Elementy Modelowania Matematycznego

Wykład 7

Modele Markova

Spis treści

- Wstęp
- Łańcuch i procesy Markova
- Przykłady procesów Markova

Wstęp

- Andrey Andreyevich Markov (14 czerwca 1856   20 czerwca 1922) wybitny rosyjski matematyk.
- Znany jest przede wszystkim ze swych prac na temat procesów stochastycznych, zwanych później łańcuchami Markova.
- On i jego młodszy brat Vladimir Andreevich Markov (1871-1897) udowodnili tzw. Nierówność Markova.

Wstęp

- Procesy Markowa w reprezentacji dyskretnej lub ciągłej, jako wyodrębniona grupa procesów przypadkowych (losowych), są dziś najlepiej zbadaną dziedziną procesów losowych i znalazły zastosowanie w modelowaniu wielu zjawisk z życia codziennego

Wstęp

- Prognozowanie cen akcji giełdowych
- Niemal wszystkie akcje zmieniają swoje ceny codziennie, a duża ich część w sposób niemal ciągły.
- Wykresy cen przedstawiają, w zależności od nastawienia obserwatora, efekt równoważenia się popytu i podaży, dyskontowanie przyszłych zdarzeń, reakcje na wydarzenia historyczne, efekt manipulacji akcjami, wpływ kosmosu bądź też całkowicie przypadkowe ruchy Browna.

Wstęp

- Te bądź jeszcze inne przyczyny zmian cen usiłuje się wykorzystać w analizie historycznych przebiegów i próbie prognozowania przyszłego zachowania cen.
- Zachowanie poszczególnych akcji zapisane w postaci kolejnych cen i przekształcone do postaci graficznej to dla każdego inwestora wykres ceny.
- Bardzo podobne przebiegi i szeregi liczb są znane i używane w wielu dziedzinach nauki

Wstęp

- Jako że szeregi te opisują za pomocą kolejnych liczb zachowanie pewnego zjawiska w czasie, bardzo często określa się je mianem szeregów czasowych.
- Kolejne zdarzenia występujące w szeregach czasowych tworzą pewien proces.
- Tak więc to, co dla inwestora jest wykresem cen, dla specjalisty zajmującego się na przykład teorią informacji, jest graniczną interpretacją szeregu czasowego opisującego proces zmian cen.

Wstęp

- Proces stochastyczny jest to takie zjawisko (reprezentowane liczbowo przez szereg czasowy), w którym przyszła wartość opisująca stan zjawiska nie jest pewna (przyszłe liczby opisujące je mogą przyjmować różne wartości, przy czym żadna z nich nie pojawi się z prawdopodobieństwem równym 1).

Wstęp

- Klasycznymi przypadkami procesów stochastycznych są przyszłe wartości zmiennych opisujących pogodę
  - temperatura,
  - ciśnienia ,
  - kierunek bądź siła wiatru).
- Dobrym przykładem może być wypełnianie się niżu.

Wstęp

- Można nawet w pewien sposób oszacować drogę, którą się przesunie i czas potrzebny na podniesienie się ciśnienia wewnątrz niżu do wartości średniej.
- Nie da się tego jednak zrobić w sposób dokładny.
- Czyli mimo pewnych ściśle określonych ram zachowania, dokładne zachowanie nie jest znane.
- Podobnie jest ze zmianami cen na giełdzie.
- Jakkolwiek każdy silny spadek kiedyś musi się skończyć, nigdy nie mamy pewności kiedy to nastąpi.

Łańcuch Markowa

- Ciąg Markowa to taki proces stochastyczny, w którym określone są związki probabilistyczne przyszłych zdarzeń w zależności od wcześniej występujących.

Ciąg Markowa

- ciąg Markova pierwszego rzędu - jutrzejsze zachowanie zależy (w sensie statystycznym) tylko i wyłącznie od dzisiejszej zmiany
- ciąg Markova drugiego rzędu - prawdopodobieństwo jutrzejszego zachowania zależy od dzisiejszej i wczorajszej zmiany
- ciąg Markova zerowego rzędu - jutrzejsze zachowanie jest całkowicie niezależne od wcześniejszych notowań (bez względu jakie było zachowanie historyczne przyszłe zmiany będą określone takimi samymi związkami prawdopodobieństw);
- właśnie takie założenie jest wykorzystywane w analizie portfelowej czyli fakt, że przyszłość nie zależy od przeszłości może być w jakiś sposób wykorzystany w procesie inwestycyjnym.

