WAITING TIMES FOR PATTERNS AND A METHOD OF GAMBLING TEAMS

VLADIMIR POZDNYAKOV AND MARTIN KULLDORFF

1. Introduction.

We flip a fair coin five times. Which pattern is "more difficult" to get: $HHHHH$ or $HTHTH$? If we posed this question to the typical man on the street, the most likely answer would be: the first one. Of course, we know that this answer is not correct, for both patterns have the same probability of occurring, namely, 1/32. However, there is a sense in which a street-smart person *is*, in fact, correct. If we flip the coin without stopping, then the average waiting time until the first occurrence of the pattern $HHHHH$ is 62, whereas for the pattern $HTHTH$ it is 42. From this perspective, the pattern $HHHHH$ is indeed "more difficult" to get. Now, if we ask a person familiar with probability theory (but unfamiliar with this particular topic) to rank the average waiting times until the patterns $HHHHH$, $HHHHT$, $HHHTH$, and $HTHTH$ occur, then most likely the first pattern would get rank 1 (the longest average waiting time), the second—2, the third—3, and the last one—4 (the shortest average waiting time). This ranking is based on the "intuitive" idea that long runs of the same outcomes, such as $HHHH$ or $HHHHH$, require more time until they occur. In fact, the average waiting times are 62, 32, 34, and 42, respectively.

All the foregoing waiting times are easily and elegantly obtained by using martingale theory and the "optional stopping theorem," as shown in the classical paper by Li [16] and briefly described in section 2. Our focus in the present paper is the average time until we observe the first of several different patterns. Suppose, for instance, that Melanie flips a coin until she observes either $HHHTH$ or $HTHTH$, while Kyle flips another coin until he observes either $HHHHT$ or $HHHTH$. Since Kyle was assigned the two patterns with the shortest waiting times, 32 and 34 versus 34 and 42, one would expect him to have a shorter average waiting time. In fact, the averages are the same—22 for both Melanie and Kyle.

Let us present another counterintuitive fact. Consider the two patterns $HHHHT$ and $HHHTH$. What is the probability that in a stochastic sequence of heads and tails the pattern $HHHHT$ appears earlier than $HHHTH$? Since the average waiting times (32 and 34, respectively) are close to each other, one might guess that the probability would be reasonably close to 1/2. However, the exact answer is 2/3! As we will see, this probability is determined by the relationship between patterns rather than by their individual average waiting times.

Finally, consider two special patterns: a run of Hs of length r and a run of Ts of length ρ . The expected waiting time until the run of Hs is $2^{r+1} - 2$, while for the run of Ts it is $2^{\rho+1} - 2$. We can ask: what is the expected time until either of these two runs happens for the first time? Using results presented in this article

we can show that the answer is given by the following nice formula:

$$\left(\frac{1}{2^{r+1}-2}+\frac{1}{2^{\rho+1}-2}\right)^{-1}.$$

We now give a formal statement of the problem. Consider a discrete random variable Z that assumes values from a finite alphabet Σ, and let $\{Z,Z_k\}_{k\geq 1}$ be a sequence of independent, identically distributed random variables. Assume that we are given a finite collection of patterns (words) $\{A_j\}_{1\leq j\leq K}$ over Σ. We denote by τ_{A_j} the waiting time until A_j occurs for the first time as a run in the sequence $Z_1,Z_2,....$ The main objects of interest are the expected time of

$$\tau=\min\{\tau_{A_1},...,\tau_{A_K}\}, \tag{1}$$

and the probabilities $\pi_j=\mathbf{P}\big(\tau=\tau_{A_j}\big)$.

These two quantities were obtained for general patterns by Li [16] and Gerber and Li [13]. Their central observation was that information on the waiting time could be obtained from the values taken by a specially designed martingale at a relevant stopping time. However, they used the martingale technique only to evaluate the expected waiting time until one pattern occurs, whereas they treated the transition to the situation with many patterns by using a Markov chain embedding.

The goal of this article is to present an idea of multiple gambling teams that provides a simpler and more elegant solution than the method of Markov chain embedding does. The basic idea is to introduce a weighted sum of martingales (which is a itself martingale) and then choose the weights in such way that the expression of the stopped martingale is *the same* for all ending scenarios. After a routine application of the optional stopping theorem, the expression for the expected waiting time is found.

