Случајно талкање и Винеров процес

[Илија Боневски]

30 септември 2025 год.

1 Вовед

Изучувањето на случајни талкања и Винеровиот процес започнува во 19-ти кон ран 20-ти век, се изучуваат прости талкања, нивните законитости за рекуренција и транзитивност од Поја, Луи Башеље и Ајнштајн користат случајни движења за моделирање на финансии и Брауново движење, и Норберт Винер прв ригорозно дефинира непрекинато Брауново движење. Понатаму се поврзуваат дискретните талкања со континуираната дифузија преку функционалната гранична теорема (Донскеров принцип на инваријантност), се развиваат Левиевите процеси и на мартингалите, и врз нив се основа современото стохастичкото сметање (стохастичка анализа), како и теоријата на потенцијалот.

Како модели во реалниот свет: Брауновото движење моделира дифузија во хемија и физика, а случајни талкања опишуваат модели за миграции на животни, редици, пребарувачки алгоритми, многу финансиски модели се основат на нив (Блек Шоулз). Математички, тие поврзуваат веројатност со термодинамика, мартингали и времиња на запирање, и се основа на стохастични диференцијални равенки во инженерство, машинско учење и статистика.

2 Случајно талкање

Случајно талкање е случаен процес кој дефинира последователни чекори на некој математички простор, претставен со:

$$S_n = \sum_{i=1}^{n} X_i, \quad n \geq 1$$

каде X_i се независни и еднакво распределени случајни променливи, такви што

$$\mathbb{E}[X] = \mu, \qquad \mathrm{Var}(X) = \mathbb{D}[X] = \sigma^2.$$

Просто напротив сложено случајно талкање. Во случај на *просто* случајно талкање, вредноста на S_n, од една итерација до друга, секогаш се разликува за некоја фиксна вредност, т.е.

$$|S_{i+1} - S_i| = d$$

или пак, за вредностите на X_i да важи дека

$$\mathbb{P}\big(X_1 \in a + d\mathbb{Z}\big) = 1,$$

каде a е реален број а d е некој коефициент поголем од 0, ова исто така се нарекува Lattice или случајно талкање во решетка, поради тоа што вредностите на S_n можат да се претставaт на решетка, каде вредноста во даден момент може да се помести само на соседните полиња од решетката.

Во случај кога не е просто случајното талкање, X_i може да земе вредности на некој непрекинат интервал, според некоја случајна распределба. Ова би важело за $X_i \sim \mathcal{N}(0,1)$ и би опишало пословен пат.

2.1 Едно димензионално случајно талкање

Карактеристиките на случајно талкање. Ќе го разгледуваме најпростиот пример за случајно талкање, каде секој X_i може да е 1 или -1 со подеднаква веројатност. Ваквото случајно талкање има

$$E(S_n) = \sum_{j=1}^{n} E(X_j) = 0,$$

$$D(S_n) = E(S_n^2) = \sum_{i=1}^{n} E(X_i^2) + 2 \sum_{1 \le i < j \le n} E(X_i X_j) = n.$$

Очекуваната транзлација/поместување, $E|S_n|$, т.е. колку се има поместено вредноста од x-оската, може да се добие на следниот начин. Прво ќе изразиме очекуваната вредност рекурзивно

$$\mathbb{E}|S_{n+1}| = \mathbb{E}\big(\mathbb{E}\big(|S_{n+1}| \,\big|\, S_n\big) \big) = \mathbb{E}\Big[\tfrac{|S_n+1|+|S_n-1|}{2}\Big] = \mathbb{E}|S_n| + \mathbb{P}(S_n = 0),$$

вредноста се менува само ако на претходниот член вредноста е нула, префрлувајќи очекуваните вредности на едната страна, и собирајќи ги рекурзивните формули за сите S_n од 1 до n, се добива следната формула

$$\mathbb{E}|S_n| - \mathbb{E}|S_0| = \sum_{k=0}^{n-1} \mathbb{P}(S_k = 0).$$

Веројатноста $\mathbb{P}(S_k = 0)$, е 0, во случај k да е непарен, а кога k е парен, таа е веројатноста 1 и -1 да се појавиле исто толку често кај променливите X_i,

$$\mathbb{P}(S_{2m} = 0) = \binom{2m}{m} 2^{-2m},$$

според Stirling формулата, $\binom{2m}{m}$, кога m тежи кон бесконечно, после соодветно поедноставување, има вредност $\frac{4^m}{\sqrt{\pi m}}$, со ова добиваме дека,

$$\mathbb{E}|S_n| = \sum_{m \le n/2} \mathbb{P}(S_{2m} = 0) \sim \sum_{m \le n/2} \frac{1}{\sqrt{\pi m}} \approx \frac{1}{\sqrt{\pi}} \cdot 2\sqrt{\tfrac{n}{2}} = \sqrt{\tfrac{2n}{\pi}}\,.$$

Поместувањето, кога n тежи кон бесконечност, е од ред $\sqrt{n}$, со коефициент $\sqrt{\frac{2}{\pi}}$.

