

Lecture-09: Limit Theorems

1 Growth of renewal counting processes

Consider a renewal sequence $S : \Omega \rightarrow \mathbb{R}_+^{\mathbb{N}}$ with *i.i.d.* inter renewal time sequence $X : \Omega \rightarrow \mathbb{R}_+^{\mathbb{N}}$ such that $S_n \triangleq \sum_{i=1}^n X_i$ for each $n \in \mathbb{N}$. The associated renewal counting process $N : \Omega \rightarrow \mathbb{Z}_+^{\mathbb{R}_+}$ is defined as $N_t \triangleq \sum_{n \in \mathbb{N}} \mathbb{1}_{\{S_n \leq t\}}$ and renewal function $m \in \mathbb{R}_+^{\mathbb{R}_+}$ is defined as $m_t \triangleq \mathbb{E}N_t$ for all $t \in \mathbb{R}_+$. Let $\mathcal{F}_\bullet \triangleq (\mathcal{F}_n : n \in \mathbb{N})$ be the natural filtration for process X , such that $\mathcal{F}_n \triangleq \sigma(X_1, \dots, X_n)$ for each $n \in \mathbb{N}$.

Lemma 1.1. *If X_1 has finite mean, then $N_\infty \triangleq \lim_{t \rightarrow \infty} N_t$ is almost surely infinite.*

Proof. Since $\mathbb{E}[X_n] < \infty$ and $X_n \geq 0$, it follows that $P\{X_n = \infty\} = 0$. Further, we observe that

$$P\{N_\infty < \infty\} = P\left(\bigcup_{n \in \mathbb{N}} \{N_\infty < n\}\right) = P\left(\bigcup_{n \in \mathbb{N}} \{S_n = \infty\}\right) = P\left(\bigcup_{n \in \mathbb{N}} \{X_n = \infty\}\right) \leq \sum_{n \in \mathbb{N}} P\{X_n = \infty\} = 0.$$

□

Corollary 1.2. *If the first renewal instant and subsequent inter renewal times for a delayed renewal processes N^D have finite means, then $P\{\lim_{t \rightarrow \infty} N_t^D = \infty\} = 1$.*

Proof. For each $n \in \mathbb{N}$, the inter renewal time X_n is positive and has finite mean, and hence is finite almost surely. Therefore, $P\{N_\infty < \infty\} \leq \sum_{n \in \mathbb{N}} P\{X_n = \infty\} = 0$. □

1.1 Growth of counting process

We observed that the number of renewals N_t increases to infinity with the length of the duration t . We will show that the growth of N_t is asymptotically linear with time t , and we will find this coefficient of linear growth of N_t with time t .

Theorem 1.3 (Strong law for renewal process). $\lim_{t \rightarrow \infty} N_t/t = 1/\mathbb{E}X_1$ almost surely.

Proof. We first consider the case when $\mathbb{E}X_1 = \infty$. It follows that $\alpha \triangleq P\{X_1 < \infty\} < 1$ and hence $P\{N_\infty = n\} = \alpha^n(1 - \alpha)$ for all $n \in \mathbb{Z}_+$. It follows that $\bigcup_{n \in \mathbb{Z}_+} \{N_\infty = n\} = \bigcup_{n \in \mathbb{Z}_+} \{N_\infty \leq n\}$ is an almost sure event, and hence $\lim_{t \rightarrow \infty} N_t/t = 0$ almost surely.

Next, we assume that $0 \leq \mathbb{E}X_1 < \infty$. For this case, $\lim_{t \rightarrow \infty} N_t = \infty$ almost surely from Lemma 1.1. We note that $S_{N_t} \leq t < S_{N_t+1}$, and dividing by N_t , we get

$$\sum_{n \in \mathbb{N}} \frac{S_n}{n} \mathbb{1}_{\{N_t=n\}} = \frac{S_{N_t}}{N_t} \leq \frac{t}{N_t} < \frac{S_{N_t+1}}{N_t} = \sum_{n \in \mathbb{N}} \frac{S_{n+1}}{n} \mathbb{1}_{\{N_t=n\}}.$$

Since $\lim_{t \rightarrow \infty} N_t = \infty$ almost surely, we have $\lim_{t \rightarrow \infty} \frac{S_{N_t}}{N_t} = \lim_{n \rightarrow \infty} \frac{S_n}{n}$ and $\lim_{t \rightarrow \infty} \frac{S_{N_t+1}}{N_t} = \lim_{n \rightarrow \infty} \frac{S_{n+1}}{n}$ almost surely. The result follows from L^1 strong law of large numbers. □