In section 3 we present a simple alternative proof of Li's result based on the method of gambling teams. In section 4 we extend the results to multiple series of independent identically distributed random variables and the occurrence of two-dimensional patterns that depend on more than one of these series. In the last section we present a list of relevant articles on the occurrence of patterns and provide some comments on further possible developments.

2. A Single Pattern.

We first consider the classical case in which we have only one stopping pattern $A=a_1...a_m$, where $\mathbf{P}(Z=a_i)>0$ $(i=1,...,m)$. What is the expected value $\mathbf{E}\tau_A$ of τ? The standard solution that was discovered by Li [16] goes as follows. Assume that a new gambler arrives just before each time $n=1,2,....$ He bets \$1 that

$$Z_n=a_1.$$

If he loses, he leaves the game. If he wins, he gets $1/\mathbf{P}(Z=a_1)$ dollars. Then he bets the whole amount on the event that

$$Z_{n+1}=a_2.$$

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

$$(1/\mathbf{P}(Z=a_1))\times(1/\mathbf{P}(Z=a_2))$$

dollars, and he bets his whole fortune on the next event

$$Z_{n+2}=a_3,$$

and so on through the m letters of the pattern A. If the gambler is lucky and finishes the pattern, he leaves the game with his winnings.

Let X_n be the net amount of money collected by the casino from all gamblers up to and including time n. Since the amount of the bets in round n depends only on the history up to time $n-1$ and the odds are fair for each gambler, the sequence $\{X_n\}_{n\geq 0}$ with $X_0 = 0$ is a *martingale*, i.e.,

$$\mathbf{E}(X_{n+1}|X_0, X_1, ..., X_n) = X_n, \text{ for } n = 0, 1, ... \tag{2}$$

The next step is to find the value of X_{τ_A}—the value of the martingale $\{X_n\}$ stopped at time τ_A. We first consider a simple example.

Example 1. Assume that we flip a fair coin until the first time τ_A when pattern $A = HTHT$ occurs. At this moment exactly τ_A gamblers have entered the game, each of them has paid a dollar, and almost all of them have lost their money. Only two gamblers have won: the one who entered the game right before time $\tau_A - 4$ and the one who started his betting at time $\tau_A - 2$. At time τ_A, the first gambler has \$16 and the second \$4. Thus, we find that $X_{\tau_A} = \tau_A - 16 - 4$.

Mimicking Li [16] we introduce the following measure of the amount of overlap between two patterns. Let $A = a_1...a_m$ and $B = b_1...b_k$ be patterns over Σ. For each pair (i, j) write

$$\delta_{i\,j} = \begin{cases} 1/\mathbf{P}(Z = b_j) & \text{if } 1 \leq i \leq m, 1 \leq j \leq k, \text{ and } a_i = b_j, \\ 0 & \text{otherwise.} \end{cases}$$

Then define

$$A * B = \delta_{1\,1}\delta_{2\,2}\cdots\delta_{m\,m} + \delta_{2\,1}\delta_{3\,2}\cdots\delta_{m\,m-1} + ... + \delta_{m\,1}. \tag{3}$$

It is easy to see that the casino's net gain by time τ_A is given by

$$X_{\tau_A} = \tau_A - A * A.$$

To verify this note that, as before, τ_A is the total amount of money spent by all the gamblers by time τ_A, while $A * A$ represents the winnings collected by the few lucky ones. For example, if $A = HTHT$, then $A * A$ is given by

$$A * A = \frac{1}{.5} \times \frac{1}{.5} \times \frac{1}{.5} \times \frac{1}{.5} + 0 + \frac{1}{.5} \times \frac{1}{.5} + 0 = 16 + 0 + 4 + 0 = 20.$$

