Let $Z$ be a discrete random variable. Let $\Sigma$ the set of all values of $Z$. Let \( \{Z_n\}_{n=1}^{\infty} \) be a sequence of iid random variables with distribution as $Z$.

Let $A$ and $B$ to finite sequences over $\Sigma$, such that $B$ is not a connected subsequence of $A$. Define

\[
N_B = \min\{k : B \text{ is a connected subsequence of } (Z_1, \cdots, Z_k)\},
\]

Then $N_B$ is a stopping times. The goal is to find $EN_B$ and $EN_{AB}$ for a given sequence $A$.

Notice that since $B$ is finite, and every finite sequence has a positive probability, then $N_B \leq G$, for some geometric random variable. ($G$ for example could be the smallest $n$, such that $Z_{nk}, Z_{nk+1}, \cdots, Z_{(n+1)k-1}$ matches the pattern we want.

**Example 1** Let a die, which shows $x, y$ and $z$ with respective probabilities $\frac{1}{2}, \frac{1}{3},$ and $\frac{1}{6}$ be rolled repeatedly. Let $B$ be the sequence $(x, z, x)$. Compute $N_B$.

**Solution:**

We introduce the following fair game. A gambler bets 1 dollar. At the first roll if $x$ appears, he receives 2 dollars (including his bet) and must parlay the 2 dollars on the occurrence of $z$ at the second roll. If he wins, he receives 12 dollars and must parlay the whole amount of 12 dollars on the occurrence of $x$ at the third roll. If he wins three times in a row, he receives 24 dollars and the game is over.

Now suppose that, before each roll, a new gambler joins the game and starts betting 1 dollar on the same sequence $B$. We continue the game until the first person wins. For example if the rolls turn out to be $(y, x, x, z, y, x, x, z, x)$, then we have 9 participants. The gambler 7 wins 24, and the gambler 9 wins 2 dollars. Since this game is a martingale, then the sequence of the participants gain $\{X_{N_B} \wedge n\}$ forms a martingale. Therefore

\[
EX_{N_B} \wedge n = 0
\]

We will argue that

\[
EX_N = \lim_{N \to \infty} EX_{N \wedge n} = 0
\]

Equation(1) will let us to calculate $EN$. This is because of the following relation between the net gain $X$ of the participants and $N_B$. Given $N_B = n$, then all...
n participants except for n – 2th who wins 24 and n th who wins 2 dollars, the rest of n participants lose 1 dollar, therefore

\[ 0 = EX = E(E(X|N)) = E[24 + 2 - (N_B)] \]

Therefore

\[ EN = 26 \]

In General we can design a fair game by defining the random net gain of the \( j \) th gambler at the time \( k \) as

\[ M^{(j)}_k = \begin{cases} 0 & \text{if } k < j \\ \frac{1}{P(Z=b_1)\cdots P(Z=b_{k-j+1})} - 1 & \text{if } k - j + 1 \text{ terms in } \omega_k \text{ are identical with } b_1, \ldots, b_{k-j+1}, \\ -1 & \text{otherwise} \end{cases} \]

where \( \omega_k \) denote the sequence \((Z_1, \ldots, Z_k)\).

Therefore if add the net gain for the all participants we get

\[ \sum_{j=1}^{\infty} M^{(j)}_k = \sum_{j=1}^{k} M^{(j)}_k = \omega_k * B - k \]

(2)

where \( \omega_k * B \) is define as follows

**Definition 1** Let \( A = (a_1, \ldots, a_m) \) and \( B = (b_1, \ldots, b_n) \) be a sequence over \( \Sigma \).

For every pair \((i, j)\) of integers, write

\[ \delta_{ij} = \begin{cases} P(Z = b_j)^{-1} & \text{if } 1 \leq i \leq m, \ 0 \leq j \leq n, \text{ and } a_i = b_j \\ 0 & \text{otherwise} \end{cases} \]

and define

\[ A * B = \delta_{11}\delta_{22}\cdots\delta_{mm} + \delta_{21}\delta_{32}\cdots\delta_{m,m-1} + \delta_{m1} \]

**Lemma 1** Given a starting sequence \( B \), the expected waiting time for a sequence \( B \) is \( EN_B = B * B \).

**Proof.** Define \( X_k = \omega_k * B - k \). Then by (2)

\[ \{X_k \wedge N_B\} \]

is a martingale. Let \( k \to \infty \), to get

\[ X_{N_B} = B * B - N_B \]

Since \( EX_{N_B} \leq B * B - EN_B < \infty \), and on the set \( \{N_B > k\} \) we have \( |X_k| \leq B * B + N_B \) Then \( EN_B = B * B \).
Let $A_1, \ldots, A_n$ be sequences over $\Sigma$. For each $i$, we want to calculate the probability that $A_i$ precedes all the remaining $n-1$ sequences in a realization of the process $Z_1, Z_2, \ldots$. Naturally we assume that none of the sequences contain any other as a connected subsequence. Write $N_i$ for $N_{A_i}$. Let $N$ be the minimum among $N_1, \ldots, N_n$. We want to compute $P(N = N_j)$ for each $j$.

**Theorem 1** Let $Z, Z_1, Z_2, \ldots$ be discrete iid random variables and $A_1, \ldots, A_n$ be finite sequences of possible values of $Z$ not containing one another. Let $A$ be another such sequence not containing any $A_i$. Let $p_i$ be the probability that $A_i$ precedes the remaining $n-1$ sequences in a realization of the process $Z_1, Z_2, \ldots$

The for every $i$,

$$\sum_{j=1}^n p_j A_j \ast A_i = EN,$$

where $N$ is the stopping time when any $A_j$’s occurs.