За вакво случајно талкање, ако има бесконечно чекори, секоја можна вредност ќе се помине бесконечно многу пати. Нека a и b се два природни броеви, за кои ќе разгледаме веројатностите дека ќе се погоди b пред -a, како и очекуваниот број чекори да се достигне таква состојба.

Веројатноста случајното талкање попрво да добие вредност b При добивање на оваа веројатност, потребно е случајното талкање да е мартингал, т.е. треба да докажеме дека очекуваната вредност на талкањето во некој n+1 момент, при тоа знаејќи ги сите претходни реализации во талкањето, е еднакво на вредноста на случајното талкање во n-тиот момент, доказот е следниот

$$\mathbb{E}[S_{n+1} \mid \mathcal{F}_n] = \mathbb{E}[\ S_n + X_{n+1} \,|\, \mathcal{F}_n] = S_n + \mathbb{E}[X_{n+1} \mid \mathcal{F}_n].$$

Од тоа што X_{n+1} е независен од $\mathcal{F}_n$ и $\mathbb{E}[X_{n+1}] = 0$,

$$\mathbb{E}[S_{n+1} \mid \mathcal{F}_n] = S_n.$$

Имајќи го ова во предвид, ние можеме да го искористиме теоремата за запирање на Дуб. Нека времето на стопирање ни се дефинира со првиот момент во кој S_n добива вредност b или -a, нека $\tilde{S}_n$ биде дефинирана со

$$\tilde{S}_n = \begin{cases} S_n, & n < \tau, \\ -a \ \text{ or } \ b, & n \geq \tau. \end{cases}$$

со ова, $\tilde{S}_n$, го исполнува првиот услов од Теоремата за озапирање на Дуб, од кое следува дека

$$\mathbb{E}\big[\,|\tilde{S}_\tau|\,\big] \;=\; \mathbb{E}\big[\,|\tilde{S}_0|\,\big] \;=\; 0.$$

Нека веројатноста случајното талкање прво да достигне вредност b е p,

$$\mathbb{P}(S_\tau = b) = p, \qquad \mathbb{P}(S_\tau = -a) = 1 - p,$$

$$\mathbb{E}[\tilde{S}_\tau] = p \cdot b + (1-p)(-a) = 0 \quad \Longrightarrow \quad p = \frac{a}{a+b}.$$

Очекуваниот број чекори за S_n да добие вредност b или -a За ова ни треба знаењето дека $S_n^2 - n$ е исто така мартингал, па со помош на погорната теорема на запирање се добива следното

$$\mathbb{E}[S_\tau^2 - \tau] = \mathbb{E}[S_0^2] = 0 \quad \Rightarrow \quad \mathbb{E}[\tau] = \mathbb{E}[S_\tau^2] = b^2 p + a^2(1-p) = ab.$$

Обопштување на веојатностите за погодување и очекуваниот број чекори

Веројатностите на премин се постојани но различни меѓу себе, p!=q За поедноставување на пресметките, нека почетната состојба ни е a, што значи дека бараме веројатноста да се погоди a+b, нека е N, наспроти 0. Прво да дефинираме веројатноста од состојба k т.е. $S_i = k$, да се погоди N пред да се погоди 0,

$$h_k := \mathbb{P}_k(\text{погоди } N \text{ пред } 0).$$

Од ова следува дека $h_0 = 0$ и $h_N = 1$, како и следната рекурентна релација

$$h_k = p h_{k+1} + q h_{k-1}, \quad k = 1, \ldots, N-1,$$

$$q h_{k-1} - h_k + p h_{k+1} = 0.$$

На сличен начин можеме да го претставиме очекуваното време на погодување

$$t_k := \mathbb{E}_k[T], \quad \text{каде } T = \text{време до апсорпција/погодување на } \{0, N\},$$

следува дека $t_0 = 0, t_N = 0$.

$$t_k = 1 + pt_{k+1} + qt_{k-1}, \quad k = 1, \ldots, N-1,$$

$$qt_{k-1} - t_k + pt_{k+1} = -1.$$

Со решавањето на овие системи се добиваат следните формули соодветно

$$h_k = \begin{cases} \dfrac{1 - \left(\frac{q}{p}\right)^k}{1 - \left(\frac{q}{p}\right)^N}, & p \neq q, \\ \dfrac{k}{N}, & p = q = \frac{1}{2}. \end{cases}$$

$$t_k = \frac{N}{p-q} \cdot \frac{\left(\frac{q}{p}\right)^k - 1}{\left(\frac{q}{p}\right)^N - 1} - \frac{k}{p-q} \quad (p \neq q).$$