One can show that $\mathbf{E}(\tau_A) < \infty$ and that the increments of X_n are uniformly bounded with probability 1. Indeed, first observe that the waiting time τ_A is bounded by mT, where

$$T = \min\{k : Z_{mk+1}...Z_{m(k+1)} = A\}.$$

Since T is a random variable with a geometric distribution (see Feller [6, p. 166]), $\mathbf{E}(\tau_A) < \infty$. Second, at any given time n there are at most m gamblers who bet and the total possible gain (and loss) of a gambler is bounded by $\prod_{j=1}^{m} 1/\mathbf{P}(Z = a_j)$. Therefore, the increments of X_n are also bounded. We can now invoke the optional stopping theorem (see Williams [23, p. 100]). This theorem simply asserts that if we are playing a fair game and if we are given the choice of when to stop the game, then we cannot beat the system to make a profit that on average exceeds 0. By an appeal to the optional stopping theorem we thus obtain

$$0 = \mathbf{E}(X_0) = \mathbf{E}(X_{\tau_A}) = \mathbf{E}(\tau_A) - A * A$$

and conclude that

$$\mathbf{E}(\tau_A) = A * A.$$

In Example 1, we have $0 = \mathbf{E}(X_{\tau_A}) = \mathbf{E}(\tau_A - 20)$, so $\mathbf{E}(\tau_A) = 20$.

Remark 1. According to Hall and Heyde [14] the concept of martingale (without the use of the word "martingale" though) was introduced in works of S. Bernstein and P. Lévy in the 1920s and 1930s. The martingale theory was considered as a natural extension of theory of processes with independent increments. Since the martingale condition (2) also implies that the processes does not change on average, martingales can be employed to model a fair game. But the martingale approach is more general than the one based on the random walk, because some forms of dependence are allowed. J. Ville brought in the name martingale in an article that was published in 1939. But it was J. Doob whose work in the 1940s and early 1950s added crucial developments that transformed martingale theory into an important area of probability theory.

A non-mathematical meaning of the word martingale is a strap of horse's harness that joins the noseband to the girth. This device is designed to prevent the horse from throwing its head back. Another, more relevant, meaning of the term martingale appears in a gambling context. As Hall and Heyde [14] note "the Oxford English Dictionary dates this usage back to 1815." It refers to a family of betting systems the simplest example of which is a doubling strategy. Consider a game where a gambler wins his stake if a fair coin comes up heads and he loses if it comes up tails. The gambler's betting system is as follows. The gambler starts the game with bet of 1 dollar, he doubles the next stake after a loss, and he leaves the game after the first win. More specifically, if he wins in the first round he leaves the game with a profit of 1 dollar. If he loses he bets 2 dollars in the second round. If he wins he gets 4 dollars that cover 1 dollar that was lost in the first round and the stake of 2 dollars in the second round, so he has a profit of 1 dollar again. If he is not lucky, and he loses again, he bets 4 dollars in the third round and so forth. It is clear that eventually the coin will come up heads and the gambler will increase his capital by 1 dollar. If V_n denotes the gambler's total gain through the round n, then one can show that the process $\{V_n\}_{n\geq 0}$ with $V_0 = 0$ forms a martingale in the mathematical sense, i.e., $\mathbf{E}(V_{n+1}|V_0, V_1, ..., V_n) = V_n$. But it is not a process with independent increments. In this case the size of the next bet really depends on observed values of the process V_n.

Perhaps, with some stretch of imagination, one can see a parallel between the martingale as a piece of a horse's harness that pulls the horse's head down and the martingale as a gambling strategy that via the bet doubling recovers losses in previous rounds. Before we go further let us note also that in spite of the simplicity of the described betting system it is physically impossible to implement. The problem is that this strategy requires access to an infinite capital and a gambler should be ready to play the game for an unbounded time. In fact, this process is a standard example of a bad martingale for which the optional stopping theorem does not hold.

3. Multiple Patterns: A Method of Gambling Teams.

Li [16] extended the result that we have just described to determine the expected time until any one of a finite number of patterns occurs. This was done by developing a more general and extensive formula for the expected waiting time for each of

the patterns A_1, A_2, A_3, ..., conditioned on each one of the other patterns being the starting pattern of the sequence of random letters. As we will see, this elaborate step can in fact be skipped.

