Enkel slumpvandring

Sven Erick Alm

9 april 2002
(modifierad 8 mars 2006)

Innehåll

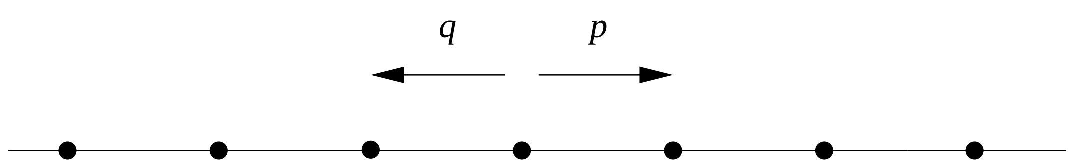
1 Inledning

Med en *slumpvandring* menas allmänt en stokastisk följd $\{S_n\}$, med $S_0 = 0$, definierad som

$$S_n = \sum_{k=1}^{n} X_k,$$

där $\{X_k\}$ är oberoende och likafördelade stokastiska variabler (olf).

Slumpvandringen sägs vara *enkel* om $X_k = \pm 1$, med $P(X_k = 1) = p$ och $P(X_k = -1) = 1 - p = q$. Man kan tänka sig en partikel som utför en slumpvandring på heltalspunkterna på linjen och i varje steg går till en av de två grannpunkterna, se Figur 1.



Figur 1: Enkel slumpvandring

Anm 1. Man kan också studera slumpvandring i flera dimensioner. I två dimensioner har varje punkt 4 grannar och i tre dimensioner finns 6 grannar. □

En enkel slumpvandring sägs vara symmetrisk om partikeln har lika stor sannolikhet att gå till var och en av grannpunkterna.

Allmän slumpvandring behandlas i kapitel 7 i Ross. Här kommer vi endast att studera enkel slumpvandring, huvudsakligen i en dimension.

Vi ska bl.a. försöka besvara följande frågor:

- Vad är sannolikheten att partikeln någonsin når punkten a?
  (Fallet $a = 1$ brukar kallas ”apan och stupet”.)
- Hur lång tid tar det att nå a?
- Vad är sannolikheten att den når $a > 0$ innan den når $-b < 0$? (”Spelarens ruinproblem”)
- Om partikeln vid tidpunkten n befinner sig i $a > 0$, vad är sannolikheten att
  - den befunnit sig till höger om 0 sedan första steget?
  - den aldrig befunnit sig till vänster om 0?

  (”The Ballot problem”)
- Hur långt bort kommer partikeln som längst i n steg?

Vid analys av slumpvandring har man glädje av, dels allmänna metoder, som

- betingning,
- genererande funktioner,
- differensekvationer,

- teorin för Markovkedjor,
- teorin för förgreningsprocesser,
- martingaler,

men även mera speciella, som

- vägräkning,
- spegling,
- tidsomvändning.

2 Apan och stupet

En apa står ett steg från ett stup och tar upprepade gånger, oberoende av varandra, slumpmässigt ett steg framåt, med sannolikhet p, eller ett steg bakåt, med sannolikhet q.

2.1 Passagesannolikheter

Vad är sannolikheten att apan, förr eller senare, ramlar utför stupet?

Kalla denna sannolikhet P_1. Då är

$$P_1 = P(\text{en slumpvandrande partikel någonsin når } x = 1).$$

Vi kan också studera

$$P_k = P(\text{en slumpvandrande partikel någonsin når } x = k),$$

vilket svarar mot att apan startar k steg från stupet.
På grund av oberoendet (och den starka Markovegenskapen) gäller

$$P_k = P_1^k.$$

För att bestämma P_1 kan man betinga med avseende på det första steget.

$$P_1 = p \cdot 1 + q \cdot P_2 = p + q \cdot P_1^2,$$

så att

$$P_1^2 - \frac{1}{q}P_1 + \frac{p}{q} = 0,$$

med lösningar

$$\frac{1}{2q} \pm \sqrt{\frac{1}{4q^2} - \frac{p}{q}} = \frac{1}{2q} \pm \frac{\sqrt{1-4pq}}{2q}.$$

Observera att

$$1 = p + q = (p+q)^2 = p^2 + q^2 + 2pq,$$

så att

$$1 - 4pq = p^2 + q^2 - 2pq = (p-q)^2$$

och alltså

$$\sqrt{1-4pq} = |p-q|.$$

Lösningarna kan alltså skrivas

$$\frac{1 \pm (p-q)}{2q} = \begin{cases} 1, \\ \frac{p}{q}. \end{cases} \tag{1}$$

Om $p > q$ förkastas lösningen $\frac{p}{q} > 1$, så att, för $p \geq q$ (dvs. $p \geq 1/2$), vi får $P_1 = 1$, och därmed $P_k = 1$ för $k \geq 1$.

Det är litet besvärligare att se att för $p < q$ den korrekta lösningen är $P_1 = p/q < 1$, vilket ger $P_k = (p/q)^k$.

Man kan visa detta med hjälp av genererande funktioner, t.ex. genom att utnyttja teorin för förgreningsprocesser, där extinktionssannolikheten är den minsta positiva roten till ekvationen $g(s) = s$; se Problem 2.

Ett mera direkt sätt är att studera

$$P_k(n) = P(\text{att nå } x = k \text{ i de } n \text{ första stegen}).$$

Då gäller nämligen att

$$P_1(n) = p + q \cdot P_2(n-1) \leq p + q \cdot P_1^2(n-1).$$

Eftersom $P_1(1) = p \leq p/q$, kan vi, med induktion, visa att $P_1(n) \leq p/q$ för alla $n \geq 1$, om vi kan visa att $P_1(n) \leq p/q$ medför att även $P_1(n+1) \leq p/q$. Antag därför att $P_1(n) \leq p/q$. Då gäller

$$P_1(n+1) \leq p + q \cdot P_1^2(n) \leq p + q \cdot (p/q)^2 = p + \frac{p^2}{q} = \frac{p}{q}.$$

Eftersom

$$P_1 = \lim_{n \to \infty} P_1(n) \leq \frac{p}{q}$$

kan vi, för $p < q$, förkasta lösningen $P_1 = 1$.

Vi har alltså visat

Sats 1. Med P_k som ovan gäller, för $k \geq 1$,

$$P_k = \begin{cases} 1 & \text{om } p \geq q, \\ (\frac{p}{q})^k & \text{om } p < q. \end{cases}$$

Anm 2. Detta betyder att en symmetrisk slumpvandring, med sannolikhet 1, kommer att besöka **alla** punkter på linjen! □

Problem 1. Låt $p < q$. Bestäm fördelningen för $Y = \max_n S_n$. Vad blir $E(Y)$?
Ledning: Studera $P(Y \geq a)$. □

Problem 2. Låt $p < q$. Visa att P_1 är extinktionssannolikheten i en förgreningsprocess med reproduktionsfördelning $p_0 = p$, $p_2 = q$. □

2.2 Passagetider

Hur lång tid tar det tills apan ramlar utför stupet?

Låt $T_{jk} =$ ”tiden det tar att gå från $x = j$ till $x = k$”, så att
$T_{0k} =$ ”tiden tills partikeln når $x = k$ för första gången (vid start i $x = 0$)”,
och låt vidare

$$E_k = E(T_{0k}),$$

om väntevärdet existerar. I så fall måste gälla, för $k > 0$,

$$E_k = k \cdot E_1.$$

Betingning ger

$$E_1 = 1 + p \cdot 0 + q \cdot E_2 = 1 + 2q \cdot E_1.$$

Vi skiljer på fallen $p < q$, $p = q$ och $p > q$.

För $p < q$ gäller, enligt Sats 1, att $P(T_{01} = \infty) = 1 - P_1 > 0$, vilket medför att $E_1 = +\infty$.

För $p = q = 1/2$ får vi, om E_1 antas vara ändligt, att

$$E_1 = 1 + E_1,$$

dvs. en motsägelse, så även i det fallet är $E_1 = +\infty$.