Corollary 1.4. *For a delayed renewal process with finite $\mathbb{E}X_2$, we have $\lim_{t \rightarrow \infty} N_t^D/t = 1/\mathbb{E}X_2$.*

Proof. From L^1 strong law of large numbers, we observe that $\lim_{n \in \mathbb{N}} S_n/n = \mathbb{E}X_2$, and the result follows. □

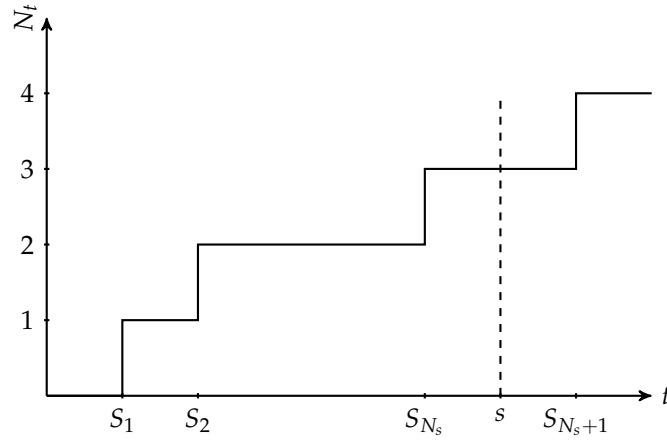


Figure 1: Time of last renewal

Example 1.5. Suppose, you are in a casino with infinitely many games. We assume that $X : \Omega \rightarrow [0, 1]^{\mathbb{N}}$ is an *i.i.d.* uniform sequence where X_i is the random probability of win in the game $i \in \mathbb{N}$. One can continue to play a game or switch to another one. We are interested in a strategy that maximizes the long-run proportion of wins. Let $N(n)$ denote the number of losses in n plays. Then the fraction of wins $P_W(n)$ is given by $P_W(n) = \frac{n - N(n)}{n}$. We pick a strategy where any game is selected to play, and continue to be played till the first loss. We show that $\lim_{n \rightarrow \infty} P_W(n) = 1$ for this proposed strategy. Let T_i be the number of times a game i is played. We observe that the conditional probability mass function for the number of plays for each game i is geometrically distributed as

$$\mathbb{E}[\mathbb{1}_{\{T_i=k\}} \mid \sigma(X_i)] = X_i^{k-1}(1 - X_i), \quad k \in \mathbb{N}.$$

Hence, it follows that T_i are *i.i.d.* random variables with mean $\mathbb{E}T_i = \mathbb{E}[\mathbb{E}[T_i \mid X_i]] = \mathbb{E}\left[\frac{1}{1-X_i}\right] = \infty$. It follows that each loss is a renewal event, and from the strong law of renewal process, we obtain

$$\lim_{n \rightarrow \infty} \frac{N(n)}{n} = \frac{1}{\mathbb{E}[\text{Time till first loss}]} = \frac{1}{\mathbb{E}T_i} = 0.$$

1.2 Growth of renewal function

Basic renewal theorem implies N_t/t converges to $1/\mathbb{E}X_1$ almost surely. We are next interested in convergence of the ratio m_t/t . Note that this is not obvious, since almost sure convergence doesn't imply convergence in mean. To illustrate this, we have the following example.

Example 1.6. Consider a Bernoulli random sequence $X : \Omega \rightarrow \{0, 1\}^{\mathbb{N}}$ with probability $P\{X_n = 1\} = \frac{1}{n}$, and another random sequence $Y : \Omega \rightarrow \mathbb{Z}_+^{\mathbb{N}}$ defined as $Y_n \triangleq nX_n$ for $n \in \mathbb{N}$. Then, $P\{Y_n = 0\} = 1 - \frac{1}{n}$. That is $Y_n \rightarrow 0$ a.s. However, $\mathbb{E}[Y_n] = 1$ for all $n \in \mathbb{N}$. So $\mathbb{E}[Y_n] \rightarrow 1$.

Even though, basic renewal theorem does **NOT** imply it, we show that m_t/t converging to $1/\mathbb{E}X_1$.