Consider the situation when we have a collection of K patterns. Without loss of generality we require that no pattern be a subpattern of another. Assume now that we have K teams of betters and that the first team bets on the pattern A_1, the second team on A_2, and so on. Let y_j be the initial amount of money with which each of the gamblers from the jth team starts his or her betting. As before let X_n be the net gain of the casino at time n. Since any weighted sum of martingales is a martingale the stochastic sequence $\{X_n\}$ is a martingale. The stopped martingale X_τ is given by

$$X_\tau = \begin{cases} (y_1 + ... + y_K)\tau - (A_1 * A_1 \times y_1 + ... + A_1 * A_K \times y_K) & \text{if } \tau = \tau_{A_1}, \\ (y_1 + ... + y_K)\tau - (A_2 * A_1 \times y_1 + ... + A_2 * A_K \times y_K) & \text{if } \tau = \tau_{A_2}, \\ \qquad\qquad ... \qquad\qquad & \qquad ... \\ (y_1 + ... + y_K)\tau - (A_K * A_1 \times y_1 + ... + A_K * A_K \times y_K) & \text{if } \tau = \tau_{A_K}. \end{cases} \tag{4}$$

By the same arguments as earlier we find that

$$0 = \mathbf{E}(X_\tau) = (y_1 + ... + y_K)\mathbf{E}(\tau) - \Pi^\top MY, \tag{5}$$

where $\Pi = (\pi_1, ..., \pi_K)^\top$ ($\pi_j = \mathbf{P}(\tau = \tau_{A_j})$), $Y = (y_1, ..., y_n)^\top$, and

$$M = \begin{bmatrix} A_1 * A_1 & A_1 * A_2 & ... & A_1 * A_K \\ A_2 * A_1 & A_2 * A_2 & ... & A_2 * A_K \\ ... & ... & ... & ... \\ A_K * A_1 & A_K * A_2 & ... & A_K * A_K \end{bmatrix}. \tag{6}$$

The most important benefit of introducing several gambling teams is that we have K parameters (the weights $y_1, ..., y_K$) that are under *our* control. A good choice for those parameters helps us achieve our goal of finding $\mathbf{E}(\tau)$ and the probabilities π_j ($j = 1, 2, ..., K$). In particular, to prove Proposition 1 (see [16, p. 1174]) we choose the vector of weights (or initial bets) in such a way that the expression for the stopped martingale X_τ does not depend on how the game is stopped.

Proposition 1. *If there exist a solution* $Y^* = (y_1^*, ..., y_K^*)^\top$ *to the linear system* $MY = \mathbf{1}$, *where* $\mathbf{1} = (1, ..., 1)^\top$, *then*

$$\mathbf{E}(\tau) = \frac{1}{y_1^* + \cdots + y_K^*}.$$

The proof based on multiple gambling teams is very simple:

Proof. Since $\Pi^\top MY^* = \sum_{j=1}^K \pi_j = 1$, we infer from (5) that

$$0 = (y_1^* + ... + y_K^*)\mathbf{E}(\tau) - 1.$$

□

The next proposition shows how a different choice of bets can identify the probabilities $\pi_1, ..., \pi_K$. Here we introduce K different vectors of bets. Each vector gives us a linear equation for $(\pi_1, ..., \pi_K)$. Note that in each case only one initial bet is different from zero, so most of the teams only *pretend* that they are betting. We can call them *wise gamblers*.

Proposition 2. *The probability vector* $\Pi = (\pi_1, ..., \pi_K)^\top$ *satisfies the equation* $M^\top \Pi = \mathbf{E}(\tau)\mathbf{1}$.

Again, with the method of multiple gambling teams the proof is extremely simple:

Proof. We consider the following vector of initial bets

$$Y^j = (0, ..., \underbrace{1}_{j\text{th position}}, ..., 0)^\top.$$

From (5) we learn that

$$0 = \mathbf{E}(\tau) - \Pi^\top M Y^j = \mathbf{E}(\tau) - (A_1 * A_j, ..., A_K * A_j)^\top \Pi.$$

Since this equation holds for $j = 1, 2, ..., K$, the proposition follows. □

Remark 2. According to Li [16, p. 1175] the matrix M in (6) is nonsingular provided that no pattern contains another as a subpattern.

