О ЗАКОНЕ АРКСИНУСА ДЛЯ ВРЕМЕНИ ПРЕБЫВАНИЯ ВЫШЕ НУЛЯ СЛУЧАЙНОГО БЛУЖДАНИЯ

Панфилов И.А.

В работе рассматривается симметричное случайное блуждание $\{S_n\}_{n\geq 0}$. Вводится случайная величина μ_n – число отрезков $[k-1,k]$, $k=1,...,n$, для которых хотя бы одна из величин S_{k-1} или S_k положительна. Устанавливается, что величина $\frac{\mu_n}{n}$ сходится по распределению при $n\to\infty$ к случайной величине μ, распределённой по закону арксинуса.

1 Введение

Рассмотрим независимые одинаково распределённые бернуллиевские случайные величины $X_1, X_2, ...$ на некотором вероятностном пространстве $(\Omega, F, \mathbf{P})$, где $\mathbf{P}(X_1 = +1) = \mathbf{P}(X_1 = -1) = \frac{1}{2}$. Положим

$$S_0 = 0, \qquad S_n = \sum_{k=1}^{n} X_k,\ n\geq 1.$$

Введём некоторые обозначения:

$$I(n) = \{1, ..., n\},\ I_0(n) = \{0, 1, ..., n\}.$$

Для каждой реализации X_k, $k\in I(n)$, на плоскости в координатах (x,y) отметим точки (m, S_m), где $m\in I_0(n)$. Соединим отрезками прямых соседние отмеченные точки. Полученную ломаную назовем *траекторией случайного блуждания длины* n.

Можно рассматривать величину S_n как координату блуждающей частицы в момент времени n, где $n\geq 0$. В таком случае траектория – это график изменения координаты с течением времени.

Скажем, что частица на отрезке $[k-1,k]$, где $k\in\mathbf{N}$, находится на положительной стороне, если хотя бы одно из значений S_{k-1} и S_k положительно. Тогда можно рассмотреть величину

$$\mu_n = \#\{k\in I(n) : S_{k-1} > 0 \textit{ или } S_k > 0\},$$

то есть количество отрезков, на которых частица находилась выше нуля. В этом случае $\dfrac{\mu_n}{n}$ – это доля времени, когда частица находилась выше нуля.

Введем некоторые виды сходимости.

Определение 1. Последовательность случайных величин $\{\xi_n\}_{n\in\mathbf{N}}$ *сходится по распределению* к случайной величине ξ при $n\to\infty$, если

$$\mathbf{E}f(\xi_n)\to\mathbf{E}f(\xi),$$

для любой непрерывной ограниченной функции $f:\mathbf{R}\to\mathbf{R}$. Обозначение: $\xi_n \xrightarrow{d} \xi$.

Определение 2. Последовательность функций распределения $\{F_n(\cdot)\}_{n\in\mathbf{N}}$ *сходится в основном* к функции распределения $F(\cdot)$ при $n\to\infty$, если

$$F_n(x)\to F(x),$$

в любой точке x, в которой F непрерывна.

Известно, что сходимость случайных величин по распределению эквивалентна сходимости в основном их функций распределения.

Справедлив следующий результат для случайных величин $\frac{\mu_n}{n}$, получивший название *закона арксинуса для времени пребывания выше нуля.*

Теорема 1. *При* $n \to \infty$

$$\frac{\mu_n}{n} \xrightarrow{d} \mu,$$

где $\mathbf{P}(\mu \le x) = \frac{2}{\pi} arcsin\sqrt{x}, \quad 0 \le x \le 1.$

2 Вспомогательные результаты

Введем некоторые вероятности для $n \in \mathbf{N},\ k \in I_0(n)$:

$$u_{2k} = \mathbf{P}(S_{2k} = 0),$$

$$f_{2k} = \mathbf{P}(S_1 \neq 0, \dots, S_{2k-1} \neq 0, S_{2k} = 0),$$

$$P_{2k,2n} = \mathbf{P}(\mu_{2n} = 2k),$$

то есть u_{2k} – вероятность возвращения частицы в ноль в момент времени $2k$, f_{2k} – вероятность того, что первое возвращение частицы в ноль произойдёт в момент времени $2k$, а $P_{2k,2n}$ – вероятность того, что число отрезков, на которых частица находилась выше нуля, равно $2k$.