Proces Markowa

- Proces Markowa bazuje wyłącznie na rozkładzie prawdopodobieństw warunkowych.
- Może się więc zdarzyć, że mamy do czynienia z deterministycznym procesem chaotycznym, w którym jutrzejsze zachowanie określone jest ścisłym wzorem, a mimo to proces będzie sprawiał wrażenie, że jest zerowego rzędu (to znaczy zupełnie nie zależy od przeszłości).

Proces Markowa

- Wynika to z faktu, że bardzo podobne, niemal identyczne zachowanie historyczne może skutkować zupełnie różnym zachowaniem w przyszłości.
- Tak więc mimo tego, że proces chaotyczny oznacza się istnieniem tak zwanej długoterminowej pamięci zachowania wykrycie tej zależności może być trudne bądź niemożliwe.

Proces Markowa

- Najważniejszym problemem w prognozowania cen jest brak stacjonarności procesu.
- Niestacjonarność to zjawisko, które jest źródłem większości niepowodzeń inwestorów giełdowych, próbujących wyznaczyć przyszłe ceny akcji na giełdzie.

Proces Markowa

- Proces stacjonarny to taki proces, w którym związki probabilistyczne są stałe i nie zależą od zmiennej niezależnej, czyli prawdopodobieństwo wystąpienia pewnej sytuacji nie zmienia się w miarę upływu czasu.

Proces Markowa

- Gdyby przyjąć, że zachowanie cen akcji jest procesem niestacjonarnym o nieznanej zmianie sposobu zachowania oznaczałoby to, że do prognozowania przyszłych cen potrzebna byłaby wiedza o przyszłym charakterze tego procesu, natomiast zupełnie nieprzydatna byłaby wiedza o wcześniejszym zachowaniu.

Proces Markowa

- W skrócie oznacza to, że wyłącznie osoby manipulujące rynkiem (przy założeniu, że jest to możliwe na większą skalę) mogłyby posiadać wiedzę jak zarobić na inwestycjach giełdowych.

Proces Markowa

- Należy rozróżnić niestacjonarność procesu od efektywności rynku.
- Rynek efektywny, jest skutkiem tego, że zmiany cen są procesem Markowa zerowego rzędu.
- Dodatkowo, charakteryzuje go tak zwana słaba stacjonarność, która cechuje się stałością średniej i wariancji.
- Czyli ostatecznie na rynku efektywnym ceny nie zależą od wcześniejszych.
- Natomiast w przypadku braku stacjonarności ceny zależą od poprzednich, lecz nie ma pewności, że wiemy w jaki sposób.

Proces Markowa

- W praktyce sprawa nie jest taka beznadziejna.
- Zmiany cen nie są procesem stacjonarnym, jednak zmienność zależności jest bardzo powolna.
- To znaczy system, który był dobry wczoraj będzie dobry jeszcze dzisiaj, a jutro będzie tylko trochę gorszy.
- Kiedyś oczywiście może utracić swoje właściwości. Ponadto można podejrzewać, że zmiany cen składają się z kilku (zapewne trzech) procesów o różnych charakterach.
- Bardzo prawdopodobne, że przynajmniej jeden z nich jest stacjonarny, czyli jego parametry ustalone w przeszłości będą w przyszłości takie same.