Slutligen, för $p > q$, får vi

$$E_1 = \frac{1}{1 - 2q} = \frac{1}{p - q} < \infty.$$

Vi har alltså visat

Sats 2. För $k \geq 1$ gäller

$$E_k = \begin{cases} +\infty & \text{om } p \leq q, \\ \frac{k}{p-q} & \text{om } p > q. \end{cases}$$

Med hjälp av Sats 1 och 2 kan vi också studera återkomster till startpunkten. Låt

$$\begin{aligned} P_0 &= P(\text{att partikeln någonsin kommer tillbaka till startpunkten}), \\ T_{00} &= \text{"tiden tills första återkomsten"}, \\ E_0 &= E(T_{00}). \end{aligned}$$

Då gäller

Sats 3. $E_0 = +\infty$ för alla p och

$$P_0 = \begin{cases} 1 & \text{om } p = q, \\ 1 - |p - q| & \text{om } p \neq q. \end{cases}$$

Bevis: Med beteckningar som tidigare får vi genom betingning att

$$P_0 = p \cdot P_{-1} + q \cdot P_1.$$

Om $p = q$ gäller, enligt Sats 1, $P_{-1} = P_1 = 1$, så att även $P_0 = 1$. Om $p \neq q$ kommer endera $P_{-1} < 1$ eller $P_1 < 1$, så att $P_0 < 1$.

I fallet $p < q$ är $P_{-1} = 1$ och $P_1 = p/q$, så att

$$P_0 = p + q \cdot \frac{p}{q} = 2p = 1 - (q - p) < 1.$$

I fallet $p > q$ gäller i stället $P_{-1} = q/p$ och $P_1 = 1$, så att

$$P_0 = p \cdot \frac{q}{p} + q = 2q = 1 - (p - q) < 1.$$

För $p \neq q$ gäller alltså $P_0 = 1 - |p - q| < 1$ och följdaktligen är $P(T_{00} = \infty) > 0$, så att $E_0 = +\infty$.
Om $p = q$ får vi

$$E_0 = 1 + \frac{1}{2} E_{-1} + \frac{1}{2} E_1 = +\infty,$$

enligt Sats 2. □

Anm 3. Den symmetriska slumpvandringen kommer alltså, med sannolikhet 1, tillbaka till 0. Detta gäller efter varje återkomst, så att

$$P(S_n = 0 \text{ i.o.}) = 1.$$

Trots detta är väntevärdet för tiden mellan två återkomster oändligt! Stora talens lag kan alltså inte tolkas som att partikeln i allmänhet är nära 0. I själva verket är partikeln sällan nära 0 och en stor del av tiden långt från 0, även i en symmetrisk slumpvandring! Se Avsnitt 4.5. □

Anm 4. Man kan visa att även den symmetriska slumpvandringen i två dimensioner återkommer till origo med sannolikhet 1, medan däremot i tre dimensioner sannolikheten är ≈ 0.35. □

Problem 3. Låt $p \neq q$. Bestäm fördelningen för $Y =$ # återkomster till 0. Beräkna $E(Y)$. □

3 Spelarens ruinproblem

Apan och stupet kan uppfattas som att vi placerat en absorberande barriär i $x = 1$ (eller $x = k$). Om man studerar två absorberande barriärer på var sin sida om startpunkten kan man lösa *Spelarens ruinproblem* (Gambler's ruin):

Två spelare, A och B, spelar ett spel med oberoende omgångar där, i varje omgång, en av spelarna vinner 1 krona av motståndaren; A med sannolikhet p och B med sannolikhet $q = 1-p$. A börjar spelet med a kronor och B med b kronor. Spelet slutar när någon av spelarna blivit ruinerad.

3.1 Absorptionssannolikheter

Vad blir spelarnas ruinsannolikheter?

Detta motsvarar en slumpvandring där partikeln startar i 0 och absorberas i tillstånden b och $-a$, eller, ekvivalent, startar i a och absorberas i 0 och $a + b$.
Låt

$$A_k = P(\text{A vinner då han har } k \text{ kronor}).$$

Då gäller $A_0 = 0$, $A_{a+b} = 1$ och vi söker A_a. Betinga!

$$A_k = p \cdot A_{k+1} + q \cdot A_{k-1}. \tag{2}$$

Denna homogena differensekvation kan man lösa genom att bestämma nollställena till det karakteristiska polynomet

$$z = p \cdot z^2 + q \Leftrightarrow z^2 - \frac{1}{p} \cdot z + \frac{q}{p} = 0,$$

med lösningar $z_1 = 1$ och $z_2 = q/p$. (Jämför med (1).) Detta ger, för $p \neq q$, följande allmänna lösning till (2)

$$A_k = C_1 \cdot 1^k + C_2 \cdot (\frac{q}{p})^k,$$

där konstanterna C_1 och C_2 bestäms av randvillkoren.

$$A_0 = 0 \Rightarrow C_1 + C_2 = 0,$$
$$A_{a+b} = 1 \Rightarrow C_1 + C_2(\frac{q}{p})^{a+b} = 1,$$

så att

$$C_1 = \frac{-1}{(\frac{q}{p})^{a+b} - 1},$$
$$C_2 = \frac{1}{(\frac{q}{p})^{a+b} - 1},$$

och

$$A_k = \frac{-1}{(\frac{q}{p})^{a+b} - 1} + \frac{1}{(\frac{q}{p})^{a+b} - 1}(\frac{q}{p})^k = \frac{(\frac{q}{p})^k - 1}{(\frac{q}{p})^{a+b} - 1},$$
$$A_a = \frac{(\frac{q}{p})^a - 1}{(\frac{q}{p})^{a+b} - 1}.$$

För $p = q$ får vi differensekvationen

$$A_k = \frac{1}{2}A_{k+1} + \frac{1}{2}A_{k-1}.$$

Det karakteristiska polynomet $z^2 - 2z + 1$ har dubbelroten $z_1 = z_2 = 1$. Vi behöver därför ytterligare en lösning. $A_k = k$ fungerar, så att

$$A_k = C_1 \cdot 1^k + C_2 \cdot k.$$

$$A_0 = 0 \Rightarrow C_1 = 0,$$
$$A_{a+b} = 1 \Rightarrow C_2 = \frac{1}{a+b},$$

så att

$$A_k = \frac{k}{a+b},$$
$$A_a = \frac{a}{a+b}.$$

Vi har alltså visat

Sats 4. Sannolikheten att A ruinerar B (partikeln absorberas i $x = b$) är

$$A_a = \begin{cases} \dfrac{(\frac{q}{p})^a - 1}{(\frac{q}{p})^{a+b} - 1} & \text{om } p \neq \frac{1}{2}, \\[2ex] \dfrac{a}{a+b} & \text{om } p = \frac{1}{2}. \end{cases}$$

Anm 5. $a = +\infty$, $b = 1$ svarar mot *Apan och stupet*. $P_1 = P(\text{apan faller utför stupet}) = \lim_{a\to\infty} P(\text{A vinner})$. För $p = q$ får vi

$$P_1 = \lim_{a\to\infty} \frac{a}{a+1} = 1$$

och för $p \neq q$

$$P_1 = \lim_{a\to\infty} \frac{(\frac{q}{p})^a - 1}{(\frac{q}{p})^{a+1} - 1} = \begin{cases} 1 & \text{om } p > q, \\ \frac{p}{q} & \text{om } p < q. \end{cases}$$

□

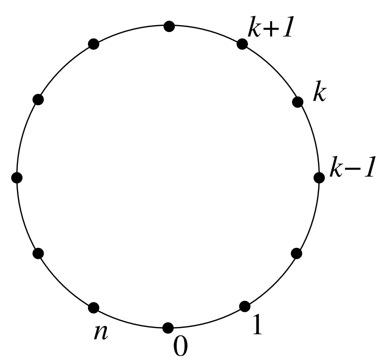
Exempel 1. Betrakta en symmetrisk slumpvandring på punkterna $0, 1, \ldots, n$, belägna på periferin till en cirkel; se Figur 2. Slumpvandringen startar i punkten 0. Man inser lätt att, med sannolikhet 1, alla punkter kommer att besökas. (Varför?)

Vad är sannolikheten att punkten k ($k = 1, \ldots, n$) är den sista som besöks?