Назовём путём любую ломаную, соединяющую на плоскости две точки с целочисленными координатами, каждое звено которой как вектор есть либо $(1,1)$, либо $(1,-1)$. Все вершины такой ломаной – целочисленные точки, обозначим их координаты как (k, s_k), где $k \in N_0$. Обозначим такой путь $s_0, s_1, \dots$. Можно рассмотреть конечный участок пути $s_0, s_1, ..., s_n$, про него будем говорить, что это путь длины n.

Ясно, что некоторые пути являются траекториями случайного блуждания. Скажем, что путь l, который задается набором значений $s_0, ..., s_n$, является допустимым для случайного события A, если левый конец пути l, точка $(0, s_0)$ есть точка $(0,0)$, и при этом с путём l совпадает некоторая возможная траектория, которая задается реализацией $i_1, ..., i_n \in \{+1, -1\}$, для которой $\{X_1 = i_1, ..., X_n = i_n\} \subseteq A$. Будем считать, что $p_n(A)$ – число допустимых путей длины n для события A.

Отметим, что $p_n(\Omega) = 2^n$ для любого $n \in \mathbf{N}$, поскольку подойдет любой путь длины n, выходящий из начала координат, а такой путь задается n выборами одного из двух значений для вектора-звена.

Основная польза от введенного понятия пути – удобный способ вычисления вероятностей событий. Событие A обозначим как $A^{(n)}$, если A выражается через события $\{X_k = +1\}$ и $\{X_k = -1\}$, где $1 \le k \le n$, а n – некоторое натуральное число. Все траектории равновероятны, а значит, легко подсчитать вероятность такого события:

$$\mathbf{P}(A^{(n)}) = \frac{p_n(A^{(n)})}{2^n}.$$

Вычислим u_{2k}. Событие $\{S_{2k} = 0\}$ происходит тогда и только тогда, когда

$$\#\{i \in I(2k) : X_i = +1\} = \#\{i \in I(2k) : X_i = -1\} = k.$$

Число допустимых путей есть C_{2k}^k. Тогда

$$u_{2k} = C_{2k}^k \frac{1}{2^{2k}}. \tag{1}$$

Согласно формуле Стирлинга

$$n! = n^n e^{-n} \sqrt{2\pi n}(1 + o(n)), \tag{2}$$

при $n \to \infty$. Из (1) и (2) получаем, что при $k \to \infty$

$$u_{2k} = \frac{(2k)!}{k!k!}\frac{1}{2^{2k}} = \frac{(2k)^{2k}e^{-2k}\sqrt{2\pi 2k}(1+o(2k))}{k^{2k}e^{-2k}2\pi k(1+o(k^2))}\frac{1}{2^{2k}} = \frac{1+o(2k)}{\sqrt{\pi k}(1+o(k^2))},$$

следовательно

$$u_{2k} \sim \frac{1}{\sqrt{\pi k}}. \tag{3}$$

Далее, нам понадобятся следующие леммы.

Лемма 1. *Для любого* $k \in \mathbf{N}$

$$u_{2k} = \sum_{r=1}^{k} f_{2r}u_{2k-2r}. \tag{4}$$

Доказательство. Пусть τ_{2k} – момент первого возвращения частицы в ноль на отрезке $[0, 2k]$. События $\{\tau_{2k} = 2r\},\ r \in I(k)$, вместе с событием $A_{2k} = \{S_1 \neq 0, ..., S_{2k} \neq 0\}$ образуют разбиение вероятностного пространства. Тогда по формуле полной вероятности

$$\mathbf{P}(S_{2k} = 0) = \mathbf{P}(S_{2k} = 0 \mid A_{2k})\mathbf{P}(A_{2k}) + \sum_{r=1}^{k} \mathbf{P}(S_{2k} = 0 \mid \tau_{2k} = 2r)\mathbf{P}(\tau_{2k} = 2r). \tag{5}$$