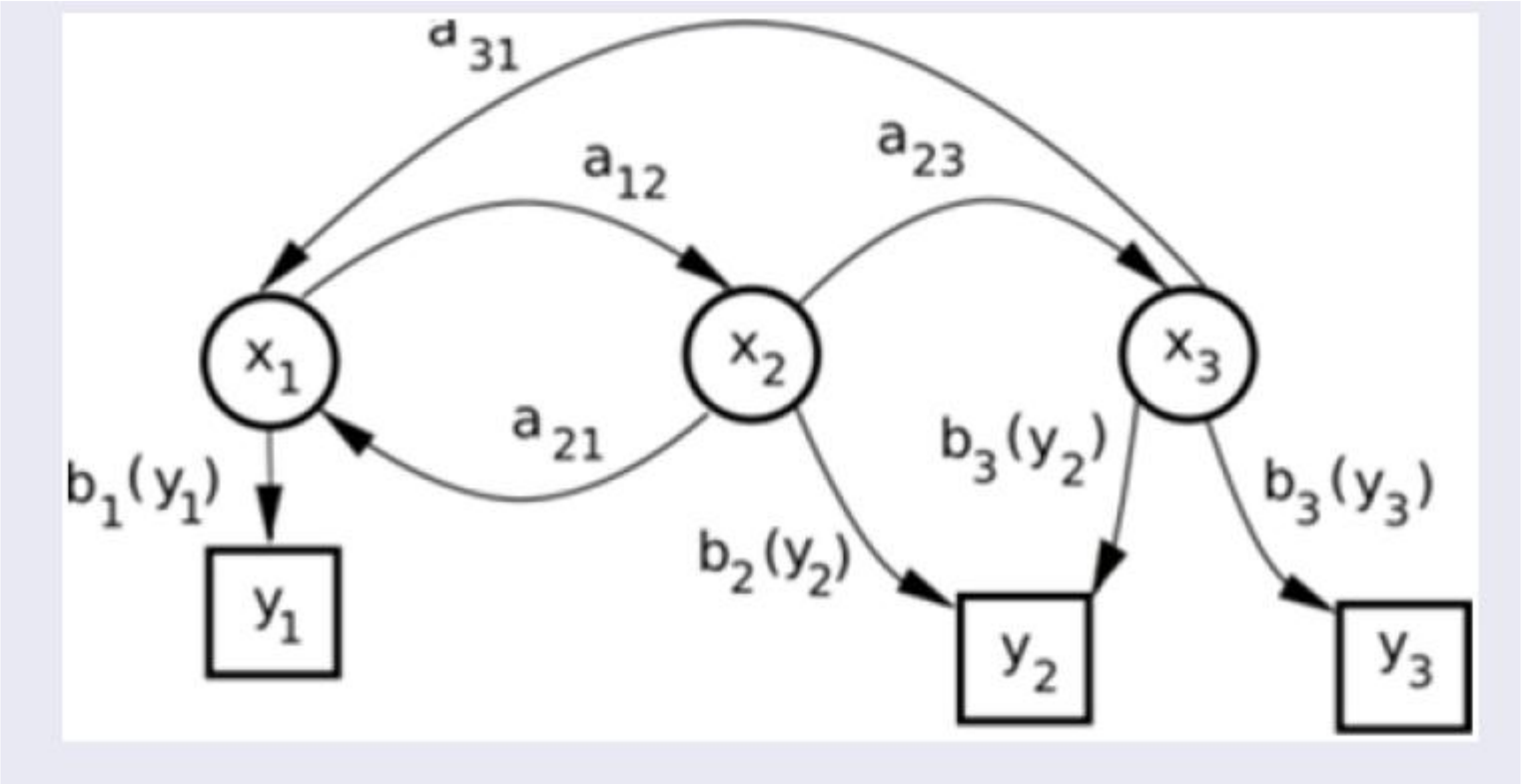
Proces Markowa

- Niech układ Ω może przyjmować stany $\omega_1, \omega_2\dots$ - zbiór skończony lub przeliczalny i niech w pewnej jednostce czasu może przejść z jednego stanu do innego z pewnym prawdopodobieństwem, to

$$P(\omega_j^{(n)} | \omega_i^{(n-1)})$$

- prawdopodobieństwo warunkowe, że układ znajdujący się w chwili n-1 w stanie ω_i przejdzie do stanu ω_j w chwili n.

Proces Markowa



Łańcuch jednorodny

- Jeśli prawdopodobieństwo nie zależy od czasu, tzn.:

$$P(\omega_j^{(n)}|\omega_i^{(n-1)}) = P(\omega_j^{(n+1)}|\omega_i^{(n)})$$

- to łańcuch Markova jest jednorodny, a macierz złożona z elementów

$$p_{ij} = P(\omega_j^{(n)}|\omega_i^{(n-1)})$$

- to macierz przejścia.
- Mamy $p_{ij} \geq 0, \sum_j p_{ij} = 1.$

Łańcuch jednorodny

Prawdopodobieństwo, że w n przejściach układ przejdzie ze stanu ω_i do stanu ω_j wynosi

$$p_{ij}(n) = \sum_k p_{ik}(m)\, p_{kj}(n-m)$$

m – liczba całkowita, $1 \leq m < n$.

Przykład łańcucha Markova: proces urodzin i śmierci — zmiana liczebności populacji na skutek narodzin i śmierci.

Łańcuch Pochłaniający

- Łańcuch Markova nazywamy pochłaniającym, jeśli istnieje taki stan i , z którego nie można wyjść, czyli:

$$p_{ii} = 1 \quad \wedge \quad \forall_{i \neq j} \, p_{ij} = 0$$

- Stan taki nazywamy stanem pochłaniającym (ang. absorbing
- state).
- Stan nie będący stanem pochłaniającym nazywamy stanem przejściowym (ang. transient state).