Innan k besöks måste någon av $k-1$ och $k+1$ besökas. Betrakta tidpunkten då detta sker för första gången. På grund av symmetrin kan vi anta att det då är $k-1$ som besöks. Att k är den sista punkten som besöks betyder att $k+1$ måste besökas före k och detta kan bara ske genom att slumpvandringen går medsols från $k-1$ till $k+1$ innan den besöker k. Sannolikheten för detta är densamma som ruinsannolikheten för en spelare som har $n-1$ kronor och möter en motståndare med 1 krona, dvs. $1/n$. Vi har alltså visat det överraskande resultatet att

$$P(k \text{ besöks sist}) = \frac{1}{n} \text{ för } k = 1, \ldots, n.$$

□



Figur 2: Slumpvandring på en cirkel

Exempel 2. Betrakta en symmetrisk slumpvandring och ett tillstånd $a > 0$.
Låt Y_a = "# besök i a innan slumpvandringen återänder till 0". För att a över huvud taget ska besökas måste första steget gå åt höger, så att $P(Y_a > 0) = \frac{1}{2} \cdot P(Y_a > 0|S_1 = 1)$. Den betingade sannolikheten är vinstchansen för en spelare med 1 krona som spelar mot en motståndare med $a-1$ kronor, dvs. $P(Y_a > 0|S_1 = 1) = \frac{1}{a}$, så att $P(Y_a > 0) = \frac{1}{2a}$. På liknande sätt kan vi beräkna $P(Y_a = 1|Y_a > 0)$. Betrakta nämligen slumpvandringen då a besöks, vilket vi vet inträffar om $Y_a > 0$. För att detta ska vara det sista besöket i a, före nästa besök i 0, krävs att det första steget går åt vänster, så att $P(Y_a = 1|Y_a > 0) = \frac{1}{2} \cdot P(0 \text{ nås före } a \text{ vid start i } a-1) =$

$\frac{1}{2} \cdot \frac{1}{a} = \frac{1}{2a}$. Samma situation uppstår vid varje besök i a, så att $Y_a | Y_a > 0$ är ffg$(1/2a)$ och $P(Y_a > 0) = 1/2a$. Detta ger att

$$E(Y_a) = \frac{1}{2a} \cdot 2a = 1$$

för alla $a > 0$! (Detta gäller på grund av symmetrin även för negativa a.)

Vi har alltså visat något mycket överraskande:

Mellan två besök i 0 kommer den symmetriska slumpvandringen att göra i genomsnitt 1 besök i **alla** *andra tillstånd!*

För att detta ska vara möjligt måste rimligen $E_0 = \infty$, vilket också visades i Sats 3. □

Problem 4. Utnyttja Sats 1 för att visa att, för alla p, a och b, spelet kommer att ta slut med sannolikhet 1. □

Problem 5. Antag att Du har 10 kronor och din motståndare har 100 kronor. Du får chansen att välja att spela med insatsen 1, 2, 5 eller 10 kronor per omgång. Hur ska Du välja, och vad blir dina vinstchanser, om din vinstchans i ett enskilt spel är
a) $p = 0.5$, b) $p = 0.4$, c) $p = 0.8$. □

3.2 Absorptionstider

Hur lång tid tar det innan någon blir ruinerad?

Låt

$$Y_k = \text{"\# återstående spelomgångar då A har } k \text{ kronor"},$$
$$E_k = E(Y_k).$$

Betingning ger då

$$E_k = 1 + p \cdot E_{k+1} + q \cdot E_{k-1}, \tag{3}$$

med

$$E_0 = E_{a+b} = 0.$$

Ekvation (3) är en icke-homogen differensekvation. För att lösa den behöver vi dels lösa motsvarande homogena ekvation, som ovan, men också hitta en partikulärlösning till den inhomogena.

Vi börjar med det symmetriska fallet ($p = q = 1/2$). Differensekvationen är då

$$E_k = 1 + \frac{1}{2} \cdot E_{k+1} + \frac{1}{2} \cdot E_{k-1},$$

med homogen lösning $A + B \cdot k$. En partikulärlösning är $-k^2$, så att den allmänna lösningen är

$$E_k = A + B \cdot k - k^2.$$

Randvillkoren ger $A = 0$ och $B = a + b$, så att

$$E_k = k \cdot (a + b - k),$$
$$E_a = a \cdot b.$$

I det asymmetriska fallet är den homogena lösningen, som tidigare, $A + B \cdot (q/p)^k$ och en partikulärlösning ges av $k/(q-p)$, vilket ger den allmänna lösningen

$$E_k = \frac{k}{q-p} + A + B \cdot (\frac{q}{p})^k. \tag{4}$$

Genom att utnyttja randvillkoren kan vi bestämma

$$\begin{aligned} A &= \frac{a-b}{(q-p)((\frac{q}{p})^{a+b}-1)}, \\ B &= -A. \end{aligned}$$

Insättning i (4) ger

Sats 5. Förväntade antalet spelomgångar tills någon av spelarna ruinerats är

$$E_a = \begin{cases} a \cdot b & \text{om } p = q = \frac{1}{2}, \\ \dfrac{a}{q-p} - \dfrac{a+b}{q-p} \cdot \dfrac{(\frac{q}{p})^a - 1}{(\frac{q}{p})^{a+b}-1} & \text{om } p \neq q. \end{cases}$$

Exempel 3. I ett rättvist spel med $a + b = 100$ får vi

a	b	A_a	E_a
1	99	1/100	99
2	98	1/50	196
5	95	1/20	475
10	90	1/10	900
20	80	1/5	1600
25	75	1/4	1875
50	50	1/2	2500

□

Problem 6. Beräkna motsvarande tabell som i Exempel 3 då
a) $p = 0.6$, b) $p = 0.2$. □

Problem 7. För vilket värde på a maximeras/minimeras E_a när $a + b = 100$ och
a) $p = 0.4$, b) $p = 0.8$. □

Problem 8. Betrakta slumpvandringen i Exempel 1 och låt $E_k = \#$ steg tills $x = k$ besöks för första gången. Beräkna E_k för $k = 1, \ldots, n$ då
a) $n = 2$, b) $n = 3$, c) $n = 4$. □

Problem 9. Låt i Exempel 1, $P(\text{motsols}) = p$ och $P(\text{medsols}) = q = 1 - p$. Visa att, för alla $0 \leq p \leq 1$, med sannolikhet 1, alla punkter kommer att besökas. □

3.3 Reflekterande barriärer

Vi har hitills studerat absorberande barriärer. Man kan också tänka sig *reflekterande barriärer*. Att a är en reflekterande barriär innebär att så fort man kommer till $x = a$ återvänder man i nästa steg med sannolikhet 1 till den föregående positionen. Med två reflekterande barriärer i a och b, $a < b$, finns bara $b - a + 1$ tänkbara tillstånd, och alla kommer att besökas oändligt många gånger (om $0 < p < 1$).

Exempel 4. Låt 0 vara en reflekterande barriär, dvs. $P(S_{n+1} = 1 | S_n = 0) = 1$. Betrakta en symmetrisk slumpvandring med $S_0 = 0$ och låt

$$E_{j,k} = E(\text{tiden att gå från } j \text{ till } k).$$

Då gäller

$$\begin{aligned} E_{0,1} &= 1, \\ E_{1,2} &= 1 + \frac{1}{2} \cdot E_{0,2} = 1 + \frac{1}{2} \cdot (E_{0,1} + E_{1,2}) = \frac{3}{2} + \frac{1}{2} \cdot E_{1,2}, \end{aligned}$$

så att $E_{1,2} = 3$ och $E_{0,2} = 1 + 3 = 4$. På samma sätt ser vi att

$$E_{2,3} = 1 + \frac{1}{2} \cdot E_{1,3} = 1 + \frac{1}{2} \cdot (E_{1,2} + E_{2,3}) = \frac{5}{2} + \frac{1}{2} \cdot E_{2,3},$$

så att $E_{2,3} = 5$ och $E_{0,3} = 4 + 5 = 9$. Det gäller uppenbarligen, för $k = 1, 2, 3$, att

$$\begin{aligned} E_{k-1,k} &= 2k - 1 \\ E_{0,k} &= k^2. \end{aligned}$$

Antag att detta gäller för $k \leq n$. Då får vi

$$\begin{aligned} E_{n,n+1} &= 1 + \frac{1}{2} \cdot E_{n-1,n+1} = 1 + \frac{1}{2} \cdot (E_{n-1,n} + E_{n,n+1}) = 1 + \frac{2n-1}{2} + \frac{1}{2} \cdot E_{n,n+1} \\ &= \frac{2n+1}{2} + \frac{1}{2} \cdot E_{n,n+1} = 2n + 1, \\ E_{0,n+1} &= E_{0,n} + E_{n,n+1} = n^2 + 2n + 1 = (n+1)^2. \end{aligned}$$