По определению f_{2r}, где $r \in I(k)$,

$$\mathbf{P}(\tau_{2k} = 2r) = f_{2r}.$$

События $\{S_{2k} = 0\}$ и A_{2k} несовместны, значит

$$\mathbf{P}(S_{2k} = 0 \mid A_{2k}) = 0.$$

Подсчитаем остальные условные вероятности из формулы (5):

$$\mathbf{P}(S_{2k} = 0 \mid \tau_{2k} = 2r) = \mathbf{P}(S_{2r} + X_{2r+1} + ... + X_{2k} = 0 \mid S_2 \neq 0, \dots, S_{2r-2} \neq 0, S_{2r} = 0) =$$

$$= \mathbf{P}(X_{2r+1} + ... + X_{2k} = 0 \mid S_2 \neq 0, \dots, S_{2r-2} \neq 0, S_{2r} = 0) = \mathbf{P}(X_{2r+1} + ... + X_{2k} = 0) =$$

$$= \mathbf{P}(X_1 + ... + X_{2k-2r} = 0) = u_{2k-2r}.$$

Теперь подставим вычисленные вероятности в (5), получим

$$\mathbf{P}(S_{2k} = 0) = \sum_{r=1}^{k} f_{2r}u_{2k-2r},$$

а это есть утверждение леммы. $\square$

Лемма 2. *Если* $k \in \mathbf{N}$, *то*

$$\mathbf{P}(S_{2k} = 0) = \mathbf{P}(S_1 \geq 0, ..., S_{2k} \geq 0) = \mathbf{P}(S_1 \leq 0, ..., S_{2k} \leq 0). \tag{6}$$

Доказательство. Для любого допустимого пути $s_0, ..., s_{2k}$ для события $\{S_{2k} = 0\}$, рассмотрим r – момент первого достижения минимума. Отразим участок пути на отрезке $[0, r]$ относительно вертикальной прямой $x = r$. Затем сместим левый конец отраженного участка пути в точку $(2k, 0)$. Далее, перенесем начало координат в точку (r, s_r). Получим некоторый допустимый путь для события $\{S_1 \geq 0, ..., S_{2k} \geq 0\}$.

Обратно, возьмём допустимый путь для события $\{S_1 \geq 0, ..., S_{2k} \geq 0\}$. Пусть у такого пути $s_{2k} = 2m$ и p – самое большое число, для которого $s_p = m$. Отразим участок пути на отрезке $[p, 2k]$ относительно прямой $x = p$ и сместим правый конец отраженной части пути в точку $(2k, 0)$. Полученный путь является допустимым для события $\{S_{2k} = 0\}$.

Построенное соответвествие между множествами взаимно однозначно. Взаимная однозначность неположительных и неотрицательных путей очевидна из соображений симметрии относительно оси Ox. Таким образом, лемма полностью доказана. $\square$

Лемма 3. *Если* $k \in I_0(n)$, *где* $n \in \mathbf{N}$, *то*

$$P_{2k,2n} = u_{2k}u_{2n-2k}. \tag{7}$$

Доказательство. Если $k = n$ или $k = 0$, то из (6) следует (7).

Пусть теперь $1 \leq k \leq n-1$. Допустимые пути $s_0, ..., s_{2n}$ для события $\{\mu_{2n} = 2k\}$ есть пути, двигаясь по которым, частица находилась выше нуля ровно на $2k$ отрезках, что также означает, что частица находилась ниже нуля ровно на $2n-2k$ отрезках. Ясно, любой что такой путь пересечет ось абсцисс, то есть найдется s_l, равное нулю, и пусть s_{2r} – первая такая нулевая вершина. В этом случае на отрезке $[0, 2r]$ частица была либо все время выше нуля, и тогда число допустимых путей есть