4. Multiple Sequences with Multiple Patterns.

To demonstrate some applications of Propositions 1 and 2 we look at multiple sequences of simultaneously generated random letters. For example, rather than a single gambling table we can consider a room with multiple tables operating at the same time and at the same speed, so that in round n a random letter is generated at the same time at each of the tables. The problem is now to find the expected time until one of a predetermined set of patterns occurs at the first table and/or one of another set of patterns occurs on the second table, and so forth for an arbitrary number of tables. For example, in a sequences of coin tosses, we might require three straight heads at table one at the same time that there are at least two tails out of three at table two.

At first glance, this problem seems conceptually much more complex than having only one sequence to deal with, but as we will see it is only computationally more complex. To illustrate this point we use two simple examples, one involving two sequences without pattern dependence and the other involving three sequences with pattern dependence.

Example 2. Let $\{Z_k^{(1)}\}_{k\geq 1}$ and $\{Z_k^{(2)}\}_{k\geq 1}$ be two independent sequences of independent, identically distributed random variables (say, referring to bets at tables one and two, respectively), where

$$\mathbf{P}(Z_k^{(i)} = A) = p_i, \quad \mathbf{P}(Z_k^{(i)} = B) = q_i, \quad p_i + q_i = 1 \quad (i = 1, 2).$$

Let $A_1 = AA$ and $A_2 = AAA$, and set $\tau = \min\{\tau_{A_1}, \tau_{A_2}\}$, where τ_{A_i} is the waiting time for the occurrence of the pattern A_i at table i $(i = 1, 2)$. What is $\mathbf{E}(\tau)$?

The expected time can be found via a straightforward application of Proposition 1. Indeed, we can treat each pair $(Z_n^{(1)}, Z_n^{(2)})^\top$ as a random letter of a new four-letter alphabet:

$$\left[\begin{array}{c} A \\ A \end{array}\right] \quad \left[\begin{array}{c} A \\ B \end{array}\right] \quad \left[\begin{array}{c} B \\ A \end{array}\right] \quad \left[\begin{array}{c} B \\ B \end{array}\right].$$

In this new alphabet the waiting time τ is the first time when we observe one of the following blocks (or patterns from the new alphabet):
(1) when $\tau_{A_1} < \tau_{A_2}$,

$$\begin{array}{cccc} \text{BAA} & \text{AA} & \text{AA} & \text{AA} \\ \text{BAA} & \text{AB} & \text{BA} & \text{BB} \end{array};$$

(2) when $\tau_{A_1} = \tau_{A_2}$,

$$\begin{matrix} \text{BAA} \\ \text{AAA} \end{matrix};$$

(3) when $\tau_{A_2} < \tau_{A_1}$,

$$\begin{matrix} \text{BBB} \\ \text{AAA} \end{matrix} \quad \begin{matrix} \text{BBA} \\ \text{AAA} \end{matrix} \quad \begin{matrix} \text{BAB} \\ \text{AAA} \end{matrix} \quad \begin{matrix} \text{ABB} \\ \text{AAA} \end{matrix} \quad \begin{matrix} \text{ABA} \\ \text{AAA} \end{matrix}.$$

For instance, if $p_1 = .1$ and $p_2 = .2$, then the matrix M (rounded to whole numbers) is given by

$$M = \begin{bmatrix} 3472 & 50 & 0 & 0 & 0 & 0 & 0 & 0 & 50 & 50 \\ 0 & 625 & 13 & 13 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 50 & 625 & 0 & 0 & 0 & 0 & 0 & 50 & 50 \\ 0 & 0 & 13 & 169 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 50 & 0 & 0 & 13889 & 0 & 0 & 0 & 50 & 50 \\ 0 & 0 & 0 & 0 & 6 & 208 & 36 & 6 & 0 & 0 \\ 0 & 50 & 0 & 0 & 278 & 0 & 1543 & 278 & 50 & 50 \\ 0 & 0 & 0 & 0 & 6 & 6 & 6 & 1549 & 278 & 278 \\ 0 & 0 & 0 & 0 & 6 & 36 & 36 & 6 & 1543 & 0 \\ 0 & 50 & 0 & 0 & 278 & 0 & 0 & 278 & 50 & 13939 \end{bmatrix}.$$