Vi har alltså visat att, för $n \geq 1$,

$$\begin{aligned} E_{n-1,n} &= 2n - 1 \\ E_{0,n} &= n^2. \end{aligned}$$

□

Problem 10. Bestäm rekursiva uttryck för $E_{n-1,n}$ och $E_{0,n}$ i det allmänna fallet ($0 < p < 1$). Visa att $E_{0,n} < \infty$ för alla $n \geq 1$. Vad blir $E_{0,2}$ uttryckt i p? □

4 Vägräkning

Inför händelserna

$$\begin{aligned} F_n &= \text{"partikeln befinner sig i 0 efter } n \text{ steg"}, \\ G_n &= \text{"partikeln återkommer till 0 för första gången efter } n \text{ steg"} \end{aligned}$$

och motsvarande sannolikheter

$$f_n = P(F_n),$$
$$g_n = P(G_n).$$

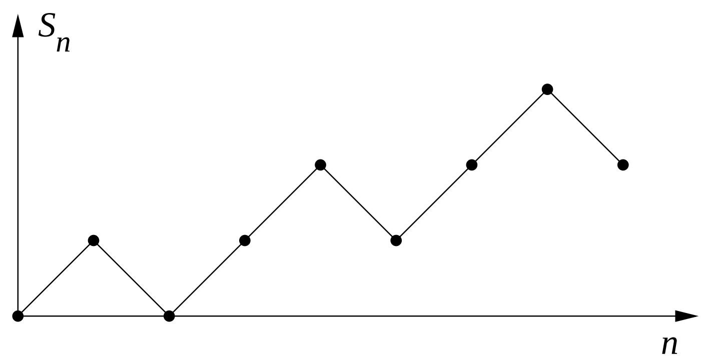
Observera att partikeln bara kan återvända efter ett jämnt antal steg så att

$$f_{2n+1} = g_{2n+1} = 0.$$

För jämna n måste partikeln ha tagit lika många steg åt höger och åt vänster, så att

$$f_{2n} = \binom{2n}{n} p^n q^n.$$

Ett sätt att bestämma g_n är att räkna vägar. Detta underlättas om man illustrerar slumpvandringen genom att plotta talparen (n, S_n), se Figur 3.



Figur 3: Plottad slumpvandring

Anm 6. Här anges tiden längs x-axeln och partikelns förflyttningar till höger och vänster svarar mot steg uppåt och nedåt i figuren. □

Observera att varje väg som innehåller h steg åt höger (uppåt) och v steg åt vänster (nedåt) har sannolikhet $p^h \cdot q^v$, så att det ofta räcker att räkna antalet sådana för att bestämma önskade sannolikheter.

Om vi t.ex. vill bestämma

$$f_n(a, b) = P(\text{partikeln går från } a \text{ till } b \text{ i } n \text{ steg}),$$

så gäller det att

$$h + v = n \quad \text{och} \quad h - v = b - a,$$

dvs.

$$h = \frac{n + b - a}{2} \quad \text{och} \quad v = \frac{n + a - b}{2}, \tag{5}$$

så att

$$f_n(a, b) = \binom{n}{h} p^h q^v,$$

om vi definierar

$$\binom{n}{h} = 0$$

när h inte är ett heltal.

Problem 11. Utnyttja Stirlings formel

$$n! \sim \sqrt{2\pi n}\, n^n\, e^{-n}$$

för att visa att, för $p = q = 1/2$, gäller

a) $f_{2n} \to 0$ när $n \to \infty$, b) $\sum_{n=1}^{\infty} f_{2n} = \infty$, c) $f_{2n} \sim \frac{1}{\sqrt{\pi n}}$.

Ledning: CGS kan vara användbar! □

Problem 12. Vad blir f_{2n} för en symmetrisk tvådimensionell slumpvandring? □

4.1 Spegling

För att bestämma g_n inför vi

$$\begin{aligned} N_n(a,b) &= \#\text{ vägar från } a \text{ till } b \text{ i } n \text{ steg}, \\ N_n^{\neq 0}(a,b) &= \#\text{ vägar från } a \text{ till } b \text{ i } n \text{ steg som inte besöker 0 på vägen}, \\ N_n^{0}(a,b) &= \#\text{ vägar från } a \text{ till } b \text{ i } n \text{ steg som besöker 0 på vägen}. \end{aligned}$$

Observera att, med h och v som i (5), följande gäller

$$\begin{aligned} N_n(a,b) &= \binom{n}{h}, \\ N_n(a,b) &= N_n^0(a,b) + N_n^{\neq 0}(a,b), \\ N_n^{\neq 0}(a,b) &= 0 \quad \text{om } a \text{ och } b \text{ har olika tecken}, \\ N_n^0(a,b) &= N_n(a,b) \quad \text{om } a \text{ och } b \text{ har olika tecken}. \end{aligned}$$

För att beräkna

$$g_{2n} = N_{2n}^{\neq 0}(0,0) \cdot p^n \cdot q^n$$

räcker det alltså att bestämma $N_{2n}^{\neq 0}(0,0)$. Notera att

$$N_{2n}^{\neq 0}(0,0) = N_{2n-1}^{\neq 0}(1,0) + N_{2n-1}^{\neq 0}(-1,0) = 2 \cdot N_{2n-1}^{\neq 0}(1,0) = 2 \cdot N_{2n-2}^{\neq 0}(1,1).$$

För att bestämma $N_{2n-2}^{\neq 0}(1,1)$ kan man använda följande eleganta speglingsresonemang; se Figur 4.

Varje väg från 1 till 1 som besöker 0 kan, genom spegling av början av vägen, fram till första 0-besöket, överföras i en väg från -1 till 1 med lika många steg.

Alla vägar från -1 till 1 måste besöka 0!

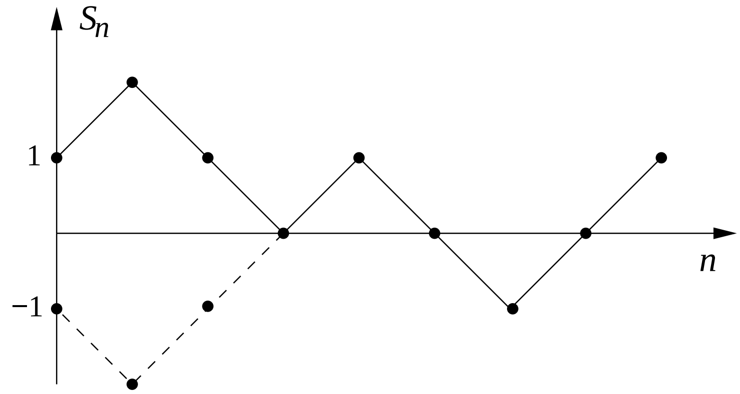
Följaktligen gäller

$$\begin{aligned} N_{2n-2}^{0}(1,1) &= N_{2n-2}^{0}(-1,1) = N_{2n-2}(-1,1), \\ N_{2n-2}^{\neq 0}(1,1) &= N_{2n-2}(1,1) - N_{2n-2}^{0}(1,1) = N_{2n-2}(1,1) - N_{2n-2}(-1,1) \\ &= \binom{2n-2}{n-1} - \binom{2n-2}{n} = \binom{2n-2}{n-1}\left(1 - \frac{n-1}{n}\right) = \frac{1}{n}\binom{2n-2}{n-1}. \end{aligned}$$

Detta ger

$$N_{2n}^{\neq 0}(0,0) = 2 \cdot N_{2n-2}^{\neq 0}(1,1) = 2\frac{1}{n}\binom{2n-2}{n-1} = \frac{1}{2n-1}\binom{2n}{n},$$

så att vi har visat



Figur 4: Spegling

Sats 6. För $n > 0$ gäller

$$g_{2n} = \frac{1}{2n-1} f_{2n}.$$

Speglingsprincipen gäller allmännare än vad vi har utnyttjat.