Postać kanoniczna łańcucha pochłaniającego

$$P = \begin{bmatrix} Q & R \\ 0 & I \end{bmatrix}$$

Q – macierz tranzytywna
0 – macierz zerowa
I – macierz identycznościowa
R – macierz przejścia do stanów pochłaniających

Łańcuch Markowa

- Łańcuch Markowa nazywamy ergodycznym, jeśli z dowolnego stanu można przejść do dowolnego innego (niekoniecznie w jednym kroku).

Procesy ergodyczne

- Centralnym zagadnieniem teorii procesów stochastycznych jest znalezienie rozkładu prawdopodobieństwa zmiennej losowej y(t) w pewnej chwili t na podstawie znajomości realizacji y(s) tej zmiennej losowej w pewnych innych chwilach s (na ogół chwila t odnosi się do przyszłości).

Procesy ergodyczne

- Jedną z podstawowych własności, dzięki którym można ocenić rozkład prawdopodobieństwa zmiennej losowej y(t) na podstawie obserwacji aktualnych przebiegów danego procesu stochastycznego, jest tzw. własność ergodyczności

Procesy ergodyczne

- Można powiedzieć, że proces stochastyczny jest ergodyczny, jeżeli prawdopodobieństwo zaobserwowania wartości y(t) należącej do jakiegoś zbioru A da się oszacować przez średni czas pobytu każdej realizacji w tym zbiorze podczas długiego czasu obserwacji

Procesy ergodyczne

- Tak więc w procesach stochastycznych ergodycznych można oszacować ich rozkład prawdopodobieństwa na podstawie obserwacji jednego przebiegu w dostatecznie długim czasie, czyli otrzymane wyniki są średnią po czasie.

Procesy ergodyczne

- Hipoteza ergodyczna
- Ewolucja klasycznego złożonego układu dynamicznego zachodzi z jednakowym prawdopodobieństwem przez wszystkie stany, które są dostępne z punktu początkowego i które podlegają ograniczeniom narzuconym przez zasadę zachowania energii.

Procesy Markowa

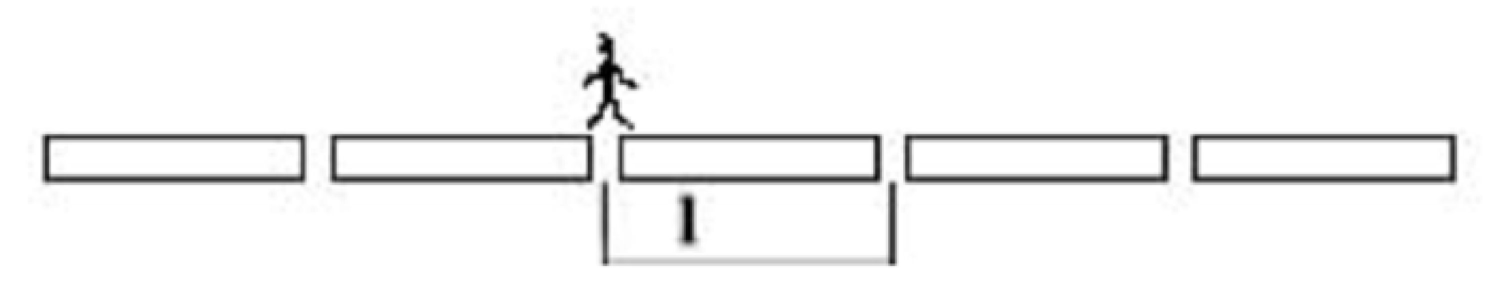
- Przykładem procesu niemarkovskiego może być np. proces zmian poziomu wody w rzece w pewnym ustalonym jej miejscu, gdzie informacja o tym, że w pewnej chwili t poziom wody wynosił y i bezpośrednio przedtem obserwowano np. tendencję obniżania się poziomu wody, pozwala na lepsze przewidywania niż sama informacja o tym, że w chwili t poziom wody wynosił y.

Procesy Markowa

- Przykłady procesów Markowa
  - Proces emisji cząstek wypromieniowanych przez substancję radioaktywną.
  - Ruch cząstki zawieszonej w cieczy tzw. ruch Browna.
  - Proces zajmowania i zwalniania łączy w centrali telefonicznej.
  - Dynamika kolejki w serwerach WWW.

Procesy Markowa

Błądzenie losowe: prawdopodobieństwo znalezienia układu w n-tym stanie zależy tylko od stanu poprzedniego $n-1$.