Applying Proposition 1 we find that $\mathbf{E}(\tau) = 65.2633$ (note that $\mathbf{E}(\tau_{A_1}) = 110$ and $\mathbf{E}(\tau_{A_1}) = 155$). By Proposition 2 we have $\mathbf{P}(\tau_{A_1} < \tau_{A_2}) = 0.5857$, $\mathbf{P}(\tau_{A_2} < \tau_{A_1}) = 0.4106$, and $\mathbf{P}(\tau_{A_2} = \tau_{A_1}) = 0.0038$.

Remark 3. Because each of the patterns A_1 and A_2 in Example 2 depends on a different single sequence, there is an alternative way of computing the expected waiting time τ. Note first that from the independence of τ_{A_1} and τ_{A_2} we get

$$\begin{aligned} \mathbf{E}(\tau) &= \sum_{n=1}^{\infty} \mathbf{P}(\tau \geq n) \\ &= \sum_{n=1}^{\infty} \mathbf{P}(\tau_{A_1} \geq n \text{ and } \tau_{A_2} \geq n) \\ &= \sum_{n=1}^{\infty} \mathbf{P}(\tau_{A_1} \geq n)\mathbf{P}(\tau_{A_2} \geq n). \end{aligned}$$

The probabilities $\mathbf{P}(\tau_{A_i} \geq n)$ can now be computed by the recurrent event theory method (see Feller [6]). However, since this method requires knowing the entire distributions of τ_{A_1} and τ_{A_2} it is much more complicated. Furthermore, it requires that the two sequences be independent, and if we assume some dependence structure for pairs $(Z_n^{(1)}, Z_n^{(2)})$, then it will not work. For instance, assume that the occurrence of A in the first sequence implies a higher probability of observing A in the corresponding column of the second sequence (i.e., $\mathbf{P}(Z_n^{(2)} = A | Z_n^{(1)} = A) = p_{12} > p_2$). The relationships between τ_{A_1} and τ_{A_2} then become too complex, and the generating function method just mentioned fails. By contrast, the approach via Proposition 1 still works. The only change is a different distribution for the letters of the new alphabet.

Example 3. Suppose that we have three independent tables, each with a sequence of independent, identically distributed random variables $\{Z_k^{(i)}\}_{k\geq 1}$ satisfying

$$\mathbf{P}(Z_k^{(i)} = A) = p_i, \quad \mathbf{P}(Z_k^{(i)} = B) = q_i, \quad p_i + q_i = 1 \quad (i = 1, 2, 3).$$

Let τ be the waiting time for the 2-by-2 block:

$$\begin{matrix} \text{A} & \text{A} \\ \text{A} & \text{A} \end{matrix}.$$

For instance, if the realization of $\{Z_k^{(i)}\}_{k\geq 1}$ $(i = 1, 2, 3)$ produced the following three sequences:

$$\begin{matrix} \text{A} & \text{B} & \text{A} & \text{A} & \text{B} & \text{A} & \text{B} & \text{A} & \text{B} & ..., \\ \text{A} & \text{B} & \text{A} & \text{A} & \text{A} & \text{B} & \text{A} & \text{A} & \text{B} & ..., \\ \text{A} & \text{B} & \text{B} & \text{B} & \text{B} & \text{A} & \text{A} & \text{A} & \text{A} & ..., \end{matrix}$$

then $\tau = 4$.

What is $\mathbf{E}(\tau)$? Again we first introduce a new (eight-letter) alphabet:

$$\begin{bmatrix} A \\ A \\ A \end{bmatrix} \quad \begin{bmatrix} A \\ A \\ B \end{bmatrix} \quad \begin{bmatrix} A \\ B \\ A \end{bmatrix} \quad \begin{bmatrix} B \\ A \\ A \end{bmatrix} \begin{bmatrix} A \\ B \\ B \end{bmatrix} \quad \begin{bmatrix} B \\ A \\ B \end{bmatrix} \quad \begin{bmatrix} B \\ B \\ A \end{bmatrix} \quad \begin{bmatrix} B \\ B \\ B \end{bmatrix}.$$