Sats 7. (**Speglingssatsen**) För $a > 0$ och $b > 0$ gäller

$$N_n^0(a,b) = N_n(-a,b),$$
$$N_n^{\neq 0}(a,b) = N_n(a,b) - N_n(-a,b).$$

Bevis: Visas på samma sätt som ovan. $\square$

Anm 7. Speglingssatsen gäller även om n har fel paritet i förhållande till a och b eftersom alla uttryck i satsen då är lika med 0. $\square$

4.2 Ballot-problemet

Om $a = 0$ i Speglingssatsen får vi

Sats 8. (**Ballot-satsen**) För $b > 0$ gäller

$$N_n^{\neq 0}(0,b) = \frac{b}{n} \cdot N_n(0,b).$$

Bevis: Låt $h = \frac{n+b}{2}$ och $v = \frac{n-b}{2}$ vara antalet steg till höger resp. vänster som krävs för att gå från 0 till b i n steg. Då gäller $N_n(0,b) = \binom{n}{h}$ och

$$\begin{aligned} N_n^{\neq 0}(0,b) &= N_{n-1}^{\neq 0}(1,b) = N_{n-1}(1,b) - N_{n-1}(-1,b) \\ &= \binom{n-1}{h-1} - \binom{n-1}{h} = \binom{n}{h} \cdot \left(\frac{h}{n} - \frac{v}{n}\right) = \binom{n}{h} \cdot \frac{b}{n}. \end{aligned}$$

$\square$

Satsen har fått sitt namn från följande klassiska sannolikhetsproblem, *The Ballot Problem*:

Exempel 5. *I ett val med två kandidater får kandidat A totalt a röster och kandidat B totalt b röster, där $a > b$. Vad är sannolikheten att A är i ledningen under hela röstsammanräkningen?*
Om vi antar att rösterna räknas i slumpmässig ordning ges svaret direkt av Ballot-satsen:

$$P(\text{A leder hela tiden}) = \frac{N_{a+b}^{\neq 0}(0, a-b)}{N_{a+b}(0, a-b)} = \frac{a-b}{a+b}.$$

□

Exempel 6. En variant av Ballot-problemet är:
Vad är sannolikheten, då $a \geq b$, att B aldrig leder under sammanräkningen?
Vi vill alltså veta

$$N_{a+b}^{\neq(-1)}(0, a-b) = N_{a+b}^{\neq 0}(1, a+1-b).$$

Enligt speglingssatsen gäller

$$\begin{aligned} N_{a+b}^{\neq 0}(1, a+1-b) &= N_{a+b}(1, a+1-b) - N_{a+b}(-1, a+1-b) \\ &= \binom{a+b}{a} - \binom{a+b}{a+1} = \binom{a+b}{a} \cdot (1 - \frac{b}{a+1}) \\ &= \binom{a+b}{a} \cdot \frac{a+1-b}{a+1} = \frac{a+1-b}{a+1} \cdot N_{a+b}(0, a-b), \end{aligned}$$

så att

$$P(\text{B aldrig leder}) = \frac{a+1-b}{a+1}.$$

Speciellt gäller att, om $a = b$, så är

$$P(\text{B aldrig leder}) = \frac{1}{a+1}.$$

□

Anm 8. Ballot-problemen är rent kombinatoriska och svaren beror inte av p (och q). Detsamma gäller vid slumpvandring om vi betingar med avseende på slutpunkten S_n. I själva verket kan vi skriva Ballot-problemen som

$$\begin{aligned} P(S_k > 0,\ \text{för } k = 1, \ldots, a+b-1 | S_{a+b} = a-b) &= \frac{a-b}{a+b}, \\ P(S_k \geq 0,\ \text{för } k = 1, \ldots, a+b-1 | S_{a+b} = a-b) &= \frac{a+1-b}{a+1}. \end{aligned}$$

□

4.3 Rekurrens

För nästa resultat behöver vi ytterligare några beteckningar. Vi utgår som tidigare från att $S_0 = 0$.

$$\begin{aligned} N_n^{\neq 0} &= \#\ \text{vägar av längd } n \text{ med } S_k \neq 0 \text{ för } k = 1, \ldots, n, \\ N_n^{>0} &= \#\ \text{vägar av längd } n \text{ med } S_k > 0 \text{ för } k = 1, \ldots, n. \end{aligned}$$

Sats 9. För $n > 0$ gäller

$$N_{2n}^{\neq 0} = N_{2n}(0,0) = \binom{2n}{n}.$$

För den symmetriska slumpvandringen gäller

$$\begin{aligned}
(i) \quad & P(S_k \neq 0,\, k = 1, \ldots, 2n) = P(S_{2n} = 0), \\
(ii) \quad & P(S_k > 0,\, k = 1, \ldots, 2n) = \frac{1}{2} \cdot P(S_{2n} = 0), \\
(iii) \quad & P(S_k \geq 0,\, k = 1, \ldots, 2n) = P(S_{2n} = 0).
\end{aligned}$$

Bevis: En väg som aldrig besöker 0 är endera hela tiden positiv eller alltid negativ.

$$N_{2n}^{\neq 0} = 2 \cdot N_{2n}^{>0} = 2 \cdot \sum_{r=1}^{n} N_{2n}^{\neq 0}(0, 2r).$$

Med samma resonemang som vid beviset av Speglingssatsen får vi, för $r > 0$,

$$N_{2n}^{\neq 0}(0, 2r) = N_{2n-1}(1, 2r) - N_{2n-1}(-1, 2r) = N_{2n-1}(0, 2r-1) - N_{2n-1}(0, 2r+1),$$

så att

$$\begin{aligned}
\sum_{r=1}^{n} N_{2n}^{\neq 0}(0, 2r) &= (N_{2n-1}(0,1) - N_{2n-1}(0,3)) + (N_{2n-1}(0,3) - N_{2n-1}(0,5)) + \ldots \\
&\quad + (N_{2n-1}(0, 2n-1) - N_{2n-1}(0, 2n+1)) \\
&= N_{2n-1}(0,1) - N_{2n-1}(0, 2n+1) = N_{2n-1}(0,1) = \frac{1}{2} \cdot N_{2n}(0,0).
\end{aligned}$$

(i) följer direkt av detta eftersom, i en symmetrisk slumpvandring, alla vägar av längd $2n$ har samma sannolikhet.
(ii) följer av att

$$P(S_1 \neq 0, \ldots, S_{2n} \neq 0) = P(S_1 > 0, \ldots, S_{2n} > 0) + P(S_1 < 0, \ldots, S_{2n} < 0)$$

och att de två fallen är lika sannolika.
För att visa (iii) observerar vi att

$$\begin{aligned}
P(S_1 > 0, \ldots, S_{2n} > 0) &= P(S_1 = 1, S_2 > 0, \ldots, S_{2n} > 0) \\
&= P(S_1 = 1) \cdot P(S_2 > 0, \ldots, S_{2n} > 0 | S_1 = 1) \\
&= \frac{1}{2} \cdot P(S_2 - S_1 \geq 0, S_3 - S_1 \geq 0, \ldots, S_{2n} - S_1 \geq 0) \\
&= \frac{1}{2} \cdot P(S_1 \geq 0, \ldots, S_{2n-1} \geq 0),
\end{aligned}$$

men eftersom $2n - 1$ är udda så gäller

$$S_{2n-1} \geq 0 \Rightarrow S_{2n-1} \geq 1 \Rightarrow S_{2n} \geq 0,$$

så att, med hjälp av (ii),

$$\frac{1}{2} \cdot f_{2n} = P(S_1 > 0, \ldots, S_{2n} > 0) = \frac{1}{2} \cdot P(S_1 \geq 0, \ldots, S_{2n} \geq 0).$$

$\square$

Anm 9. För en symmetrisk slumpvandring gäller alltså att

$$
\begin{aligned}
P(\text{aldrig återvända till } 0) &= \lim_{n\to\infty} P(S_k \neq 0,\ k = 1, \ldots, 2n) = \lim_{n\to\infty} P(S_{2n} = 0) \\
&= \lim_{n\to\infty} f_{2n} = 0,\ \text{(se Problem 11).}
\end{aligned}
$$

□

Den föregående satsen gav ett uttryck för $P(S_1 \neq 0, \ldots, S_n \neq 0)$ för en symmetrisk slumpvandring. Allmänt gäller

Sats 10. (i) För $k > 0$ gäller

$$P(S_1 > 0, \ldots, S_{n-1} > 0, S_n = k) = \frac{k}{n} \cdot P(S_n = k).$$

(ii) För $k \neq 0$ gäller

$$P(S_1 \neq 0, \ldots, S_{n-1} \neq 0, S_n = k) = \frac{|k|}{n} \cdot P(S_n = k).$$

(iii)

$$P(S_1 \neq 0, \ldots, S_n \neq 0) = \frac{E(|S_n|)}{n}.$$

Bevis: Både (i) och (ii) gäller trivialt om $P(S_n = k) = 0$. Antag därför att k är sådant att

$$
\begin{aligned}
P(S_n = k) = N_n(0, k) \cdot p^h \cdot q^v &> 0,\ \text{där} \\
h &= \frac{n+k}{2} \text{ och} \\
v &= \frac{n-k}{2} \text{ är heltal.}
\end{aligned}
$$

(i) följer då direkt av Ballot-satsen efter multiplikation med $p^h \cdot q^v$ och (ii) på grund av symmetrin, eftersom

$$
\begin{aligned}
N_n^{\neq 0}(0, k) &= N_n^{\neq 0}(0, -k) \text{ och} \\
N_n(0, k) &= N_n(0, -k).
\end{aligned}
$$

Summering av (ii) över alla $k \neq 0$ ger (iii). □

Sats 9 kan också användas för att ge ett alternativt uttryck för förstapassagesannolikheten $g_{2n} = P(T_0 = 2n)$, där $T_0 =$ ”tidpunkten för första återbesöket i 0” $= \min\{n \geq 1 : S_n = 0\}$.