Rys. 2: Losowe błądzenie pijaka: p - prawdopodobieństwo, że pijak pójdzie w lewo; $q = 1 - p$ - prawdopodobieństwo, że pijak pójdzie w prawo; $x = m\,l$ – lokalizacja pijaka wzdłuż osi x

Procesy Markowa

- Pytanie: jakie jest prawdopodobieństwo, że po wykonaniu N kroków znajdziemy pijaka w położeniu x = ml?
- Niech po n krokach pijak będzie w położeniu x = ml , m <= N.
- Niech nr - liczba kroków w prawo; nl - liczba kroków w lewo.
- Mamy więc

$$N = n_r + n_l$$

$$m = n_r - n_l = n_r - (n - n_r) = 2n_r - N$$

Procesy Markowa

Przypomnienie: rozkład dwumianowy

Prawdopodobieństwo, że pijak pokonał pewną drogę wynosi

$$p \ldots p\, q \ldots q = p^{n_r}\, q^{n_l}$$

Liczba realizacji takich dróg wynosi

$$\frac{N!}{n_r!\, n_l!}$$

więc prawdopodobieństwo wykonania n_r kroków w prawo i n_l kroków w lewo wynosi

$$P_N = \frac{N!}{n_r!\, n_l!}\, p^{n_r}\, q^{n_l}$$

Procesy Markowa

Przyjmijmy

$$n_r = \frac{1}{2}(N+m), \qquad n_l = \frac{1}{2}(N-m)$$

i wstawmy do wzoru na P_N

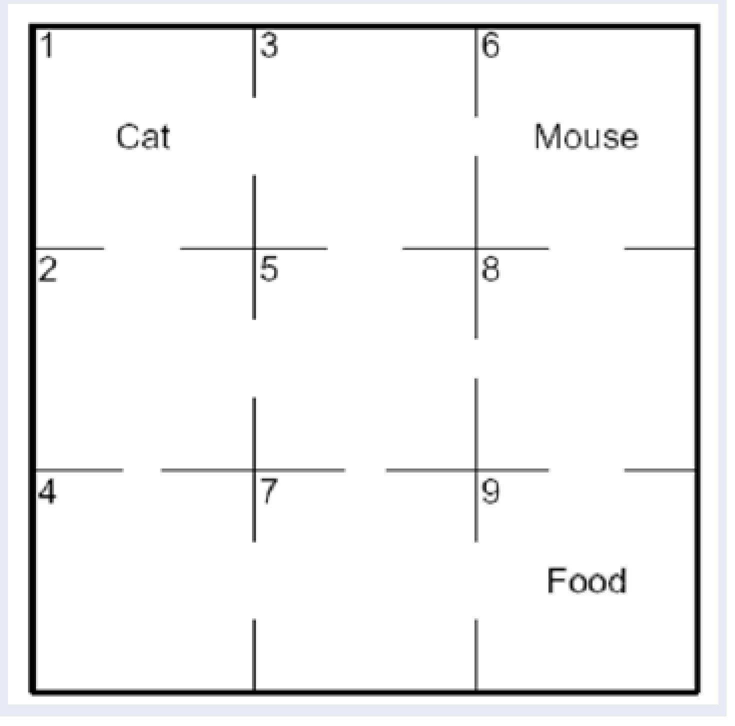
$$P_N(m) = \frac{N!}{\frac{1}{2}(N+m)!\,\frac{1}{2}(N-m)!}\,p^{(N+m)/2}q^{(N-m)/2}$$

Szczególny przypadek $p = q = 1/2$:

$$P_N(m) = \frac{N!}{\frac{1}{2}(N+m)!\,\frac{1}{2}(N-m)!}\left(\frac{1}{2}\right)^N$$

Procesy Markowa

- Mysz w labiryncie



Procesy Markowa

- Mysz w labiryncie
- Mamy kilka możliwości:
  - Kot zawsze siedzi w swojej komórce i czeka na ofiarę
  - Kot może wchodzić tylko do pomieszczeń 1,2,3 i 5, gdyż wszystkie inne otwory są dla niego za małe; do każdego sąsiedniego dopuszczalnego pomieszczenia wchodzi z jednakowym prawdopodobieństwem
  - To samo co powyżej, ale prawdopodobieństwo, że zostanie w swej komórce wynosi 1/2 , a wchodzi do sąsiednich pomieszczeń z prawdopodobieństwem 1/4