Proposition 1.7 (Wald's Lemma for renewal process). *If $\mathbb{E}X_1 \in (0, \infty)$, then $N_t + 1$ is a stopping time adapted to \mathcal{F}_\bullet , and $\mathbb{E}\sum_{i=1}^{N_t+1} X_i = (1 + m_t)\mathbb{E}X_1$.*

Proof. Recall that N_t is almost surely finite if $\mathbb{E}X_1 > 0$. We fix $n \in \mathbb{N}$, and observe that

$$\{N_t + 1 = n\} = \{S_{n-1} \leq t < S_n\} = \left\{ \sum_{i=1}^{n-1} X_i \leq t < \sum_{i=1}^{n-1} X_i + X_n \right\} \in \sigma(X_1, \dots, X_n) = \mathcal{F}_n.$$

Thus $N_t + 1$ is a stopping time adapted to \mathcal{F}_\bullet , and the result follows from Wald's Lemma. \square

Theorem 1.8 (Elementary renewal theorem). If $\mathbb{E}X_1 < \infty$, then $\lim_{t \rightarrow \infty} m_t/t = 1/\mathbb{E}X_1$.

Proof. By the assumption, we have mean $\mu \triangleq \mathbb{E}X_1 < \infty$.

- (a) Taking expectations on both sides of $S_{N_t+1} > t$ and using Wald's Lemma for renewal processes in Proposition 1.7, we have $\mu(m_t + 1) > t$. Dividing both sides by μt and taking \liminf on both sides, we get $\liminf_{t \rightarrow \infty} m_t/t \geq 1/\mu$.
- (b) We employ a truncated random variable argument to show the reverse inequality. We define truncated inter renewal times $\bar{X} : \Omega \rightarrow [0, M]^{\mathbb{N}}$ defined as $\bar{X}_n \triangleq X_n \wedge M$ for each $n \in \mathbb{N}$, and with common mean denoted by $\mu_M \triangleq \mathbb{E}X_1 \wedge M$. Since X is i.i.d., so is the truncated sequence \bar{X} , and hence we can define the corresponding renewal sequence $\bar{S} : \Omega \rightarrow \mathbb{R}_+^{\mathbb{N}}$ as $\bar{S}_n \triangleq \sum_{i=1}^n \bar{X}_i$ for each $n \in \mathbb{N}$, and the counting process $\bar{N} : \Omega \rightarrow \mathbb{Z}_+^{\mathbb{N}}$ as $\bar{N}_t \triangleq \sum_{n \in \mathbb{N}} \mathbb{1}_{\{\bar{S}_n \leq t\}}$ for each $t \in \mathbb{R}_+$. Note that since $S_n \geq \bar{S}_n$, the number of arrivals would be higher for renewal process \bar{N}_t with truncated random variables. That is, $N_t \leq \bar{N}_t$, and hence $m_t \leq \bar{m}_t$ from the monotonicity of expectation. Further, due to truncation of inter arrival time, next renewal happens within M units of time, that is $\bar{S}_{\bar{N}_t+1} \leq t + M$. From the monotonicity of expectation and Wald's Lemma for renewal processes in Proposition 1.7, we get $(1 + \bar{m}_t)\mu_M \leq t + M$. Dividing both sides by $t\mu_M$ and the fact that $m_t \leq \bar{m}_t$ for all times $t \in \mathbb{R}_+$, we obtain $\limsup_{t \rightarrow \infty} m_t/t \leq \limsup_{t \rightarrow \infty} \bar{m}_t/t \leq 1/\mu_M$. The result follows from recognizing that $\lim_{M \rightarrow \infty} \mu_M = \mu$. \square

Corollary 1.9. If a delayed renewal process has finite inter renewal durations, then $\lim_{t \rightarrow \infty} m_D(t)/t = 1/\mu_F$.

Example 1.10 (Markov chain). Consider a positive recurrent discrete time Markov chain $X : \Omega \rightarrow \mathcal{X}^{\mathbb{N}}$. Let the initial state be $X_0 = x \in \mathcal{X}$ and $\tau_y^+(0) = 0$ for $y \neq x \in \mathcal{X}$, then we can inductively define the k th recurrent time to state y as a stopping time $\tau_y^+(k) \triangleq \inf \{n > \tau_y^+(k-1) : X_n = y\}$. Since any discrete time Markov chain satisfies the strong Markov property, it follows that $\tau_y^+ : \Omega \rightarrow \mathbb{N}^{\mathbb{N}}$ form a delayed renewal process, where the first renewal distribution is $P_x \{\tau_y^+(1) = k\} = f_{xy}^{(k)}$ and the common distribution of the inter renewal duration is $P_y \{\tau_y^+(1) = k\} = f_{yy}^{(k)}$ for $k \in \mathbb{N}$. We denote the associated counting process by $N_y : \Omega \rightarrow \mathbb{Z}_+^{\mathbb{N}}$, where $N_y(n)$ denotes the number of visits to state y up to time n , and is defined as

$$N_y(n) \triangleq \sum_{k \in \mathbb{N}} \mathbb{1}_{\{\tau_y^+(k) \leq n\}} = \sum_{k=1}^n \mathbb{1}_{\{X_k = y\}}.$$