Sats 11. För en symmetrisk slumpvandring gäller

$$g_{2n} = f_{2n-2} - f_{2n}.$$

Bevis: Sats 9 säger att

$$P(T_0 > 2n) = P(S_k \neq 0,\ k = 1, \ldots, 2n) = P(S_{2n} = 0) = f_{2n},$$

vilket ger att

$$g_{2n} = P(T_0 = 2n) = P(T_0 > 2n-2) - P(T_0 > 2n) = f_{2n-2} - f_{2n}.$$

□

Anm 10. Eftersom $f_0 = 1$ och $f_{2n} \to 0$ då $n \to \infty$ ger satsen att, för den symmetriska slumpvandringen,

$$\begin{aligned} P(\text{någonsin återkomma till } 0) &= \sum_{n=1}^{\infty} g_{2n} = \sum_{n=1}^{\infty} (f_{2n-2} - f_{2n}) \\ &= (f_0 - f_2) + (f_2 - f_4) + \cdots = f_0 = 1. \end{aligned}$$

□

Ytterligare ett samband mellan $\{f_n\}$ och $\{g_n\}$ ges av

Sats 12. För en godtycklig slumpvandring gäller

$$f_{2n} = \sum_{r=1}^{n} g_{2r} \cdot f_{2n-2r}.$$

Bevis: Betinga med avseende på första återkomsten till 0, T_0.

$$\begin{aligned} f_{2n} = P(S_{2n} = 0) &= \sum_{r=1}^{n} P(T_0 = 2r) \cdot P(S_{2n} = 0 | T_0 = 2r) \\ &= \sum_{r=1}^{n} g_{2r} \cdot P(S_{2n} - S_{2r} = 0 | T_0 = 2r) \\ &= \sum_{r=1}^{n} g_{2r} \cdot P(S_{2n-2r} = 0) = \sum_{r=1}^{n} g_{2r} \cdot f_{2n-2r}. \end{aligned}$$

□

Ett annat användbart trick för att analysera slumpvandringar är *tidsomvändning*. Eftersom $\{X_k\}$ är olf så har vektorn $(X_1, X_2, \ldots, X_n)$ samma fördelning som vektorn $(X_n, X_{n-1}, \ldots, X_1)$ och därav följer att $(S_1, S_2, \ldots, S_n)$ har samma fördelning som $(X_n, X_n + X_{n-1}, \ldots, X_n + \cdots + X_1) = (S_n - S_{n-1}, S_n - S_{n-2}, \ldots, S_n)$.

Låt $T_b =$ ”tidpunkten för första besöket i b” $= \min\{n \geq 1 : S_n = b\}$.

Sats 13. För $b > 0$ gäller

$$P(T_b = n) = \frac{b}{n} \cdot P(S_n = b), \text{ för } n = b, b+1, \ldots$$

Bevis: Genom att utnyttja tidsomvändning får vi

$$\begin{aligned} P(T_b = n) &= P(S_1 < b, S_2 < b, \ldots, S_{n-1} < b, S_n = b) \\ &= P(S_n > S_1, S_n > S_2, \ldots, S_n > S_{n-1}, S_n = b) \\ &= P(S_n - S_{n-1} > 0, \ldots, S_n - S_1 > 0, S_n = b) \\ &= P(S_1 > 0, \ldots, S_{n-1} > 0, S_n = b) = \frac{b}{n} \cdot P(S_n = b), \end{aligned}$$

enligt Sats 10 (i). □

Anm 11. Med hjälp av satsen kan man beräkna

$$P(T_b > n) = \sum_{k=n+1}^{\infty} P(T_b = k) = \sum_{k=n+1}^{\infty} \frac{b}{k} \cdot P(S_k = b)$$

och därigenom

$$E(T_b) = \sum_{n=0}^{\infty} P(T_b > n).$$

Vi vet sedan tidigare att $E(T_b) = b \cdot E(T_1)$, så det räcker att beräkna $E(T_1)$. □

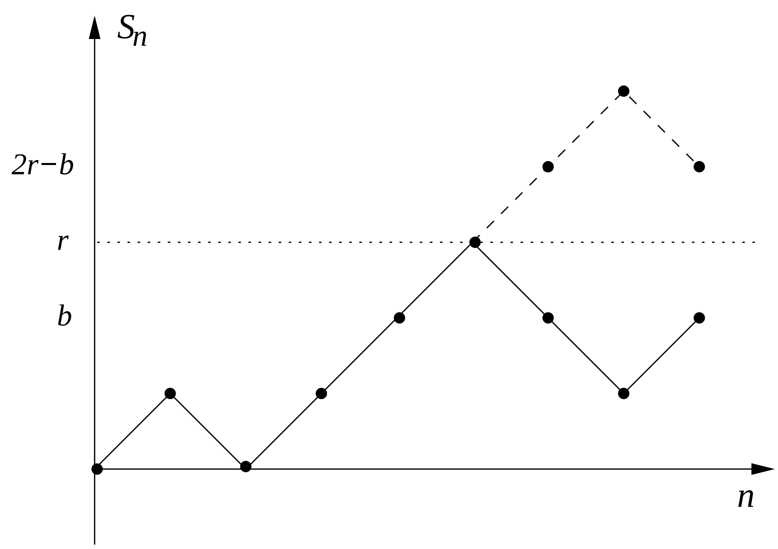
4.4 Maximum

Låt $M_n = \max(S_0, S_1, \dots, S_n)$. *Vilken fördelning har M_n?*

Lemma 1. För en symmetrisk slumpvandring gäller

$$P(M_n \geq r, S_n = b) = \begin{cases} P(S_n = b) & \text{om } b \geq r, \\ P(S_n = 2r - b) & \text{om } b < r. \end{cases}$$

Bevis: Den första delen följer direkt av att $M_n \geq S_n$. Antag därför att $b < r$. Genom att spegla slutet av vägen, från och med sista besöket i r, i linjen $y = r$ ser vi att # vägar av längd n med $M_n = r$ och $S_n = b$ är lika med # vägar av längd n med $S_n = 2r - b$; se Figur 5. Lemmat följer av att alla vägar av längd n har samma sannolikhet för en symmetrisk slumpvandring. □



Figur 5: Spegling för maximum

Sats 14. För en symmetrisk slumpvandring gäller, för $r \geq 1$,

$$P(M_n \geq r) = P(S_n = r) + 2P(S_n > r),$$
$$P(M_n = r) = P(S_n = r) + P(S_n = r + 1) = \max(P(S_n = r), P(S_n = r + 1)).$$

Bevis: Enligt lemmat gäller

$$\begin{aligned}
P(M_n \geq r) &= \sum_b P(M_n \geq r, S_n = b) = \sum_{b \geq r} P(S_n = b) + \sum_{b<r} P(S_n = 2r - b) \\
&= P(S_n \geq r) + \sum_{k>r} P(S_n = k) = P(S_n \geq r) + P(S_n > r) \\
&= P(S_n = r) + 2P(S_n > r).
\end{aligned}$$

Vidare gäller

$$\begin{aligned}
P(M_n = r) &= P(M_n \geq r) - P(M_n \geq r + 1) \\
&= P(S_n = r) + 2P(S_n > r) - (P(S_n = r + 1) + 2P(S_n > r + 1)) \\
&= P(S_n = r) + 2P(S_n = r + 1) - P(S_n = r + 1) \\
&= P(S_n = r) + P(S_n = r + 1) = \max(P(S_n = r), P(S_n = r + 1)),
\end{aligned}$$

eftersom endast en av $P(S_n = r)$ och $P(S_n = r + 1)$ kan vara skild från 0. □

4.5 Arcsinus-lagen

Vi ska visa att två stokastiska variabler med anknytning till symmetrisk slumpvandring har samma fördelning, den s.k. *arcsinus-fördelningen*.