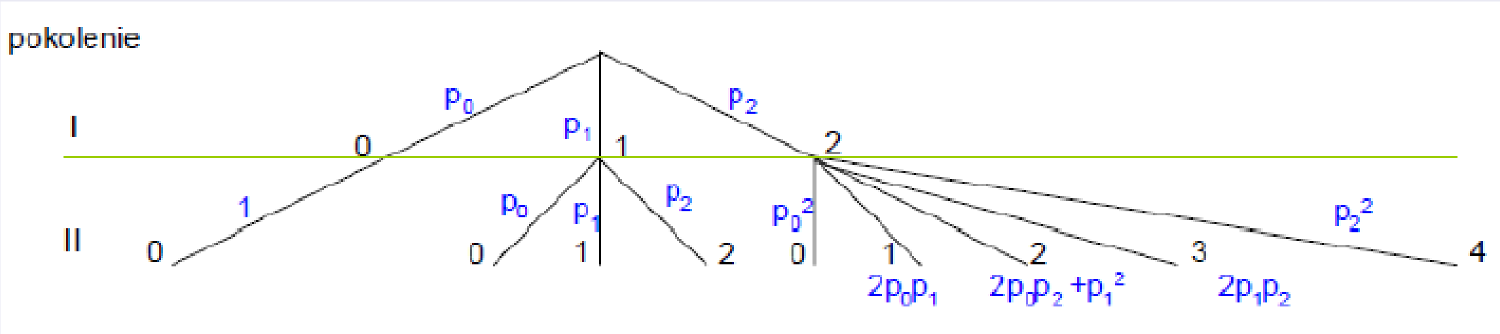
Procesy Markowa

- Procesy gałązkowe (Galtona-Watsona)
- Procesy gałązkowe modelują rozwój populacji (jednopłciowej, rozmnażającej się przez podział, np. bakterii, ameb, monet czy innych mikroorganizmów).
- Zmienne losowe y(n) (przyjmujące nieujemne wartości) określają liczbę osobników w n-tym pokoleniu.
- Przyjmujemy zawsze, że jest jeden protoplasta rodu, czyli $y(0) = 1$.
- Zmienne losowe opisujące, ile dzieci ma każdy osobnik, są niezależne o jednakowym rozkładzie.
- Główne pytanie, jakie się pojawia, to: jakie są szanse, że dana populacja przeżyje?

Procesy Markowa

Procesy gałązkowe (Galtona-Watsona)

Załóżmy, iż osobnik może mieć 0, 1 lub 2 potomków, a przez p_0, p_1 i p_2 oznaczmy prawdopodobieństwa tych zdarzeń ($p_0 + p_1 + p_2 = 1$).



Procesy Markowa

Procesy gałązkowe (Galtona-Watsona)

Niech x_k – liczba osobników k-tej generacji. Wtedy prawdopodobieństwo pojawienia się określonej liczby osobników danej generacji można zapisać:

$$p(x_0 = 1) = 1$$
$$p(x_1 = 0) = p_0 \qquad p(x_1 = 1) = p_1$$
$$p(x_1 = 2) = p_2$$
$$p(x_2 = 0) = p_0 + p_0 p_1 + p_2 p_0^2 \qquad p(x_2 = 1) = p_0 + 2 p_2 p_1 p_0$$
$$p(x_2 = 2) = p_1 p_2 + p_2(2 p_0 p_2 + p_1^2) \qquad p(x_2 = 3) = 2 p_2^2 p_1$$
$$p(x_2 = 4) = p_2^3$$

Jeśli potraktujemy x_k jako zmienną losową, to możemy wyznaczyć jej wartość oczekiwaną ($E(X_k) = \mu$).

Błądzenie losowe

$$B(t+1) = B(t) + z(t+1), \qquad B(0) = B_0$$

$z(t)$ – zakłócenie losowe opisane ciągiem niezależnych zmiennych losowych o jednakowym rozkładzie normalnym o średniej 0 i wariancji 1

t – czas mierzony w jednakowych dyskretnych odstępach

Gdy $t = 0$ – teraźniejszość. Niech $B(0) = 0$, a przyrost czasy $\Delta = \frac{1}{n}$, n – dowolna liczba naturalna.

Błądzenie losowe

Dla nowych jednostek czasu

$$B(t+\Delta) = B(t) + z(t+\Delta), \qquad B(0) = B_0$$

Zmieniła się wariancja (zmienność) zakłócenia losowego $z(t)$, a mianowicie $z(t)$ jest teraz ciągiem niezależnych zmiennych losowych o jednakowym rozkładzie $N(0, \Delta)$. Nowy proces ma takie samo średnie przesunięcie (dryf) i wariancję na przedziale o długości n odstępów (okresów), jak i wyjściowe błądzenie losowe obserwowane na jednostkowym przedziale.
Co się stanie, gdy Δ staje się nieskończenie małą wielkością, czyli dt?