$$\frac{1}{2}(2^{2r}f_{2r})(2^{2n-2r}P_{2k-2r,2n-2r}),$$

либо же частица была все время ниже нуля, в таком случае число допустимых путей есть

$$\frac{1}{2}(2^{2r}f_{2r})(2^{2n-2r}P_{2k,2n-2r}).$$

Тогда для любого $k \in I(n-1)$

$$P_{2k,2n} = \frac{1}{2}\sum_{r=1}^{k} f_{2r}P_{2k-2r,2n-2r} + \frac{1}{2}\sum_{r=1}^{n-k} f_{2r}P_{2k,2n-2r}. \tag{8}$$

Докажем формулу (7) индукцией по n. Для $n = 1$ формула верна. Предположим, что она верна для любого $m \in I(n-1)$. Тогда из (4) и (8), а также из индуктивного предположения, получаем

$$P_{2k,2n} = \frac{1}{2}u_{2n-2k}\sum_{r=1}^{k} f_{2r}u_{2k-2r} + \frac{1}{2}u_{2k}\sum_{r=1}^{n-k} f_{2r}u_{2n-2r-2k} = \frac{1}{2}u_{2n-2k}u_{2k} + \frac{1}{2}u_{2k}u_{2n-2k} = u_{2k}u_{2n-2k}. \ \square$$

3 Основной результат

Докажем теорему 1. Пусть $0 < a < b < 1$. Сперва покажем, что при $n \to \infty$

$$\mathbf{P}(a < \frac{\mu_{2n}}{2n} \leq b) \to \int\limits_a^b \frac{1}{\pi\sqrt{t(1-t)}}dt. \tag{9}$$

По определению $P_{2k,2n}$, а также по (7)

$$\mathbf{P}(a < \frac{\mu_{2n}}{2n} \leq b) = \sum_{\{k: a < \frac{2k}{2n} \leq b\}} P_{2k,2n} = \sum_{\{k: a < \frac{2k}{2n} \leq b\}} u_{2k}u_{2n-2k}. \tag{10}$$

Формула (3) означает, что для любого $\varepsilon > 0$ для всех достаточно больших k

$$|u_{2k}\sqrt{\pi k} - 1| \leq \varepsilon,$$

откуда получаем

$$\frac{1-\varepsilon}{\sqrt{\pi k}} \le u_{2k} \le \frac{1+\varepsilon}{\sqrt{\pi k}}$$

Аналогично для всех достаточно больших $n-k$

$$\frac{1-\varepsilon}{\sqrt{\pi(n-k)}} \le u_{2n-2k} \le \frac{1+\varepsilon}{\sqrt{\pi(n-k)}}.$$

Значит, при всех достаточно больших k и $n-k$, имеем

$$\frac{(1-\varepsilon)^2}{\pi\sqrt{k(n-k)}} \le u_{2k}u_{2n-2k} \le \frac{(1+\varepsilon)^2}{\pi\sqrt{k(n-k)}}.$$

При условии $a < \frac{2k}{2n} \le b$ можно подобрать достаточно большое n так, чтобы k и $n-k$ были достаточно большими и чтобы выполнялось последнее неравенство. Поэтому для всех достаточно больших n

$$\sum_{\{k: a<\frac{2k}{2n}\le b\}} \frac{(1-\varepsilon)^2}{\pi\sqrt{k(n-k)}} \le \sum_{\{k: a<\frac{2k}{2n}\le b\}} u_{2k}u_{2n-2k} \le \sum_{\{k: a<\frac{2k}{2n}\le b\}} \frac{(1+\varepsilon)^2}{\pi\sqrt{k(n-k)}}. \tag{11}$$

Заметим, что

$$\sum_{\{k: a<\frac{2k}{2n}\le b\}} \frac{1}{\pi\sqrt{k(n-k)}} = \sum_{\{k: a<\frac{2k}{2n}\le b\}} \frac{1}{\pi n\sqrt{\frac{k}{n}(1-\frac{k}{n})}}.$$