Веројатностите на премин зависат од вредноста на случајното талкање За овој случај важи следното

$$\exists\, i, j \in \{1, \ldots, N-1\},\ i \neq j \ \text{ такви што } p_i \neq p_j.$$

Според ова, рекурентните релации се следните

$$h_k = p_k h_{k+1} + q_k h_{k-1}, \quad k = 1, \ldots, N-1,$$

$$q_k h_{k-1} - h_k + p_k h_{k+1} = 0.$$

$$t_k = 1 + p_k t_{k+1} + q_k t_{k-1}, \quad k = 1, \ldots, N-1,$$

$$q_k t_{k-1} - t_k + p_k t_{k+1} = -1.$$

Има повеќе алгоритми за решавање на овие равенки, ќе се задржам на тие без матрици, здравје линеарна. За веројатноста на погодување,

$$\text{За } i = 1, \ldots, N-1, \qquad h_i = p_i h_{i+1} + q_i h_{i-1}, \qquad h_0 = 0,\ h_N = 1.$$

$$\Delta_i := h_i - h_{i-1} :$$

$$h_i - p_i h_{i+1} - q_i h_{i-1} = 0 \iff p_i(h_{i+1} - h_i) = q_i(h_i - h_{i-1}) \iff \Delta_{i+1} = \rho_i \Delta_i,$$

$$\text{Каде } \rho_i := \tfrac{q_i}{p_i}.$$

$$\text{Итереативно:} \qquad \Delta_{k+1} = s_k \Delta_1, \qquad s_0 := 1, \quad s_k := \prod_{j=1}^{k} \rho_j.$$

$$\text{Сума на разлики:} \qquad h_i = h_0 + \sum_{k=0}^{i-1} \Delta_{k+1} = \sum_{k=0}^{i-1} s_k \Delta_1.$$

$$\text{Од } h_N = 1 \implies \sum_{k=0}^{N-1} s_k \Delta_1 = 1,$$

$$\boxed{h_i = \frac{\sum_{k=0}^{i-1} s_k}{\sum_{k=0}^{N-1} s_k}, \qquad s_k = \prod_{j=1}^{k} \frac{q_j}{p_j}.}$$

Земнувајќи i со a би добиле бараната вредност. Содветно, за очекуваното време до погодок се добива следното, поради гломазноста на доказот без матрици, скриено е подолу

$$S_m := \sum_{k=0}^{m-1} s_k \ (S_0 = 0),$$

$$\boxed{t_i = \frac{S_i}{S_N} \sum_{j=1}^{N-1} \sum_{k=j}^{N-1} \frac{s_k}{p_j s_j} - \sum_{j=1}^{i-1} \sum_{k=j}^{i-1} \frac{s_k}{p_j s_j}, \qquad s_k = \prod_{r=1}^{k} \frac{q_r}{p_r}.}$$

2.2 Повеќе димензионално случајно талкање

Повеќе димензионалните случајни талкања се дефинирани со

$$S_n = \sum_{i=1}^{n} X_i, \qquad X_i \in \mathbb{Z}^d,$$

каде (X_i) се независни и еднакво распределени случајни променливи. На пример, за $d = 2$ секој чекор може да биде избран од $\{(\pm 1, 0), (0, \pm 1)\}$. Треакторијата на случајно талкање претставува множество од сите постигнати вредности/локации на талкањето. Кај 1Д lattice талкање таа претставува сите достижни вредности помеѓу долната и горната граница, каде кај повеќе димензионалните ни дава повеќе информации за самиот процес.

Табела: Прости случајна талкања на Z^d, рекурентност, веројотност две талкања да се "сретнат"т.е. да имат иста вредност во иста итерација зимајќи во предвид дека се започнале на |x| оддалеченост, и веројатност талкањето да се врати во почетната положба за дадена n-та итерација

d	**Рекурентност?** $P_0(\tau_0 < \infty)$	**Вер. на среќавање**	$P(S_{2n} = 0)$
2	$= 1$ (рекурентна)	$= 1$ (с.с. среќават)	$\sim \frac{1}{\pi n}$
3	≈ 0.340537 (транзитивно)	$= \frac{G_3(x)}{G_3(0)} \sim 0.314870244\lvert x\rvert^{-1}$	$\sim \left(\frac{3}{2\pi n}\right)^{3/2}$
4	≈ 0.193206 (транзитивно)	$= \frac{G_4(x)}{G_4(0)} \sim 0.163490616\lvert x\rvert^{-2}$	$\sim \left(\frac{4}{2\pi n}\right)^{2}$
5	≈ 0.135178 (транзитивно)	$= \frac{G_5(x)}{G_5(0)} \sim 0.109531015\lvert x\rvert^{-3}$	$\sim \left(\frac{5}{2\pi n}\right)^{5/2}$

