# Utflykter i en slumpvandring

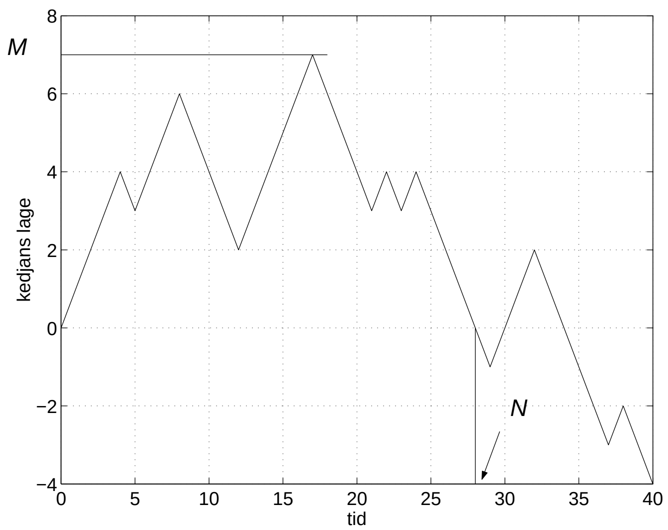
Jan Enger
Matematisk statistik
KTH

Låt $\{X_n; n \geq 0\}$, $X_0 = 0$, vara en symmetrisk slumpvandring på heltalen, dvs $X_n$ är läget av en punkt som startar i 0 och som vid varje tidpunkt hoppar till vänster med sannolikhet 1/2 och till höger med sannolikheten 1/2. Med en "utflykt" menar vi kedjan från start till den för första gången återvänder till starttillståndet 0. Vi skall undersöka denna slumpvandrings beteende och inför därför

$N$=tiden för en utflykt, dvs $X_1 \neq 0, X_2 \neq 0, \ldots, X_{N-1} \neq 0, X_N = 0$

$M$=maximala absolutvärde av kedjan under en utflykt, $M = \max\limits_{1 \leq i \leq N} (\mid X_i \mid)$.

$M$ anger sålunda hur långt ut kedjan nått under en utflykt.




Under en utflykt håller sig kedjan hela tiden på positiva axeln eller hela tiden på den negativa axeln, och det räcker av symmetriskäl att betrakta fallet att utflykten sker till höger. Vi kan alltså betrakta en kedja där hopp från 0 till 1 sker med sannolikhet 1 och övriga hopp med sannolikhet 1/2.

Betrakta först $P(M \geq n)$. Vid tid 1 befinner sig kedjan i tillstånd 1 och vi skall alltså beräkna sannolikheten att kedjan når tillstånd $n$ innan den når tillstånd 0. Gör om tillstånden 0 och $n$ till absorberande tillstånd. Genom eftertanke inses att den sökta sannolikheten då är sannolikheten att absorberas i tillstånd $n$ givet start i tillstånd 1. Från exempel 3.7 i markovkompendiet ser man att denna sannolikhet är $1/n$, och därmed är

$$P(M = n) = P(M \geq n) - P(M \geq n+1) = \frac{1}{n} - \frac{1}{n+1} = \frac{1}{n(n+1)}, n = 1, 2, \ldots$$

Av detta följer att väntevärdet är oändligt ty $\sum_{n=1}^{\infty} n \dfrac{1}{n(n+1)} = \sum_{n=1}^{\infty} \dfrac{1}{n+1} = \infty$. Härav fås också att väntevärdet för $N$ är oändligt, eftersom $N \geq M$.

Fördelningen för $N$ är svårare att komma åt. En metod är följande.

Vi inför först de stokastiska tiderna

$U_i$=tid tills kedjan hamnar i tillstånd $i-1$ vid start i tillstånd $i$, $i = 1, 2$,

dvs tid tills kedjan för första gången hamnar ett steg till vänster om ett starttillstånd. Man inser då att $N = 1 + U_1$. På grund av markoviteten är $U_1$ och $U_2$ oberoende.

Händelsen $\{N = 2\}$ inträffar om, och endast om, kedjan efter första hoppet till tillstånd 1 direkt hoppar tillbaka till tillstånd 0 och därför är $P(N = 2) = 1/2$.