From the strong law for delayed renewal sequence τ_y^+ with finite mean inter renewal time $\mu_{yy} \triangleq \mathbb{E}_y \tau_y^+(1)$ (which is also the mean recurrence time to state y), it follows that

$$P_y \left\{ \lim_{n \in \mathbb{N}} \frac{N_y(n)}{n} = \frac{1}{\mu_{yy}} \right\} = 1.$$

From the elementary renewal theorem for delayed renewal sequence τ_y^+ , it follows that

$$\lim_{n \in \mathbb{N}} \frac{1}{n} \sum_{k=1}^n p_{xy}^{(k)} = \lim_{n \in \mathbb{N}} \frac{1}{n} \mathbb{E}_x[N_y(n)] = \frac{1}{\mu_{yy}}.$$

1.3 Central limit theorem for renewal processes

Theorem 1.11. If $\mu \triangleq \mathbb{E}X_1$ and $\sigma^2 \triangleq \text{Var}(X_1)$, then $N_t \rightarrow \mathcal{N}(\frac{t}{\mu}, \sigma^2 \frac{t}{\mu^3})$ for large t in distribution. Specifically,

$$\lim_{t \rightarrow \infty} P \left\{ \frac{N_t - \frac{t}{\mu}}{\sigma \sqrt{\frac{t}{\mu^3}}} < y \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-\frac{x^2}{2}} dx.$$

Proof. Take $u = \frac{t}{\mu} + y\sigma\sqrt{\frac{t}{\mu^3}}$. We shall treat u as an integer and proceed, the proof for general u is an exercise. Recall that $\{N_t < u\} = \{S_u > t\}$. By equating probability measures on both sides, we get

$$P\{N_t < u\} = P\left\{\frac{S_u - u\mu}{\sigma\sqrt{u}} > \frac{t - u\mu}{\sigma\sqrt{u}}\right\} = P\left\{\frac{S_u - u\mu}{\sigma\sqrt{u}} > -y\left(1 + \frac{y\sigma}{\sqrt{t\mu}}\right)^{-1/2}\right\}.$$

By central limit theorem, $\frac{S_u - u\mu}{\sigma\sqrt{u}}$ converges to a normal random variable with zero mean and unit variance as t grows. We also observe that $\lim_{t \rightarrow \infty} -y\left(1 + \frac{y\sigma}{\sqrt{t\mu}}\right)^{-1/2} = -y$. These results combine with the symmetry of normal random variable to give us the result. \square

2 Patterns

Let $X : \Omega \rightarrow \mathcal{X}^{\mathbb{N}}$ be an *i.i.d.* sequence with common probability mass function $p \in \mathcal{M}(\mathcal{X})$. We denote the natural filtration of process X by $\mathcal{F}_{\bullet} \triangleq (\mathcal{F}_n : n \in \mathbb{N})$ where $\mathcal{F}_n \triangleq \sigma(X_1, \dots, X_n)$ for all $n \in \mathbb{N}$. Consider a pattern $x = (x_1, \dots, x_m) \in \mathcal{X}^m$, take $S_0^x \triangleq 0$, and inductively define k th hitting times of the pattern x as

$$S_k^x \triangleq \inf\{n > S_{k-1}^x : X_n = x_m, X_{n-1} = x_{m-1}, \dots, X_{n-m+1} = x_1\}.$$

It is easy to check that S_k^x is adapted to \mathcal{F}_{\bullet} and we will verify that S_k^x is almost surely finite for all $k \in \mathbb{N}$ given p is positive. It follows that S^x is a sequence of stopping times adapted to \mathcal{F}_{\bullet} . Since X is *i.i.d.*, it follows that $S^x : \Omega \rightarrow \mathbb{R}_+^{\mathbb{N}}$ is a delayed renewal sequence with inter renewal durations $T_k^x \triangleq S_k^x - S_{k-1}^x$ being *i.i.d.* for $k \geq 2$ and independent of T_1^x .