In terms of the new alphabet we wait for one of the following blocks (words):

$$\begin{matrix} \text{AA} & \text{AA} & \text{AA} & \text{AA} & \text{BB} & \text{BA} & \text{AB} \\ \text{AA} & \text{AA} & \text{AA} & \text{AA} & \text{AA} & \text{AA} & \text{AA} \\ \text{AB} & \text{BA} & \text{BB} & \text{AA} & \text{AA} & \text{AA} & \text{AA} \end{matrix}.$$

If $p_i = 1/2$ for each i, the matrix M for these words is given by

$$M = \begin{bmatrix} 64 & 8 & 8 & 0 & 0 & 0 & 0 \\ 8 & 64 & 0 & 8 & 0 & 0 & 8 \\ 0 & 8 & 72 & 0 & 0 & 0 & 0 \\ 8 & 0 & 0 & 72 & 0 & 0 & 8 \\ 0 & 0 & 0 & 0 & 72 & 8 & 0 \\ 8 & 0 & 0 & 8 & 0 & 64 & 8 \\ 0 & 0 & 0 & 0 & 8 & 8 & 64 \end{bmatrix},$$

and $\mathbf{E}(\tau) = 11.9245$.

If $p_1 = p_2 = 1/2$ and $p_3 = 1/4$, then M is given by

$$M = \begin{bmatrix} 85\frac{1}{3} & 5\frac{1}{3} & 5\frac{1}{3} & 0 & 0 & 0 & 0 \\ 16 & 85\frac{1}{3} & 0 & 16 & 0 & 0 & 16 \\ 0 & 5\frac{1}{3} & 33\frac{7}{9} & 0 & 0 & 0 & 0 \\ 16 & 0 & 0 & 272 & 0 & 0 & 16 \\ 0 & 0 & 0 & 0 & 272 & 16 & 0 \\ 16 & 0 & 0 & 16 & 0 & 256 & 16 \\ 0 & 0 & 0 & 0 & 16 & 16 & 256 \end{bmatrix},$$

and $\mathbf{E}(\tau) = 16.8846$. The probability that we would stop because we had the block in the top two sequences is 0.8324.

5. Concluding Remarks.

The waiting time τ is a key to many real-world questions in various fields: quality control, hypothesis testing, molecular biology, and DNA-sequencing. Because of its practical importance, the occurrence of patterns has been studied extensively by many different techniques. The case of independent trials was considered by Feller [6], who first systematically employed the recurrent event theory. Later combinatorial methods were introduced by Guibas and Odlyzko [11] and [12]. The problem has been studied by probabilistic techniques by Biggins and Cannings [2], Blom and Thorburn [3], Breen et al. [4], Chrysaphinou and Papastavridis [5], Han and Hirano [15], Robin and Daudin [18], Stefanov [20], and Uchida [22]. One of the more general techniques is the Markov chain embedding method introduced by Fu [7], which has been further developed by Antzoulakos [1], Fu [8], Fu and Chang [9], Fu and Koutras [10], Stefanov [19], and Stefanov and Pakes [21].

The method presented here—the method of gambling teams—may simply look like a more elegant way to solve a classical problem. However, it is more important than that. With suitable modifications, the method of gambling teams can also be employed to compute higher moments and the generating function of the waiting time, which in turn can be used to calculate probabilities for scan statistics that were previously too difficult to obtain analytically (see Pozdnyakov et al. [17]). In a forthcoming paper, we show that it can also be used to calculate waiting times in Markov chains.

References

[1] D. Antzoulakos, Waiting times for patterns in a sequence of multistate trials, *J. Appl. Probab.* **38** (2001) 508-518.

[2] J. D. Biggins and C. Cannings, Markov renewal processes, counters and repeated sequences in Markov chains, *Adv. Appl. Probab.* **19** (1987) 521-545.

[3] G. Blom and D. Thorburn, How many random digits are required until given sequences are obtained?, *J. Appl. Probab.* **19** (1982) 518-531.