Ruch Browna

- Proces stochastyczny B(t) nazywamy standardowym ruchem Browna (Brownian motion).
- Jest to jeden z ważniejszych modeli teoretycznych w rachunku prawdopodobieństwa.
- Nazwa pochodzi od dobrze znanego w fizyce procesu opisującego położenie cząstki w klasycznym ruchu Browna.

Ruch Browna

- Możemy go przedstawić w następującej postaci całkowej:

$$B(t) = B_0 + \int_0^t dB(s)$$

Ruch Browna

Podstawowe właściwości procesu ruchu Browna:

- prawie wszystkie realizacje $B(t)$ są ciągłe
- $B(t)$ jest procesem o przyrostach stacjonarnych i niezależnych
- przyrosty procesu $B(t)$ mają rozkład normalny $N(0, dt)$
- rozkłady warunkowe $B(u)$ przy danym $B(t)$ są normalne o rozkładzie $N(b(t), u - t)$, dla $u > t$
- wariancja $\mathrm{Var}[B(u)] \to \infty$, gdy $u \to \infty$

Ruch Browna

- Ruch Browna był po raz pierwszy wykorzystany do modelowania procesów finansowych przez Louisa Bacheliera, który w swojej pionierskiej pracy doktorskiej Thèorie de la spéculation, obronionej 29 marca 1900 r. w Paryżu, zaproponował pierwszy teoretyczny model procesu ceny akcji z paryskiej giełdy.

Przykład

- Czy możemy liczyć na kawę ?
- Maszyna z kawą może być czynna (stan 0) lub zepsuta (stan 1).
- Załóżmy, że jeśli maszyna jest czynna w danym dniu, to prawdopodobieństwo zepsucia w dniu następnym jest **d** , a jeśli jest zepsuta w danym dniu, to prawdopodobieństwo jej naprawienia na dzień następny jest **g**.
- Jakie jest prawdopodobieństwo dostania kawy z maszyny?

Przykład

- maszyna jest czynna, wczoraj też była czynna: $p(0, t_n \mid 0, t_n\text{-}1) = 1 - d$,
- maszyna jest nieczynna, wczoraj była czynna: $p(1, t_n \mid 0, t_n\text{-}1) = d$,
- maszyna jest czynna, wczoraj była nieczynna: $p(0, t_n \mid 1, t_n\text{-}1) = g$,
- maszyna jest nieczynna, wczoraj też była nieczynna: $p(1, t_n \mid 1, t_n\text{-}1) = 1 - g$,

Przykład

- Dobra i dobrze serwisowana maszyna powinna mieć **d** bliskie 0 i **g** bliskie 1 !)
- Prawdopodobieństwo dostania kawy z maszyny będzie zależało od proporcji czasu, gdy maszyna jest czynna, do całego czasu pomiaru.

Przykład

- Symulacja dla d=0.2 i g = 0.9

	LICZBA SYMULACJI (DNI)				
	10	50	100	500	1000
0 W DNIU 0	0.90	0.82	0.84	0.84	0.82
1 W DNIU 0	0.50	0.86	0.80	0.80	0.81

- Zauważmy, że uśredniając po długim czasie, prawdopodobieństwo kupienia kawy stabilizuje się na poziomie 80%, praktycznie niezależnie od stanu maszyny w dniu początkowym.

Przykład

- Brawurowa gra

- Mamy 1 zł i chcemy wygrać 5 zł.
- Krupier oferuje nam grę, w której prawdopodobieństwo naszej wygranej wynosi **p** w każdej rundzie z wypłatą podwójnej stawki w razie wygranej oraz jej stratą w razie przegranej, przy czym stawki są w całkowitych wielokrotnościach złotówki.