Betrakta en symmetrisk slumpvandring $\{S_n\}$ med $S_0 = 0$.
Låt som tidigare $f_n = P(S_n = 0)$.
Definiera $Y_{2n} = \max(k \leq 2n : S_k = 0)$, som är väldefinierad eftersom $S_0 = 0$, och $\alpha_{2n}(2k) = P(Y_{2n} = 2k)$. Då gäller

Sats 15. För $0 \leq k \leq n$,

$$\alpha_{2n}(2k) = P(Y_{2n} = 2k) = f_{2k} \cdot f_{2n-2k}.$$

Bevis:

$$\begin{aligned}
P(Y_{2n} = 2k) &= P(S_{2k} = 0, S_{2k+1} \neq 0, \dots, S_{2n} \neq 0) \\
&= P(S_{2k} = 0) \cdot P(S_{2k+1} \neq 0, \dots, S_{2n} \neq 0 | S_{2k} = 0) \\
&= P(S_{2k} = 0) \cdot P(S_1 \neq 0, \dots, S_{2n-2k} \neq 0) \\
&= P(S_{2k} = 0) \cdot P(S_{2n-2k} = 0) \quad \text{(Sats 9 (i))} \\
&= f_{2k} \cdot f_{2n-2k}.
\end{aligned}$$

□

Anm 12. Namnet arcsinus-fördelningen kommer av att

$$P(Y_{2n} \leq 2xn) \sim \frac{2}{\pi} \cdot \arcsin \sqrt{x}, \quad \text{då } n \to \infty.$$

För att se detta kan man utnyttja Stirlings formel

$$n! \sim n^n \cdot e^{-n} \sqrt{2\pi n},$$

för att visa att, för stora k,

$$f_{2k} \sim \frac{1}{\sqrt{\pi k}}, \quad \text{(Se Problem 11c.)}$$

och att därför

$$\alpha_{2n}(2k) \sim \frac{1}{n} \cdot f(\frac{k}{n}),$$

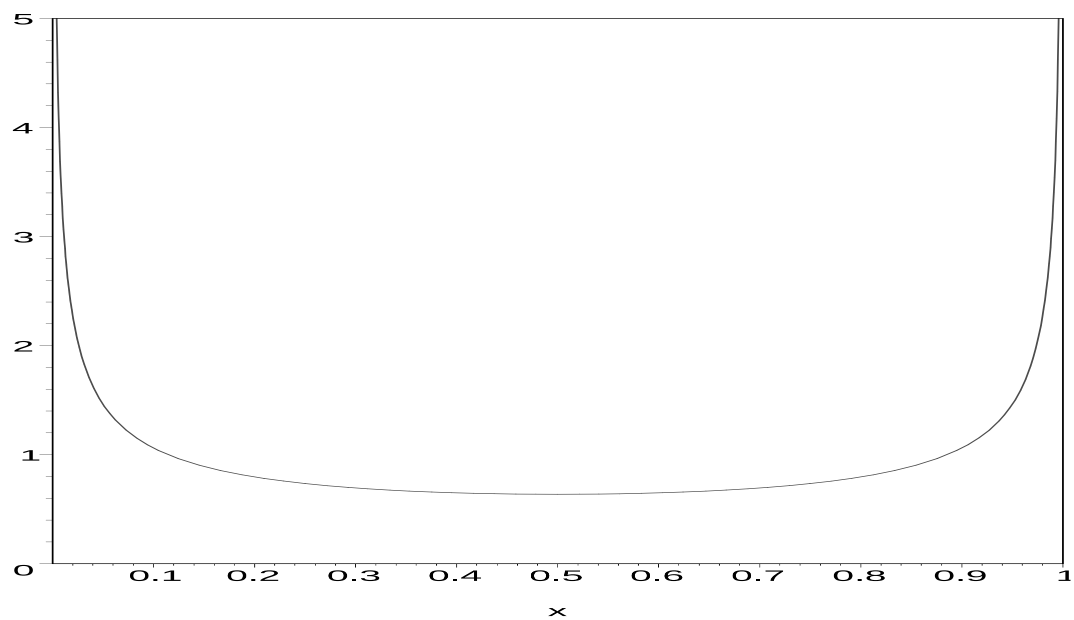
där

$$f(x) = \frac{1}{\pi\sqrt{x(1-x)}}, \quad 0 < x < 1.$$

Detta ger att

$$P(Y_{2n} \le 2xn) = \sum_{k \le xn} \alpha_{2n}(2k) \sim \frac{1}{n} \cdot \sum_{k/n \le x} f(\frac{k}{n}) \sim \int_0^x f(t)\,dt = \frac{2}{\pi} \cdot \arcsin\sqrt{x}.$$

Täthetsfunktionen $f(x) = \frac{1}{\pi\sqrt{x(1-x)}}$ finns plottad i Figur 6. □



Figur 6: Plot av $f(x) = \frac{1}{\pi\sqrt{x(1-x)}}$.

Anm 13. Den exakta fördelningen, $\alpha_{2n}(2k)$, brukar kallas den *diskreta arcsinus-fördelningen*. Den är, liksom den kontinuerliga, symmetrisk kring mittpunkten, $k = n/2$, där den också har sitt minimum. Maximum antas för $k = 0$ och $k = n$. □

Anm 14. En måhända överraskande egenskap hos den symmetriska slumpvandringen, som följer av arcsinus-lagen, är att om vi har singlat slant $2n$ gånger och är intresserade av när vi senast hade lika många krona och klave så är det mest troligt att det inträffade endera alldeles nyligen, eller alldeles i början av försöket.
Om t.ex. $n = 1000$, dvs. vi har singlat 2000 gånger, så är

$$P(Y_{2000} \le 200) \sim \frac{2}{\pi} \cdot \arcsin\sqrt{0.1} = 0.205,$$
$$P(Y_{2000} \le 20) \sim \frac{2}{\pi} \cdot \arcsin\sqrt{0.01} = 0.064.$$

□

Det finns också en arcsinuslag för *uppehållstider*. Vi säger att slumpvandringen är positiv i tidsintervallet $(k, k+1)$ om $S_k > 0$ eller $S_{k+1} > 0$. (Naturligt; se Figur 3.)
Låt Z_n = "# positiva tidsintervall mellan 0 och $2n$". Då antar Z_{2n} alltid ett jämnt värde och vidare gäller

Sats 16. För $0 \leq k \leq n$,

$$P(Z_{2n} = 2k) = \alpha_{2n}(2k) = f_{2k} \cdot f_{2n-2k}.$$

Bevis: Låt $b_{2n}(2k) = P(Z_{2n} = 2k)$.
Vi vill visa att $b_{2n}(2k) = \alpha_{2n}(2k)$ för alla n och $0 \leq k \leq n$.
Enligt Sats 9 (iii) gäller

$$b_{2n}(2n) = P(S_k \geq 0, k = 1, \ldots, 2n) = f_{2n} = f_{2n} \cdot f_0,$$

så att satsen gäller för $k = n$. Av symmetriskäl gäller också

$$b_{2n}(0) = P(S_k \leq 0, k = 1, \ldots, 2n) = f_{2n},$$

så att den också gäller för $k = 0$.
Det återstår alltså att visa att

$$b_{2n}(2k) = \alpha_{2n}(2k) (= f_{2k} \cdot f_{2n-2k}) \quad \text{för alla } n \text{ och } 0 < k < n. \tag{6}$$

Om $Z_{2n} = 2k$, där $1 \leq k \leq n-1$, måste $S_{2r} = 0$ för något r, $1 \leq r \leq n-1$, dvs. $T_0 < 2n$. Tiden fram till T_0 tillbringas med lika stor sannolikhet på den positiva sidan som på den negativa. Betingning m.a.p. T_0 ger då, för $1 \leq k \leq n-1$,

$$\begin{aligned}
b_{2n}(2k) &= \sum_{r=1}^{n-1} P(T_0 = 2r) \cdot P(Z_{2n} = 2k | T_0 = 2r) \\
&= \sum_{r=1}^{n-1} g_{2r} \cdot \frac{1}{2} \cdot P(Z_{2n-2r} = 2k) + \sum_{r=1}^{n-1} g_{2r} \cdot \frac{1}{2} \cdot P(Z_{2n-2r} = 2k - 2r) \\
&= \frac{1}{2} \cdot \sum_{r=1}^{n-1} g_{2r} \cdot b_{2n-2r}(2k) + \frac{1}{2} \cdot \sum_{r=1}^{n-1} g_{2r} \cdot b_{2n-2r}(2k-2r).
\end{aligned}$$