Дообј. G_d е дискретна Греенова функција $G_d(y) = \sum_{n \geq 0} P_0(S_n = y)$

3 Винеров процес

Винеров процес е непрекинат случаен процес кој се карактеризира со следните правила:

- $W_0 = 0$ со веројатност 1
- W има независни интервали за секое $t > 0$ т.е. интервалите $W_{t+u} - W_t,\ u \geq 0$, се независни од вредности во минатото $W_s,\ s < t$.
- W има Гаусови интевали т.е. $t \geq 0,\ u \geq 0, \quad W_{t+u} - W_t \sim \mathcal{N}(0, u)$
- W е непрекината

Карактеристики на Винеров

$$\mathbb{E}[\,|W_t|\,] = \int_0^\infty \mathbb{P}(|W_t| \geq x)\ dx = 2\int_0^\infty \mathbb{P}(W_t \geq x)\ dx = 2\int_0^\infty \Big(1 - \Phi\big(x/\sqrt{t}\big)\Big)\, dx = \sqrt{\frac{2t}{\pi}}$$

$$W_t \sim \mathcal{N}(0,t), \qquad \mathbb{E}[W_t] = 0, \qquad \mathrm{Var}(W_t) = t, \qquad \mathrm{Cov}(W_s, W_t) = \min\{s,t\}.$$

$$|W_t| \ \text{ има густина } \quad f_{|W_t|}(x) = \sqrt{\frac{2}{\pi t}}\, e^{\left(-\frac{x^2}{2t}\right)}\, \mathbf{1}_{\{x \geq 0\}}, \qquad \mathbb{E}[\,|W_t|\,] = \sqrt{\frac{2t}{\pi}}, \quad \mathrm{Var}(|W_t|) = t\left(1 - \frac{2}{\pi}\right).$$

3.1 Теорема на Донскер

Нека имаме случајно талкање S_n дефинирано со:

$$S_n = \sum_{i=1}^{n} X_i, \quad n \geq 1$$

$$\mathbb{E}[X] = 0, \qquad \mathrm{Var}(X) = \mathbb{D}[X] = \mathbf{1}$$

Ќе го интерполираме S_n така што за секој реален t ќе има вредност,

$$S_t = S_{\lfloor t \rfloor} + (t - \lfloor t \rfloor) X_{\lfloor t \rfloor + 1}, \qquad t \geq 0$$

Понатаму, сакаме да го "убрзаме"времето, т.е. сакаме S_n да е опишано на интервалот $[0,1]$, каде овој интервал ќе го поделиме на n подеднакви подинтервали, за S_n би одговарало $S_{[nt]}$. Дисперзијата на овој процес е n, т.е. со зголемувањето на бројот на чекори, би се зголемила и дисперзијата. За да се избегне ова ќе го нормализираме, нека

$$S_t^{(n)} = \frac{1}{\sqrt{n}}\, S_{\lfloor nt \rfloor}, \qquad 0 \leq t \leq 1$$

Донскеровата теорема гласи дека овој ново дефиниран процес, кога n тежи кон бесконечност, конвергира во дистрибуција кон Винеров процес.