Сумма в правой части является суммой Римана для функции $f(t) = \dfrac{1}{\pi\sqrt{t(1-t)}}$ на отрезке $[a,b]$. А значит если устремить n к ∞ в неравенстве (11), то получим

$$(1-\varepsilon)^2 \int\limits_a^b \frac{1}{\pi\sqrt{t(1-t)}}dt \le \varliminf_{n\to\infty} \sum_{\{k: a<\frac{2k}{2n}\le b\}} u_{2k}u_{2n-2k} \le$$

$$\le \varlimsup_{n\to\infty} \sum_{\{k: a<\frac{2k}{2n}\le b\}} u_{2k}u_{2n-2k} \le (1+\varepsilon)^2 \int\limits_a^b \frac{1}{\pi\sqrt{t(1-t)}}dt.$$

Устремим ε к 0, тогда

$$\lim_{n\to\infty} \sum_{\{k: a<\frac{2k}{2n}\le b\}} u_{2k}u_{2n-2k} = \int\limits_a^b \frac{1}{\pi\sqrt{t(1-t)}}dt,$$

откуда с учетом (10) следует (9).

Далее, установим, что

$$\lim_{a\to 0} \varlimsup_{n\to\infty} \mathbf{P}(\frac{\mu_{2n}}{2n} \le a) = 0. \tag{12}$$

По определению $P_{2k,2n}$, а также по (7)

$$\mathbf{P}(\frac{\mu_{2n}}{2n} \le a) = \sum_{\{k: 0\le\frac{2k}{2n}\le a\}} P_{2k,2n} = u_{2n} + \sum_{\{k: 0<\frac{2k}{2n}\le a\}} u_{2k}u_{2n-2k}. \tag{13}$$

Из (3) следует, что существует постоянная C_1 такая, что для любого $n \in N$

$$u_{2n} \le \frac{C_1}{\sqrt{\pi n}}.$$

Тогда

$$u_{2n} + \sum_{\{k:0<\frac{2k}{2n}\le a\}} u_{2k}u_{2n-2k} \le \frac{C_1}{\sqrt{\pi n}} + \frac{C_1^2}{\pi} \sum_{\{k:0<\frac{2k}{2n}\le a\}} \frac{1}{\sqrt{k(n-k)}} \le$$

$$\le \frac{C_1}{\sqrt{\pi n}} + \frac{C_1^2}{\pi} \sum_{\{k:0<\frac{2k}{2n}\le a\}} \frac{1}{\sqrt{k(n-na)}} = \frac{C_1}{\sqrt{\pi n}} + \frac{C_1^2}{\pi\sqrt{n(1-a)}} \sum_{\{k:0<\frac{2k}{2n}\le a\}} \frac{1}{\sqrt{k}}. \tag{14}$$

Из интегрального признака Коши-Маклорена следует, что последовательность частичных сумм $\{\sum\limits_{k=1}^{n} \frac{1}{\sqrt{k}}\}_{n\in\mathbf{N}}$ стремится к бесконечности при $n\to\infty$ и имеет порядок $O(\sqrt{n})$, значит,

$$\sum_{\{k:0<\frac{2k}{2n}\le a\}} \frac{1}{\sqrt{k}} \le C_2\sqrt{[na]} \le C_2\sqrt{na},$$

для некоторой константы C_2. Далее, применим эту оценку для (14):

$$u_{2n} + \sum_{\{k:0<\frac{2k}{2n}\le a\}} u_{2k}u_{2n-2k} \le \frac{C_1}{\sqrt{\pi n}} + \frac{C_1^2C_2\sqrt{na}}{\pi\sqrt{n(1-a)}}.$$

Тогда, применяя это неравенство к соотношению (13), получим оценку

$$\mathbf{P}(\frac{\mu_{2n}}{2n} \le a) \le \frac{C_1}{\sqrt{\pi n}} + \frac{C_1^2C_2\sqrt{na}}{\pi\sqrt{n(1-a)}}.$$

