1 + p^2q - p^4q}, \quad y_2^* = \frac{p^2q(1 - p^2 + p^2q)}{(1 + p^2q - p^4q)}.$$

Finally, by Theorem 1 we find that

$$\mathbf{E}(\mathcal{N}) = \frac{1 + p^2q - p^4q}{2p^2q - p^4q}.$$

Remark 2 The nature of the problem is such that making $\{Z_n, n \geq 1\}$ a Markov chain does not really complicate the problem. Since every gambler watches the game for l rounds first they know how to bet in a fair fashion even if we have Markov dependent sequence with an dependence order less than l . The exactly same algorithm works in this case.

For example, let now $\{Z_n, n \geq 1\}$ be a second-order two-state Markov chain, and again $\mathcal{C} = \{HHTHH, HHHHT\}$, $P = HH$. The profit matrix then is

$$\begin{pmatrix} \frac{1}{(1 - p_{HH|H})p_{HT|H}p_{TH|H}} & 0 \\ \frac{1}{1 - p_{HHH}} & \frac{1}{(p_{HH|H})^2(1 - p_{HH|H})} \end{pmatrix}.$$

Solving system (1) and applying Theorem 1 we obtain that

$$\mathbf{E}(\mathcal{N}) = \frac{1 + (1 - p_{HH|T})^2(1 + p_{HH|H})p_{HT|H}p_{TH|H}}{(1 - p_{HH|H})((1 - p_{HT|H}p_{TH|H})(p_{HH|H})^2 + p_{HT|H}p_{TH|H})}.$$

Remark 3 Let us now comment on computational difficulties that one can have with the usage of this approach. The algorithm is consist of three step: first we form a profit matrix, then we solve a linear system associated with the matrix, and, finally, we apply formula of Theorem 1. First of all, the examples above show us that the imposing Markov dependence structure on a stochastic sequence does not complicate calculations – unconditional probabilities should be simply substituted by conditional ones. Also making the alphabet bigger is not an issue. A computational difficulty can be caused by the size of the profit matrix. The size of W_{ij} is always $K \times K$, where K is the number of stopping patterns in list \mathcal{C} . If the number of patterns in \mathcal{C} is huge it will be not easy to use the Theorem 1. Fortunately, the intended application of this method – test of randomness – does not require a long list of stopping patterns. Also we do not

think that any other method can lead to smaller matrices. It is not clear for the moment how this problem can be solved by an appropriate Markov chain embedding. But most likely the number of needed states will be a number of all the prefixes of the compound pattern \mathcal{C} . As a consequence, the size of the transition matrix of the embedded Markov chain will be larger than $K \times K$, and it will be significantly larger if the number of patterns, K , is much smaller than the average length of patterns in \mathcal{C} .

7 Generating function of \mathcal{N}

When the martingale method of gambling teams has been already developed for the expected value, the transition to the generating function is relatively easy. All the needed ideas have been introduced. We will use the same key martingale, but now we play a bit with the amount of initial bets.

Again the gambler from j -th teams first observes l consecutive trials. If he does not see P he leaves the game without playing. If it is P he bets $y_j \alpha^k$ dollars on B_j where k is a number of overlapping occurrences of P till this moment, and $0 < \alpha < 1$. Since the size of the bet is fully determined by the previous history of the process, the total net gain of the casino, X_n , forms a martingale. Let $\alpha^N y_j W_{ij}(\alpha)$ be total winnings of the j -team when game is ended with the pattern PB_i . Then the stopped martingale X_τ is given by

$$X_\tau = \sum_{j=1}^K y_j \alpha \frac{1 - \alpha^{\mathcal{N}}}{1 - \alpha} - \alpha^{\mathcal{N}} \left[\sum_{i=1}^Q \sum_{j=1}^K y_j (W_{ij}(\alpha) + 1) I_{\tau=\tau_i} + \sum_{i=Q+1}^K \sum_{j=1}^K y_j W_{ij}(\alpha) I_{\tau=\tau_i} \right].$$

As before the key property of $W_{ij}(\alpha)$ is that it is *not* a random variable. The explicit algorithm that allows us to compute $W_{ij}(\alpha)$ will be presented a bit later. Now, let us choose weights (y_1, \dots, y_K) in such a way that

$$\begin{aligned} (W_{11}(\alpha) + 1)y_1 + \dots + (W_{1K}(\alpha) + 1)y_K &= 1 \\ \dots &\dots \\ (W_{Q1}(\alpha) + 1)y_1 + \dots + (W_{QK}(\alpha) + 1)y_K &= 1 \\ W_{Q+11}(\alpha)y_1 + \dots + W_{Q+1K}(\alpha)y_K &= 1 \\ \dots &\dots \\ W_{K1}(\alpha)y_1 + \dots + W_{KK}(\alpha)y_K &= 1 \end{aligned} \tag{2}$$