$$S_{[0,1]}^{(n)} \ \Rightarrow \ W_{[0,1]}$$

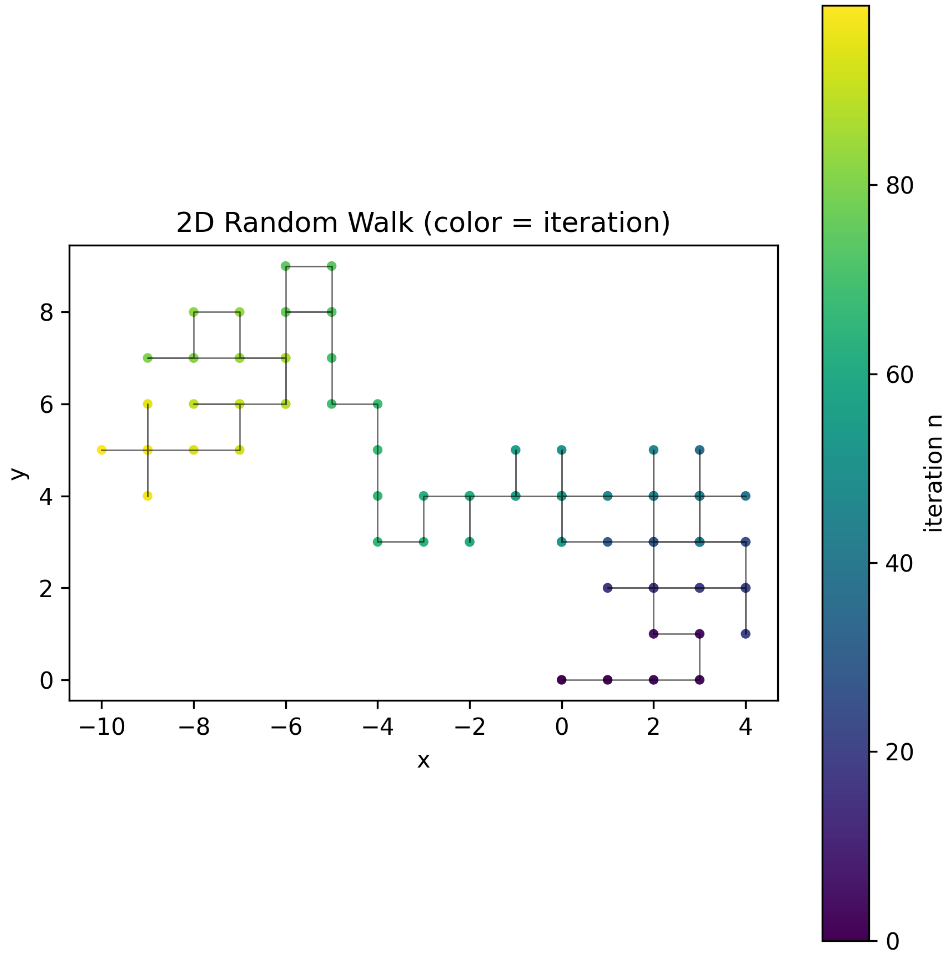
3.2 Принцип на рефлексија

Нека имаме винеров процес, поради непрекинатоста на процесот, секоја реализација/патека на $(0,t)$ што завршува со вредност поголема од a во време t (т.е. $W(t) \geq a$) мора да ја има примено вредноста a во некој $t_a \leq t$ момент за прв пат. За секоја ваква реализација, може да се дефинира $W'(t)$ на $(0,t)$ така што таа е симетрична на оригиналната реализација по оската околу a на интервалот (t_a, t). Овие пак се реализации на Винеров процес, кои поминуваат низ a, но имаат вредност помала од a во t. Од ова следува дека сите реализации кои достигнуваат вредност a на интервалот $(0,t)$, пола ќе завршат со вредност помала од a. Знаејќи го ова, може да се постигне заклучокот дека веројатноста процесот да заврши со вредност поголема од a, е пола од веројатноста да ја достигне a.

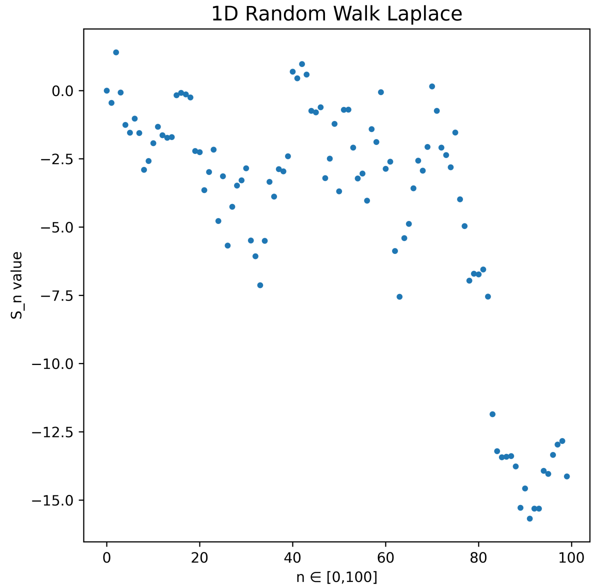
$\mathbb{P}\left(\sup_{0 \leq s \leq t} W(s) \geq a\right) = 2\,\mathbb{P}(W(t) \geq a) = 2\big(1 - \Phi(a/\sqrt{t})\big)$, sup означува максималната вредност на тој интервал