2.1 Hitting time to pattern (1)

First we consider the simplest example when the alphabet $\mathcal{X} = \{0, 1\}$, with the common mean $\mathbb{E}X_1 = p$, and the pattern $x = (1)$. One way to solve this problem is to consider S_1^1 as a random variable and find its distribution. We can write

$$P\{S_1^1 = k\} = \bar{p}^{k-1}p.$$

We observe that S_1^1 is a geometric random variable of the time to first success, with its mean as the reciprocal of *i.i.d.* success probability p . An alternative way to solve this is via renewal function approach. Recall that $\{S_1^1 = 1\} = \{X_1 = 1\}$ and $S_1^1 \mathbb{1}_{\{X_1=0\}} = (1 + S_1^1) \mathbb{1}_{\{X_1=0\}}$ in distribution where S_1^1 is independent of X_1 and distributed identically to S_1^1 . The result follows from writing

$$\mathbb{E}S_1^1 = \mathbb{E}S_1^1 \mathbb{1}_{\{S_1^1 > 1\}} + \mathbb{E}S_1^1 \mathbb{1}_{\{S_1^1 = 1\}} = \bar{p}\mathbb{E}(1 + S_1^1) + p = 1 + \bar{p}\mathbb{E}S_1^1.$$

2.2 Hitting time to pattern (0, 1)

For the alphabet $\mathcal{X} = \{0, 1\}$ with common mean $\mathbb{E}X_1 = p$, we consider the two length pattern $x = (0, 1)$, then $S_1^x = \inf\{n \in \mathbb{N} : X_n = 1, X_{n-1} = 0\}$. We can again model this hitting time as a random variable, however directly finding the distribution of S^x is slightly more complicated. We next attempt the renewal function approach. We take S_1^x to be an independent replica of S_1^x is independent of X_1 , and the following equality holds in distribution $S_1^x \mathbb{1}_{\{X_1=1\}} = (1 + S_1^x) \mathbb{1}_{\{X_1=1\}}$. In addition, the following equality holds in distribution $S_1^x \mathbb{1}_{\{X_2=0, X_1=0\}} = (1 + S_1^x) \mathbb{1}_{\{X_2=0, X_1=0\}}$. Hence, we can write

$$\mathbb{E}S_1^x = \mathbb{E}S_1^x \mathbb{1}_{\{X_1=0\}} + \mathbb{E}S_1^x \mathbb{1}_{\{X_1=1\}} = \mathbb{E}S_1^x \mathbb{1}_{\{X_2=1, X_1=0\}} + \mathbb{E}S_1^x \mathbb{1}_{\{X_2=0, X_1=0\}} + p\mathbb{E}(1 + S_1^x).$$

We recognize that the second term on the right hand side can be written as

$$\mathbb{E}S_1^x \mathbb{1}_{\{X_2=0, X_1=0\}} = \bar{p}\mathbb{E}(1 + S_1^x) \mathbb{1}_{\{X_1=0\}} = \bar{p}^2 + \bar{p}\mathbb{E}S_1^x \mathbb{1}_{\{X_1=0\}} = \bar{p}^2 + \bar{p}\mathbb{E}S_1^x - \bar{p}p\mathbb{E}(1 + S_1^x).$$

Combining the above two results, we can write

$$\mathbb{E}S_1^x = 2p\bar{p} + \bar{p}^2 + \bar{p}\mathbb{E}S_1^x + p^2\mathbb{E}(1 + S_1^x) = 1 + (\bar{p} + p^2)\mathbb{E}S_1^x.$$

It follows that $\mathbb{E}S_1^x = \frac{1}{p\bar{p}}$.

2.3 Hitting time to pattern x

For an *i.i.d.* sequence $X : \Omega \rightarrow \mathcal{X}^{\mathbb{N}}$, a general approach is to model $X_n^m = (X_n, X_{n-1}, \dots, X_{n-m+1}) \in \mathcal{X}^m$ as an m -dimensional time homogeneous irreducible positive recurrent Markov chain. We define $\mathcal{Y} \triangleq \mathcal{X}^m$, a discrete time process $Y : \Omega \rightarrow \mathcal{Y}^{\mathbb{N}}$ such that $Y_n \triangleq X_n^m$, and k th hitting time to state $x \in \mathcal{Y}$ as