With this choice of the weights denoted by (y_1^*, \dots, y_K^*) we get the following simple expression for the stopped martingale:

$$X_\tau = \sum_{j=1}^K y_j^* \alpha \frac{1 - \alpha^{\mathcal{N}}}{1 - \alpha} - \alpha^{\mathcal{N}}.$$

After the routine application of the Optional-Stopping Theorem and a bit of algebra we get the following result.

Theorem 2 *If (y_1^*, \dots, y_K^*) solves system (2), then*

$$\mathbf{E}\alpha^{\mathcal{N}} = 1 - \frac{1}{\frac{\alpha}{1-\alpha}(y_1^* + \dots + y_K^*) + 1},$$

8 Computation of $W_{ij}(\alpha)$

One can derive an explicit formula for matrix $W_{ij}(\alpha)$. Consider two patterns $B = b_1b_2\dots b_N$ and $C = c_1c_2\dots c_M$. Let us introduce the following measure of overlapping of B and C :

$$W(B, C)(\alpha) = \sum_{k=1}^{\min(N,M)-l} \delta_k(B, C)\alpha^{-n_k},$$

where $\delta_k(\cdot, \cdot)$ is the same function defined earlier, and n_k is the number of occurrences of P in the suffix $b_{N-k-l+2}\dots b_N$. The matrix $W_{ij}(\alpha)$ is given by

$$W_{ij}(\alpha) = W(PB_i, PB_j)(\alpha).$$

9 Examples: generating function

Consider a sequence of Bernoulli trials $\{Z_n, n \geq 1\}$ with $\mathbf{P}(Z_n = H) = p > 0$ and $\mathbf{P}(Z_n = T) = q > 0$. Let $\mathcal{C} = \{HHTHH, HHHHT\}$, and $P = HH$. Then the matrix $W_{ij}(\alpha), 1 \leq i, j \leq 2$ is given by

$$\begin{pmatrix} 1 & 0 \\ \alpha p^2 q & 1 \\ \frac{1}{q} & \frac{1}{\alpha^2 p^2 q} \end{pmatrix}$$

Theorem 2 then tells us that

$$\mathbf{E}(\alpha^{\mathcal{N}}) = \frac{\alpha^2 p^2 (-1 + \alpha(-1 + \alpha p^2))q}{-1 + \alpha + \alpha p^2 (-1 + \alpha^2 (-1 + p^2))q}.$$

One can check also that

$$\frac{\partial \mathbf{E}(\alpha^{\mathcal{N}})}{\partial \alpha} \Big|_{\alpha=1} = \frac{1 + p^2 q - p^4 q}{2p^2 q - p^4 q},$$

the expected value of \mathcal{N} .

Let us give another example. Again we consider Bernoulli trials $\{Z_n, n \geq 1\}$ with $\mathbf{P}(Z_n = H) = p > 0$ and $\mathbf{P}(Z_n = T) = q > 0$. Let $P = HH$, but now $\mathcal{C} = \{HHH, HHT\}$. The matrix $W_{ij}(\alpha), 1 \leq i, j \leq 2$ is given by

$$\begin{pmatrix} \frac{1}{\alpha p} & 0 \\ 0 & \frac{1}{q} \end{pmatrix},$$

$y_1^* = p^2\alpha/(1 + p\alpha), y_2^* = 1 - p$. After a bit arithmetics the formula of Theorem 2 gives us

$$\mathbf{E}(\alpha^{\mathcal{N}}) = \alpha(1 - p) + p\alpha^2.$$

Now, this answer can be easily obtained without any help of Theorem 2. Indeed, as soon as we have run HH we stop the next round. Therefore, \mathcal{N} can take only two values: 1 when we are stopped by HHT , or 2 if we stopped because of HHH . Finally, for instance, using Li (1980) one can obtain $\mathbf{P}(\tau = \tau_1) = p$ and $\mathbf{P}(\tau = \tau_2) = 1 - p$.

10 Application: test of randomness

Most standard generators of random numbers are really pseudo-random. Various deterministic procedures are employed to produce “random” sequences. As a consequence, some patterning can be observed for such generators. Therefore, testing based on occurrence of words (patterns) seems to be a natural idea (see Rukhin 2001, 2002).