Observera att $b_{2n-2r}(2k) = 0$ om $k > n - r$ och att $b_{2n-2r}(2k-2r) = 0$ om $k < r$, så att

$$b_{2n}(2k) = \frac{1}{2} \cdot \sum_{r=1}^{n-k} g_{2r} \cdot b_{2n-2r}(2k) + \frac{1}{2} \cdot \sum_{r=1}^{k} g_{2r} \cdot b_{2n-2r}(2k-2r). \tag{7}$$

Vi ska utnyttja (7) för att visa (6) med hjälp av induktion. För $n = 1$ gäller (6) trivialt. Antag att (6) gäller för $n < m$. Då är

$$\begin{aligned}
b_{2m}(2k) &= \frac{1}{2} \cdot \sum_{r=1}^{m-k} g_{2r} \cdot b_{2m-2r}(2k) + \frac{1}{2} \cdot \sum_{r=1}^{k} g_{2r} \cdot b_{2m-2r}(2k-2r) \\
&= \frac{1}{2} \cdot \sum_{r=1}^{m-k} g_{2r} \cdot f_{2k} \cdot f_{2m-2r-2k} + \frac{1}{2} \cdot \sum_{r=1}^{k} g_{2r} \cdot f_{2k-2r} \cdot f_{2m-2k} \\
&= \frac{1}{2} \cdot f_{2k} \cdot \sum_{r=1}^{m-k} g_{2r} \cdot f_{2m-2r-2k} + \frac{1}{2} \cdot f_{2m-2k} \cdot \sum_{r=1}^{k} g_{2r} \cdot f_{2k-2r}.
\end{aligned}$$

Enligt Sats 12 gäller

$$\sum_{r=1}^{k} g_{2r} \cdot f_{2k-2r} = f_{2k},$$

$$\sum_{r=1}^{m-k} g_{2r} \cdot f_{2(m-k)-2r} = f_{2m-2k},$$

så att

$$b_{2m}(2k) = \frac{1}{2} \cdot f_{2k} \cdot f_{2m-2k} + \frac{1}{2} \cdot f_{2m-2k} \cdot f_{2k} = f_{2k} \cdot f_{2m-2k} = \alpha_{2m}(2k).$$

□

Anm 15. Betrakta två spelare, A och B, som spelar ett rättvist spel, där båda kan vinna en krona av den andre med sannolikhet 1/2. Många tolkar nog Stora talens lag intuitivt som att i det långa loppet kommer båda spelarna att vara i ledningen ungefär halva tiden. *Detta är inte korrekt!* Arcsinus-lagen för uppehållstider ger att, efter många spelomgångar, det gäller att

$$\begin{aligned}
P(\text{A leder minst andelen } x \text{ av tiden}) &= \frac{2}{\pi} \cdot \arcsin \sqrt{1-x}, \\
P(\text{A leder minst 80\% av tiden}) &= \frac{2}{\pi} \cdot \arcsin \sqrt{0.2} = 0.295, \\
P(\text{någon leder minst 80\% av tiden}) &= 2 \cdot \frac{2}{\pi} \cdot \arcsin \sqrt{0.2} = 0.59, \\
P(\text{någon leder minst 90\% av tiden}) &= 2 \cdot \frac{2}{\pi} \cdot \arcsin \sqrt{0.1} = 0.41, \\
P(\text{någon leder minst 95\% av tiden}) &= 2 \cdot \frac{2}{\pi} \cdot \arcsin \sqrt{0.05} = 0.29, \\
P(\text{någon leder minst 99\% av tiden}) &= 2 \cdot \frac{2}{\pi} \cdot \arcsin \sqrt{0.01} = 0.13,
\end{aligned}$$

□

5 Blandade problem

Problem 13. Betrakta en slumpvandring med $p < 1/2$, som vid tidpunkt k befinner sig i position $a < m$. Beräkna den betingade sannolikheten att den vid tidpunkt $k+1$ befinner sig i position $a+1$ (eller i $a-1$) givet att den kommer att besöka tillståndet m i framtiden. □

Problem 14. Låt T_{0a} vara definierad som i Avsnitt 2.2, och $p > 1/2$.
a) Visa att

$$\text{Var}(T_{01}) = \frac{4pq}{(p-q)^3}.$$

Ledning: Studera $E(T_{01}^2)$ och betinga.
b) Vad blir $\text{Var}(T_{0a})$ för $a > 0$? □

Problem 15. Betrakta en spelare, med vinstsannolikhet p i ett enskilt spel, som startar med a kronor mot en oändligt rik motståndare. Vad är sannolikheten att det tar $a + 2k$ spelomgångar innan han ruineras? □

Problem 16. Visa att det finns exakt lika många vägar (x, y), som slutar i $(2n + 2, 0)$ och för vilka $y > 0$ för $0 < x < 2n + 2$, som det finns vägar som slutar i $(2n, 0)$ och för vilka $y \geq 0$ för $0 \leq x \leq 2n$. Visa också att detta, för en symmetrisk slumpvandring, medför att

$$P(S_1 \geq 0, \dots, S_{2n-1} \geq 0, S_{2n} = 0) = 2 \cdot g_{2n+2}. \qquad \square$$

Problem 17. Visa att sannolikheten att en symmetrisk slumpvandring *före* tidpunkten $2n$ återkommer exakt r gånger till 0 är densamma som sannolikheten att $S_{2n} = 0$ och att den dessförinnan återvänt minst r gånger till 0. $\square$

Problem 18. En partikel flyttar sig endera *två* steg åt höger, med sannolikhet p, eller ett steg åt vänster, med sannolikhet $q = 1 - p$. Olika steg är oberoende av varandra.
a) Om den startar i $z > 0$, vad är sannolikheten, a_z, att den någonsin kommer till 0?
b) Visa att a_1 är sannolikheten att, i en följd Bernoulliförsök som lyckas med sannolikhet p, antalet misslyckade försök någonsin överstiger dubbla antalet lyckade försök.
c) Visa att, då $p = q$,

$$a_1 = \frac{\sqrt{5} - 1}{2}. \qquad \square$$

Problem 19. Visa att, för en symmetrisk slumpvandring som startar i 0, sannolikheten att det första besöket i S_{2n} inträffar i steg $2k$ är $P(S_{2k} = 0) \cdot P(S_{2n-2k} = 0)$. $\square$

Problem 20. (Banachs tändsticksproblem)
En person har i var och en av sina två fickor en tändsticksask med n tändstickor i varje. När han behöver en tändsticka väljer han slumpmässigt en av askarna, ända tills han påträffar en tom ask. Låt, när detta inträffar, $R =$ # stickor i den andra asken.
a) Beräkna $E(R)$.
b) Om a) är för svårt; uppskatta $E(R)$ för $n = 50$ med hjälp av simulering. $\square$

Problem 21. Låt, i en tvådimensionell symmetrisk slumpvandring, startande i origo, $D_n^2 = x_n^2 + y_n^2$, där (x_n, y_n) är partikelns position efter n steg.
Visa att $E(D_n^2) = n$.
(*Ledning:* Studera $E(D_n^2 - D_{n-1}^2)$.) $\square$

Problem 22. Visa att en symmetrisk slumpvandring i d dimensioner med sannolikhet 1 kommer att återvända till en redan besökt position.
(Med sannolikhet 1 sker detta dessutom oändligt många gånger.)
(*Ledning:* I varje steg är sannolikheten att nå en ny punkt högst $(2d - 1)/(2d)$.) $\square$

6 Litteratur

En guldgruva om man vill läsa mer om slumpvandringar är
Feller, W., *An Introduction to Probability Theory and Its Applications*, Vol. 1, Third edition, Wiley 1968.
Många av exemplen och resultaten är hämtade därifrån, speciellt från Kapitel III, men även från Kapitel XIV.
Några exempel är också hämtade ur
Grimmett, G.R. & Stirzaker, D.R., *Probability and Random Processes*, Second edition, Oxford Science Publications, 1992.
En trevlig beskrivning av många klassiska sannolikhetsproblem, bl.a. slumpvandring, ges i
Blom, G., Holst, L. & Sandell, D., *Problems and Snapshots from the World of Probability*. Springer 1994.