Med sannolikhet 1/2 kommer kedjan att under de två första tidsstegen hoppa till tillstånd 2. Då återstår tid $U_2 + U_1$ tills kedjan återigen besöker tillstånd 0. Först skall kedjan gå från tillstånd 2 till tillstånd 1 och därefter återstår tiden att från tillstånd 1 gå till tillstånd 0. Dessa två tider är $U_2$ och $U_1$ och total tid är därför $2 + U_1 + U_2 = N_1 + N_2$ där $N_i = U_i + 1$. Vi har erhållit

$$N = \begin{cases} 2 \text{ med sannolikhet } 1/2 \\ N_1 + N_2 \text{ med sannolikhet } 1/2. \end{cases} \tag{1}$$

där $N_1$, och $N_2$ är oberoende med samma fördelning som $N$.

Sätt nu $p_k = P(N = 2k)$ (observera att $N$ måste vara jämnt). Ekvationen (1) ger direkt $p_1 = P(N = 2) = 0.5$. För övriga sannolikheter får man

$$\begin{aligned} p_n = P(N = 2n) = P(\text{kedjan går först till tillstånd 2})P(N_1 + N_2 = 2n) = \\ \frac{1}{2}\sum_{k=0}^{n} P(N_1 = 2k, N_2 = 2n - 2k) = \frac{1}{2}\sum_{k=0}^{n} P(N_1 = 2k)P(N_2 = 2n - 2k) = \\ \frac{1}{2}\sum_{k=0}^{n} p_k p_{n-k} = \frac{1}{2}\sum_{k=1}^{n-1} p_k p_{n-k} \end{aligned} \tag{2}$$

eftersom $p_0 = 0$. Här har vi utnyttjat att $N_1$ och $N_2$ är oberoende och har samma fördelning som $N$.

Utnyttjar vi (2) erhålls successivt

$$\begin{aligned} P(N = 2) &= p_1 = \frac{1}{2} \\ P(N = 4) &= p_2 = \frac{1}{2}p_1p_1 = \frac{1}{8} \\ P(N = 6) &= p_3 = \frac{1}{2}(p_1p_2 + p_2p_1) = \frac{1}{16} \\ P(N = 8) &= p_4 = \frac{1}{2}(p_1p_3 + p_2p_2 + p_3p_1) = \frac{5}{128} \\ &\vdots \end{aligned}$$

Man kan visa att den allmänna lösningen är

$$p_n = \frac{1 \cdot 3 \cdot 5 \cdots (2n-3)}{2 \cdot 4 \cdots 2n} = \frac{(2n-3)!!}{(2n)!!}$$

Uttrycken i täljare och nämnare kallas semifakultet och är alltså produkten av vartannat tal, udda eller jämna. Låt nu

$$a_n = \frac{(2n-1)!!}{(2n)!!}$$

Då blir

$$a_{n-1} - a_n = \frac{(2(n-1)-1)!!}{(2(n-1))!!} - \frac{(2n-1)!!}{(2n)!!} = \frac{(2n-3)!!}{(2n-2)!!}\left(1 - \frac{2n-1}{2n}\right) = \frac{(2n-3)!!}{(2n)!!} = p_n$$

Av detta erhålls att

$$P(N > 2n) = \sum_{k=n+1}^{\infty} P(N = 2k) = \sum_{k=n+1}^{\infty} p_k = \sum_{k=n+1}^{\infty} (a_{k-1} - a_k) =$$
$$a_n - a_{n+1} + a_{n+1} - a_{n+2} + a_{n+2} - a_{n+3} + \cdots = a_n = \frac{(2n-1)!!}{(2n)!!}$$

eftersom alla termer utom den första i serieutvecklingen tar ut varandra.
Utnyttjar man nu att

$$(2n)!! = 2 \cdot 4 \cdots 2n = (2 \cdot 1) \cdot (2 \cdot 2) \cdots (2 \cdot n) = 2^n n!$$
$$(2n-1)!! = \frac{(2n-1)!}{(2n-2)!!} = \text{(se ovan)} = \frac{(2n-1)!}{2^{n-1}(n-1)!}$$

och Stirlings formel, ($n! \approx e^{-n} n^n \sqrt{2\pi n}$) erhålls

$$P(N = 2n) = \frac{(2n-3)!!}{(2n)!!} \approx \frac{1}{2n\sqrt{\pi n}}$$
$$P(N > 2n) = \frac{(2n-1)!!}{(2n)!!} \approx \frac{1}{\sqrt{\pi n}}$$

Vi har således t.ex. att sannolikheten att tiden till återkomst överstiger 100000 är approximativt $\frac{1}{\sqrt{\pi 50000}} = 0.002523$, dvs i stort sett varar en utflykt av 400 längre tid än 100000.