Тогда

$$\overline{\lim_{n\to\infty}} \mathbf{P}(\frac{\mu_{2n}}{2n} \le a) \le \frac{C_1^2C_2\sqrt{a}}{\pi\sqrt{1-a}}.$$

Значит,

$$\lim_{a\to 0}\overline{\lim_{n\to\infty}} \mathbf{P}(\frac{\mu_{2n}}{2n} \le a) \to 0. \tag{15}$$

Теперь покажем, что для любого $b\in(0,1)$ при $n\to\infty$

$$\mathbf{P}(\frac{\mu_{2n}}{2n} \le b) \to \int\limits_0^b \frac{1}{\pi\sqrt{t(1-t)}}dt.$$

Для любого $a\in(0,b)$

$$\mathbf{P}(\frac{\mu_{2n}}{2n} \le b) = \mathbf{P}(\frac{\mu_{2n}}{2n} \le a) + \mathbf{P}(a < \frac{\mu_{2n}}{2n} \le b). \tag{16}$$

Тогда

$$\lim_{n\to\infty} \mathbf{P}(a < \frac{\mu_{2n}}{2n} \le b) \le \underline{\lim_{n\to\infty}} \mathbf{P}(\frac{\mu_{2n}}{2n} \le b).$$

Заметим, что предел в левой части известен из (9), а предел в правой части не зависит от a. Тогда, переходя к пределу при $a\to 0$, получим

$$\int\limits_0^b \frac{1}{\pi\sqrt{t(1-t)}}dt \le \underline{\lim_{n\to\infty}} \mathbf{P}(\frac{\mu_{2n}}{2n} \le b). \tag{17}$$

Также из (16) следует, что

$$\overline{\lim_{n\to\infty}} \mathbf{P}(\frac{\mu_{2n}}{2n} \le b) \le \overline{\lim_{n\to\infty}} \mathbf{P}(\frac{\mu_{2n}}{2n} \le a) + \lim_{n\to\infty} \mathbf{P}(a < \frac{\mu_{2n}}{2n} \le b).$$

По (15) первое слагаемое правой части стремится к нулю при $a \to 0$, а из (9) известно второе слагаемое правой части. Тогда, переходя к пределу при $a \to 0$, получим

$$\overline{\lim_{n\to\infty}} \mathbf{P}(\frac{\mu_{2n}}{2n} \leq b) \leq \int\limits_0^b \frac{1}{\pi\sqrt{t(1-t)}}dt. \tag{18}$$

Тогда в силу (17) и (18)

$$\lim_{n\to\infty} \mathbf{P}(\frac{\mu_{2n}}{2n} \leq b) = \int\limits_0^b \frac{1}{\pi\sqrt{t(1-t)}}dt. \tag{19}$$

Найдем предел в крайних точках. Для любого $n \in \mathbf{N}$

$$\mathbf{P}(\frac{\mu_{2n}}{2n} \leq 1) = 1 = \int\limits_0^1 \frac{1}{\pi\sqrt{t(1-t)}}dt. \tag{20}$$

Также по определению $P_{2k,2n}$

$$\mathbf{P}(\frac{\mu_{2n}}{2n} \leq 0) = P_{0,2n} = u_{2n}.$$

Из (3) следует, что

$$\lim_{n\to\infty} \mathbf{P}(\frac{\mu_{2n}}{2n} \leq 0) = 0. \tag{21}$$

Из (19), (20), (21) следует, что для любого $x \in [0,1]$

$$\lim_{n\to\infty} \mathbf{P}(\frac{\mu_{2n}}{2n} \leq x) = \int\limits_0^x \frac{1}{\pi\sqrt{t(1-t)}}dt = \frac{2}{\pi} arcsin\sqrt{x}.$$

Это означает сходимость в основном функций распределения случайных величин $\frac{\mu_{2n}}{2n}$ к функции распределения арксинуса, а из этого следует сходимость этих величин по распределению к некоторой случайной величине μ, которая распределена по закону арксинуса. Тем самым, доказан закон арксинуса для времени пребывания выше нуля.