4 Симулации

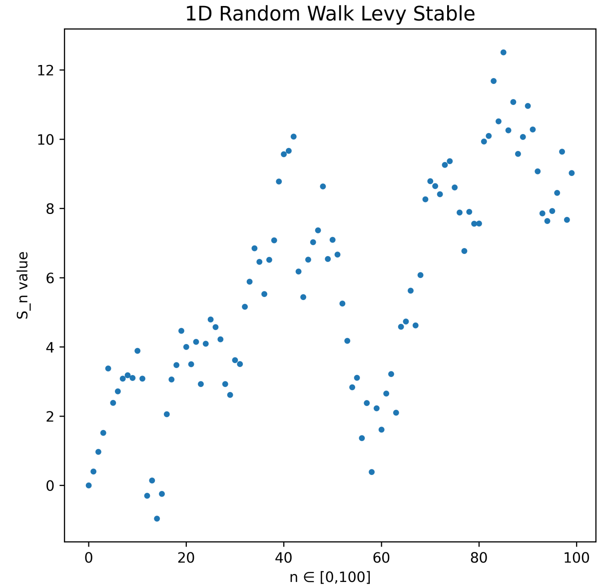
Слика 1: Просто случајно талкање



Слика 2: 2D просто случајно талкање

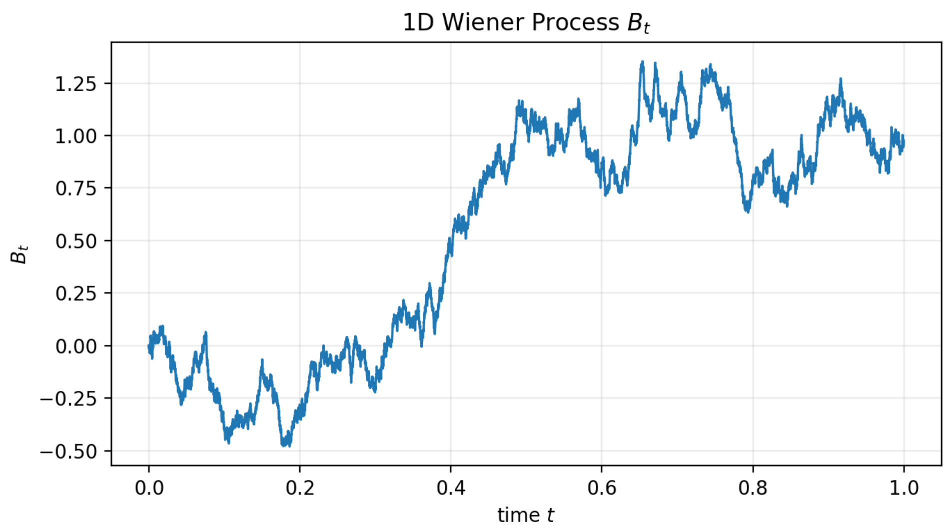


(a) Laplace-distributed increments

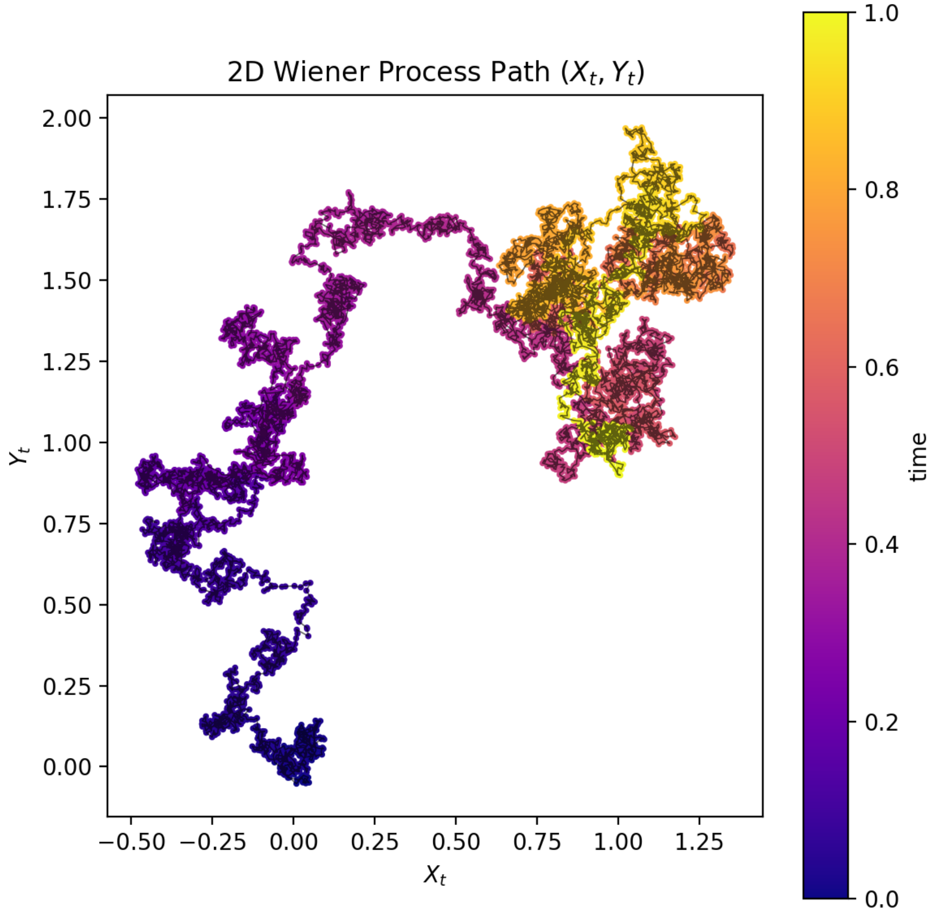


(б) Lévy (heavy-tailed) increments

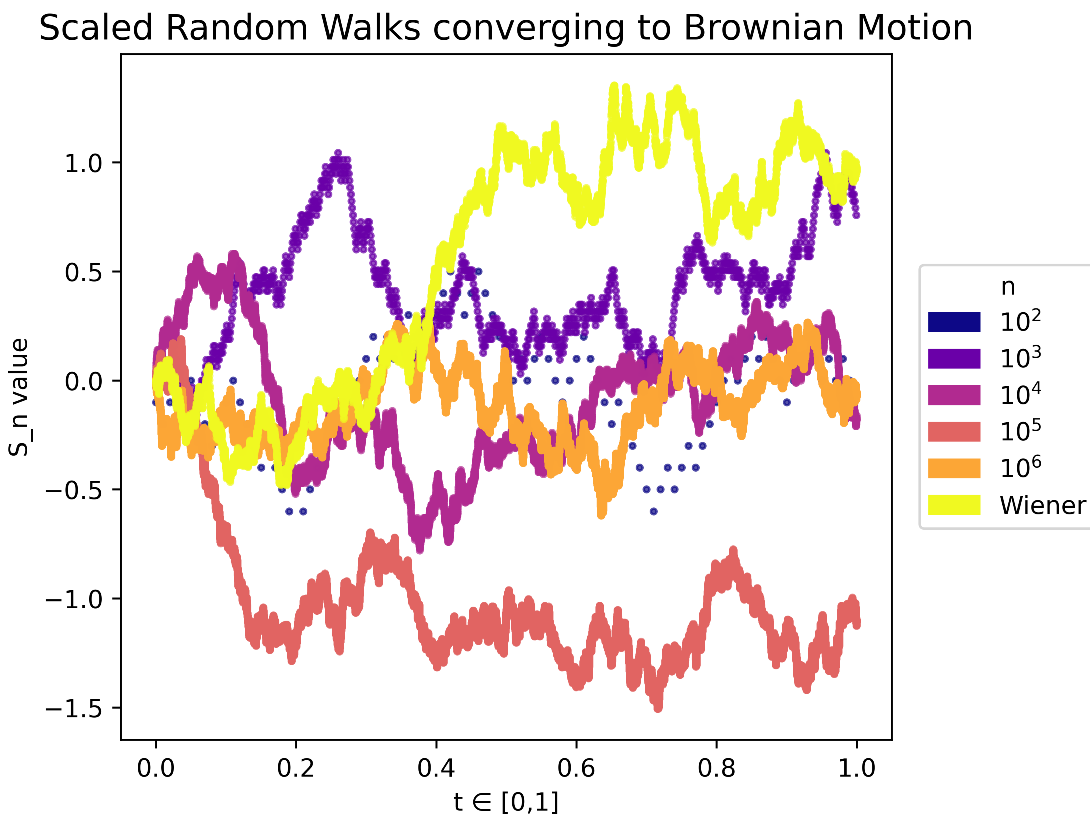
Слика 3: Случајни талкања со различни дистрибуции



Слика 4: Пример патека од 1Д Винеров процес.



Слика 5: Пример патека од 2Д Винеров процес.



Слика 6: Скалирано талкање наспроти Винеров.

A Докази и теореми

Theorem A.1 (Стерлингова апроксимација)**.** *$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$ кога $n \to \infty$.*

Definition A.2 (Мартингал)**.** Мартингал е случаен процес за кои важи дека очекуваната вредност на следниот примерок, имајќи/знаејќи ги вредностите на претходните примероци, е еднаква на вредноста во претходната итерација, т.е.

$$\mathbb{E}[X_{n+1} \mid \mathcal{F}_n] = X_n$$

каде $\mathcal{F}_n$ е множество од претходните реализации т.е. примероците x_1,x_2,...x_n.

Definition A.3. Време на запирање е случајно време во кое престанува набљудувањето на некој случаен процес, основано само на информациите добиени до таа точка

[Време на запирање] Нека $(\mathcal{F}_n)_{n\geq 0}$ е филтрација на веројатносниот простор . Случајна променлива $\tau : \Omega \to \{0, 1, 2, \dots\} \cup \{\infty\}$ се нарекува а *време на запирање* во однос на $(\mathcal{F}_n)$ ако за секој $n \geq 0$,

$$\{\tau = n\} \in \mathcal{F}_n.$$

Еквивалентно, "$\tau \leq n$" се одредува по информациите достапни до време n.