[4] S. Breen, M. Waterman, and N. Zhang, Renewal theory for several patterns, *J. Appl. Probab.* **22** (1985) 228-234.

[5] O. Chrysaphinou and S. Papastavridis The occurrence of a sequence of patterns in repeated dependent experiments, *Theory Probab. Appl.* **35** (1990) 145-152.

[6] W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. 1, 3rd ed., Wiley, New York, 1968.

[7] J. C. Fu, Reliability of consecutive-k-out-of-n: F systems with $(k-1)$-step Markov dependence, *IEEE Trans. Reliability* **R35** (1986) 602-606.

[8] —, Distribution of the scan statistics for a sequence of bistate trials, *J. Appl. Probab.* **38** J. (2001) 908-916.

[9] J. C. Fu and Y. Chang, On probability generating functions for waiting time distribution of compound patterns in a sequence of multistate trials, *J. Appl. Probab.* **39** (2002) 70-80.

[10] J. C. Fu and M. V. Koutras, Distribution theory of runs: A Markov chain approach, *J. Amer. Statist. Assoc.* **78** (1994) 168-175.

[11] L. Guibas and A. Odlyzko, Periods of strings, *J. Combin. Theory Ser. A* **30** (1981) 19-42.

[12] —, String overlaps, pattern matching and nontmsitive games, *J. Combin. Theory Ser. A*, **30** (1981) 183-208.

[13] H. Gerber and S. Li, The occurrence of sequence patterns in repeated experiments and hitting times in a Markov chain, *Stochastic Process. Appl.* **11** (1981) 101-108.

[14] P. Hall and C.C. Heyde, *Martingale Limit Theory and Its Applications*, Academic Press, New York, 1960.

[15] Q. Han and K. Hirano, Sooner and later waiting time problems for patterns in Markov dependent trials, *J. Appl. Probab.* **40** (2003) 73-86.

[16] S. Li, A martingale approach to the study of occurrence of sequence patterns in repeated experiments, *Ann. Probab.* **8** (1980) 1171-1176.
[17] V. Pozdnyakov, J. Glaz, M. Kulldorff, and J.M. Steele, A martingale approach to scan statistics, *Ann. Inst. Statist. Math.* (to appear).
[18] S. Robin and J.-J. Daudin, Exact distribution of word occurrences in a random sequence of letters, *J. Appl. Probab.* **36** (1999) 179-193.
[19] V. T. Stefanov, On some waiting time problems, *J. Appl. Probab.* **37** (2000) 756-764.
[20] —, The intersite distances between pattern occurrences in strings generated by general discrete- and continuous-time models: an algorithmic approach, *J. Appl. Probab.* **40** (2003) 881-892.
[21] V.T. Stefanov and A. G. Pakes, Explicit distributional results in pattern formation, *Ann. Appl. Probab.* **7** (1997) 666-678.
[22] M. Uchida, On generating functions of waiting time problems for sequence patterns of discrete random variables, *Ann. Inst. Statist. Math.* **50** (1998) 655-671.
[23] D. Williams, *Probability with Martingales*, Cambridge University Press, Cambridge, 1991.

VLADIMIR POZDNYAKOV received his Ph.D. in 2001 from the University of Pennsylvania. Since 2001 he has been at the University of Connecticut, where he is currently assistant professor of statistics. His research interests are limit theorems, mathematical finance, sequential analysis, and the occurrence of patterns.
Department of Statistics,University of Connecticut, 215 Glenbrook Road, U-4120, Storrs, CT 06269-4120
vladimir.pozdnyakov@uconn.edu

MARTIN KULLDORFF is associate professor of ambulatory care and prevention at Harvard Medical School and Harvard Pilgrim Health Care. He studied mathematics and statistics at Umeå University, graduating in 1984, and received his Ph.D. in operations research from Cornell University in 1989. He is interested in the probabilistic and algorithmic properties of different types of scan statistics and their practical use for geographical disease surveillance, drug safety monitoring, and the early detection of disease outbreaks.
Department of Ambulatory Care and Prevention, Harvard Medical School and Harvard Pilgrim Health Care, 133 Brookline Avenue, Boston, MA 02215-3920
martin_kulldorff@hms.harvard.edu