In particular, in Rukhin (2001, p. 118) a number of occurrences of a given (nonoverlapping with itself or nonperiodic) word in a random text is suggested as a test statistic for randomness. Depending on the relationship between the length of the word and the length of the random text, a Poisson limiting distribution or normal can be used to make inference.

Clearly, that random variable \mathcal{N} can be also employed to design a similar test of randomness. The difference between our test and one suggested by Rukhin (2001) that here we stop at random time τ . But it is not bad. The stopping time is defined by collection \mathcal{C} the choice of which is completely up to us. As a result, we can control the *average* stopping time. In general, to increase $\mathbf{E}(\tau)$ we need longer patterns in \mathcal{C} and a smaller K , cardinality of \mathcal{C} . In return, we have an exact distribution, and we are not restricted to nonperiodic words. Also the previous section demonstrates that \mathcal{C} provides a good control over the distribution of \mathcal{N} as well.

Let us consider a four-letter alphabet $\{1, 2, 3, 4\}$ with the uniform distribution over it. We assume that letters are drawn in an independent way. The testing word is $P = 111$, and

Table 1 Test of randomness

	Exact	Simulated
$\mathbf{P}(\mathcal{N} = 1)$	0.18750	0.18788
$\mathbf{P}(\mathcal{N} = 2)$	0.15234	0.15019
$\mathbf{P}(\mathcal{N} = 3)$	0.17456	0.17714
$\mathbf{P}(\mathcal{N} = 4)$	0.11961	0.11831
$\mathbf{P}(\mathcal{N} = 5)$	0.09183	0.09156
$\mathbf{P}(\mathcal{N} = 6)$	0.06848	0.06797
$\mathbf{P}(\mathcal{N} = 7)$	0.05143	0.05125
$\mathbf{P}(\mathcal{N} = 8)$	0.03856	0.03909
$\mathbf{P}(\mathcal{N} = 9)$	0.02892	0.02891
$\mathbf{P}(\mathcal{N} = 10)$	0.02169	0.02255
$\mathbf{P}(\mathcal{N} = 11)$	0.01627	0.01629
$\mathbf{P}(\mathcal{N} > 11)$	0.04879	0.04886

$$\mathcal{C} = \{11111, 11122, 11133, 11144\}.$$

In this case, $\mathbf{E}(\tau) = 287.15789$, by Theorem 1 $\mathbf{E}(\mathcal{N}) = 4.42105$, and Theorem 2 gives us

$$\mathbf{E}(\alpha^{\mathcal{N}}) = \frac{12\alpha + 3\alpha^2 + 4\alpha^3}{64 - 36\alpha - 9\alpha^2}.$$

Taylor series of this rational function gives us needed probabilities. Results of 100,000 simulations together with the exact probabilities are presented in Table 1. The simulated expected number of occurrences is 4.42671. We can see that there is a good agreement between the exact and simulated probabilities. Thus we can claim that `splus` function `sample` passed this test of randomness.

References

Aki, S., Balakrishnan, N., Mohanty, S.G. (1996). Sooner and later waiting time problems and failure runs in higher order Markov dependent trials. *Annals of the Institute of Statistical Mathematics*, 48, 773–787.

Aki, S., Hirano, K. (1994). Distributions of numbers of failures and successes until the first consecutive k successes. *Annals of the Institute of Statistical Mathematics*, 46, 193–202.

Antzoulakos, D. (2001) Waiting times for patterns in a sequence of multistate trials. *Journal of Applied Probability*, 38, 508–518.

Balakrishnan, N. and Koutras, M.V. (2002). *Runs and Scans with Applications*. Wiley, Inc., New York.

Biggins, J. D. and Cannings, C. (1987). Markov renewal processes, counters and repeated sequences in Markov chains. *Advances in Applied Probability*, 19, 521–545.

Blom, G. and Thorburn, D. (1982). How many random digits are required until given sequences are obtained?. *Journal of Applied Probability*, 19, 518–531.

Breen, S., Waterman, M., Zhang, N. (1985). Renewal theory for several patterns. *Journal of Applied Probability*, 22, 228–234.

Chadjicostantinidis, S., Koutras, M. (2001). Distributions of the numbers of failures and successes in a waiting time problem. *Annals of the Institute of Statistical Mathematics*, 53, 576–598.