Theorem A.4 (Теоремата на запирање на Дуб)**.** *Нека $(X_n, \mathcal{F}_n)_{n\geq 0}$ е мартингал и τ нека биде време на стопирање во однос на $(\mathcal{F}_n)$. Ако τ и X_n задоволуваат било кој од следните услови:*

(i) *τ е скоро сигурно ограничено, т.е. $\tau \leq c$ за некое*

(ii) *$(X_{n\wedge\tau})$ is uniformly integrable, or*

(iii) *$\mathbb{E}[\tau] < \infty$ и интервалите се ограничени: $|X_{n+1} - X_n| \leq C$,*

тогаш
$$\mathbb{E}[X_\tau] = \mathbb{E}[X_0].$$

Доказ дека $M_n = S_n^2 - n$ е мартингал.
$$M_n := S_n^2 - n$$

$$\mathbb{E}[M_{n+1} \mid \mathcal{F}_n] = M_n \quad \text{a.s.}$$

$$S_{n+1}^2 = (S_n + X_{n+1})^2 = S_n^2 + 2S_n X_{n+1} + X_{n+1}^2$$

$$\mathbb{E}[S_{n+1}^2 \mid \mathcal{F}_n] = S_n^2 + 2S_n\,\mathbb{E}[X_{n+1}] + \mathbb{E}[X_{n+1}^2] = S_n^2 + 1$$

$$\mathbb{E}[M_{n+1} \mid \mathcal{F}_n] = \mathbb{E}\big[S_{n+1}^2 - (n+1) \mid \mathcal{F}_n\big] = (S_n^2 + 1) - (n+1) = S_n^2 - n = M_n$$

□

Доказ за очекуваното време до достигнување на состојба.
$$t_0 = t_N = 0, \qquad t_i = 1 + p_i t_{i+1} + q_i t_{i-1} \quad (i = 1, \dots, N-1).$$

$$\text{Нека } w_i := t_i - t_{i-1}, \qquad \rho_i := \frac{q_i}{p_i}, \qquad s_0 := 1, \quad s_k := \prod_{j=1}^{k} \rho_j \ (k \geq 1).$$

$$t_i - p_i t_{i+1} - q_i t_{i-1} = 1 \iff p_i(t_{i+1} - t_i) = q_i(t_i - t_{i-1}) - 1 \iff w_{i+1} = \rho_i w_i - \frac{1}{p_i}.$$

$$\text{Нека } v_i := \frac{w_i}{s_{i-1}} \quad (i \geq 1).$$

$$v_{i+1} = \frac{\rho_i w_i - 1/p_i}{s_i} = \frac{w_i}{s_{i-1}} - \frac{1}{p_i s_i} = v_i - \frac{1}{p_i s_i},$$

$$v_{k+1} = v_1 - \sum_{j=1}^{k} \frac{1}{p_j s_j}, \qquad w_{k+1} = s_k \left(v_1 - \sum_{j=1}^{k} \frac{1}{p_j s_j} \right).$$

$$S_m := \sum_{k=0}^{m-1} s_k \ (S_0 = 0), \qquad A_m := \sum_{j=1}^{m} \frac{1}{p_j s_j} \ (A_0 = 0).$$

$$t_i = \sum_{k=0}^{i-1} w_{k+1} = \sum_{k=0}^{i-1} s_k v_1 - \sum_{k=0}^{i-1} s_k A_k = S_i v_1 - \sum_{k=0}^{i-1} s_k A_k.$$

$$t_N = 0: \quad 0 = S_N v_1 - \sum_{k=0}^{N-1} s_k A_k \implies v_1 = \frac{\displaystyle\sum_{k=0}^{N-1} s_k A_k}{S_N}.$$

$$\boxed{t_i = \frac{S_i}{S_N} \sum_{k=0}^{N-1} s_k A_k - \sum_{k=0}^{i-1} s_k A_k.}$$

$$\sum_{k=0}^{m-1} s_k A_k = \sum_{k=0}^{m-1} s_k \sum_{j=1}^{k} \frac{1}{p_j s_j} = \sum_{j=1}^{m-1} \frac{1}{p_j s_j} \sum_{k=j}^{m-1} s_k = \sum_{j=1}^{m-1} \sum_{k=j}^{m-1} \frac{s_k}{p_j s_j},$$

$$\boxed{t_i = \frac{S_i}{S_N} \sum_{j=1}^{N-1} \sum_{k=j}^{N-1} \frac{s_k}{p_j s_j} - \sum_{j=1}^{i-1} \sum_{k=j}^{i-1} \frac{s_k}{p_j s_j}, \qquad s_k = \prod_{r=1}^{k} \frac{q_r}{p_r}.}$$

$\square$