$$S_k^x \triangleq \inf \{n > S_{k-1}^x : Y_n = x\}.$$

We are interested in the mean of first hitting time S_1^x to state $x \in \mathcal{Y}$ of the joint process $X^m : \Omega \rightarrow \mathcal{Y}^{\mathbb{N}}$. It follows that the successive times $S^x : \Omega \rightarrow \mathbb{N}^{\mathbb{N}}$ for process X^m to hit a pattern $x \in \mathcal{Y}$ is a delayed renewal process in general. Defining the on times when X^m hits x , it follows from the strong law for renewal processes that the average number of visits to state x is the reciprocal of mean inter renewal duration. That is, defining $T_k^x \triangleq S_k^x - S_{k-1}^x$ for all $k \in \mathbb{N}$, we have

$$\lim_{n \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mathbb{1}_{\{X_n^m = x\}} = \lim_{n \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mathbb{1}_{\{X_n = x_m, \dots, X_{n-m+1} = x_1\}} = \prod_{i=1}^m p_{x_i} = \frac{1}{\mathbb{E}T_k^x}.$$

Definition 2.1. For each pattern $x \in \mathcal{Y}$, we define sub-patterns $w^k \triangleq (w_1, \dots, w_k) \in \mathcal{X}^k$ for $k \in [m]$. We say that w^k is a *sub-pattern* of x if $w^k = (x_j, \dots, x_{j+k-1})$ for each $j \in [m-k]$ and $(w_1, \dots, w_k) = (x_{m-k+1}, \dots, x_m)$.

Remark 1. We observe that hitting time to a pattern x is also a hitting time to sub-pattern w^k . Therefore, first hitting time to pattern x is $\mathbb{E}S_1^{w^k} + \mathbb{E}_{w^k}S_1^x$. If x has no sub-patterns, i.e. the initial part of x is not one of the final parts, i.e. $(x_1, \dots, x_k) \neq (x_{m-k+1}, \dots, x_m)$ for any $k \in [m]$, then we observe that S^x is a renewal sequence and $\mathbb{E}S_1^x = \frac{1}{\prod_{i=1}^m p_{x_i}}$.

Example 2.2. Consider patterns (1) and (01) for *i.i.d.* Bernoulli sequence $X : \Omega \rightarrow \{0,1\}^{\mathbb{N}}$ with common mean $\mathbb{E}X_1 = p$. Clearly, (1) has no sub-pattern and hence $\mathbb{E}S_1^1 = \frac{1}{p}$. Similarly, (01) has no sub-pattern and hence $\mathbb{E}S_1^{01} = \frac{1}{\bar{p}p}$.

We can generalize this to arbitrary number of sub-patterns of x . If there exists a non empty $I \subseteq [m]$ such that for each $k \in I$ there is an initial sub-pattern x^k such that $(x^k) = (x_{m-k+1}^m)$ is a final sub-pattern of x , then the mean hitting time to pattern x is equal to the telescopic sum of mean hitting time to sub-patterns. That is, denoting $I \triangleq \{i_1, \dots, i_\ell\}$, we can write

$$\mathbb{E}S_1^x = \sum_{j=1}^{\ell} \mathbb{E}_{x^{i_j}} S_1^{x^{i_{j+1}}}.$$

The mean time duration between two successive hits to $x^{i_{j+1}}$ is $\mathbb{E}T_2^{x^{i_{j+1}}} = \frac{1}{\prod_{r=1}^{i_{j+1}} p_{x_r}}$. This is the same mean time from x^{i_j} to $x^{i_{j+1}}$. Therefore,

$$\mathbb{E}S_1^x = \sum_{j=1}^{\ell} \frac{1}{\prod_{r=1}^{i_{j+1}} p_{x_r}}.$$

Example 2.3. Consider pattern (101) and (1011) for *i.i.d.* Bernoulli sequence $X : \Omega \rightarrow \{0,1\}^{\mathbb{N}}$ with common mean $\mathbb{E}X_1 = p$. Pattern (101) has sub-patterns (1) that appears at the end of (101). Therefore,

$$\mathbb{E}S_1^{(101)} = \mathbb{E}S_1^1 + \mathbb{E}_1 S_1^{101} = \frac{1}{p} + \frac{1}{p^2 \bar{p}}.$$

Pattern (1011) has a sub-pattern (1) that appears at the end of (1011). Thus,

$$\mathbb{E}S_1^{(1011)} = \mathbb{E}S_1^1 + \mathbb{E}_1 S_1^{1011} = \frac{1}{p} + \frac{1}{p^3 \bar{p}}.$$