- Chrysaphinou, O., Papastavridis, S. (1990). The occurrence of a sequence of patterns in repeated dependent experiments. *Theory of Probability and Applications*, 35, 145–152.
- Feller, W. (1968). An introduction to probability theory and its applications, vol. 1, 3rd edn. Wiley, New York.
- Fu, J. C. (1986). Reliability of consecutive- k -out-of- n : F systems with $(k - 1)$ -step Markov dependence. *IEEE Transactions on Reliability*, R35, 602–606.
- Fu, J. (2001). Distribution of the scan statistics for a sequence of bistate trials. *Journal of Applied Probability*, 38, 908–916.
- Fu, J., Chang, Y. (2002). On probability generating functions for waiting time distribution of compound patterns in a sequence of multistate trials. *Journal of Applied Probability*, 39, 70–80.
- Fu, J. C., Koutras, M. V. (1994). Distribution theory of runs: A Markov chain approach. *Journal of The American Statistical Association*, 78, 168–175.
- Fu, J. C., Lou, W. Y. W. (2003). *Distribution Theory of Runs and Patterns*, World Scientific Publishing, Singapore.
- Gerber, H., Li, S.-Y.R. (1981). The occurrence of sequence patterns in repeated experiments and hitting times in a Markov chain. *Stochastic Processes and their Applications*, 11, 101–108.
- Glaz, J., Kulldorff, M., Pozdnyakov, V., Steele, J. M. (2006). Gambling teams and waiting times for patterns in two-state Markov chains. *Journal of Applied Probability*, 43, 127–140.
- Guibas, L., Odlyzko, A. (1981a). Periods of strings. *Journal of Combinatorial Theory. Series A*, 30, 19–42.
- Guibas, L., Odlyzko, A. (1981b). String overlaps, pattern matching and nontmsitive games. *Journal of Combinatorial Theory. Series A*, 30, 183–208.
- Han, Q., Hirano, K. (2003). Sooner and later waiting time problems for patterns in Markov dependent trials. *Journal of Applied Probability*, 40, 73–86.
- Hirano, K., Aki, S., Uchida, M. (1997). Distributions of numbers of success-runs until the first consecutive k successes in higher order Markov dependent trials. In: *Advances in combinatorial methods and applications to probability and statistics*, Balakrishnan N., (Ed.), Statistics for Industry and Technology, Birkhauser Boston, Boston.
- Hirano, K., Aki, S. (2003). Number of occurrences of subpattern until the first appearance of a pattern and geometric distribution. *Statistics and Probability Letters*, 65, 259–262.
- Li, S.-Y.R. (1980). A martingale approach to the study of occurrence of sequence patterns in repeated experiments. *Annals of Probability*, 8, 1171–1176.
- Pozdnyakov, V., Glaz, J., Kulldorff, M., Steele, J. M. (2005). A martingale approach to scan statistics. *Annals of the Institute of Statistical Mathematics*, 57, 21–37.
- Pozdnyakov, V., Kulldorff, M. (2006). Waiting times for patterns and a method of gambling teams. *American Mathematical Monthly*, 113, 134–143
- Robin, S., Daudin, J.-J. (2001). Exact distribution of the distances between any occurrence of a set of words. *Annals of the Institute of Statistical Mathematics*, 53, 895–905.
- Rukhin, A. L. (2001). Testing randomness: a suite of statistical procedures. *Theory of Probability and Applications*, 45, (2001), 111–132
- Rukhin, A. L. (2002). Distribution of the number of words with a prescribed frequency and tests of randomness. *Advances in Applied Probability*, 34, 775–797.
- Stefanov, V. T. (2000). On some waiting time problems. *Journal of Applied Probability*, 37, 756–764.
- Stefanov, V. T. (2003). The intersite distances between pattern occurrences in strings generated by general discrete- and continuous-time models: an algorithmic approach. *Journal of Applied Probability*, 40, 881–892.
- Stefanov, V. T., Pakes, A. G. (1997). Explicit distributional results in pattern formation. *Annals of Applied Probability*, 7, 666–678.
- Stefanov, V. T., Pakes, A. G. (1999). Explicit distributional results in pattern formation. II. *Australian & New Zealand Journal of Statistics*, 41, 79–90.
- Uchida, M. (1998). On generating functions of waiting time problems for sequence patterns of discrete random variables. *Annals of the Institute of Statistical Mathematics*, 50, 655–671.
- Williams, D. (1991). *Probability with martingales*. Cambridge University Press, Cambridge.