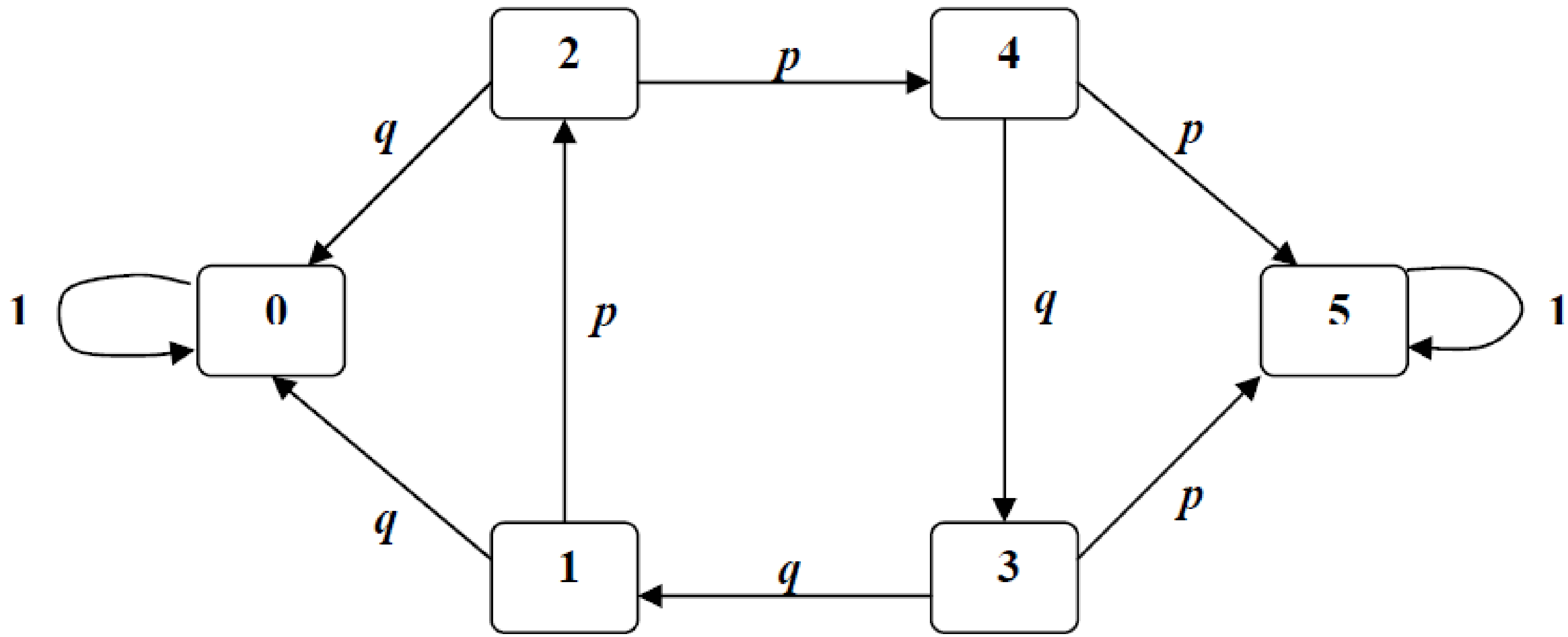
Przykład

- Wybieramy następującą strategie
- brawurowa: w każdej grze stawiamy wszystko co mamy, jeśli ewentualna wygrana pozwoli osiągnąć cel (osiągnąć 5 zł), lub mniej niż cel.
- W przeciwnym razie, stawiamy tyle aby ewentualnie wygrać 5 zł.
- Jaka jest szansa wygrania w **k** lub mniejszej liczbie gier?

Przykład

- Ponumerujmy stany liczbą posiadanych przez nas złotówek.
- Zaczynamy grę od stanu nr 1.
- Wygrać grę – znaczy przejść od stanu 1 do stanu 5.
- Stan 0 oznacza przegraną.
- Prawdopodobieństwo wygrania lub przegrania gry nie zależy od historii wygranych i przegranych w poprzednich grach – własność Markowa jest zatem spełniona.
- Prawdopodobieństwo wygranej w stanie **i** nie zależy od czasu, wiec proces ten jest jednorodny w czasie.

Przykład



Przykład

- Interesują nas tylko ścieżki, które kończą się w stanie 5. Nazwiemy je istotnymi.
- Prawdopodobieństwo każdej ścieżki jest iloczynem prawdopodobieństw przejść jednokrotnych.
- Istotna ścieżka o długości 3 o prawdopodobieństwie $\boldsymbol{p^3}$ S: (1) -> (2) -> (4) -> (5)
- Zatem prawdopodobieństwo wygrania w trzech grach wynosi p^3

Przykład

- Istotna ścieżka o długości 4
- R: (1) -> (2) -> (3) -> (5)
- ma prawdopodobieństwo p^3q.
- Zatem prawdopodobieństwo wygrania w 4 lub mniej grach wynosi $p^3(1+q)$.

Przykład

- Nie ma istotnych ścieżek o długości 6.
- Istotna ścieżka o długości 7 ma jedna pętle
- L: (1) -> (2) -> (4) -> (3) -> (1),
- po czym następuje ścieżka S.

Przykład

- Prawdopodobieństwo pętli L wynosi $\lambda=p^2q^2$, zatem prawdopodobieństwo dla ścieżki L*S wynosi λq^3.
- Zatem prawdopodobieństwo wygrania w 7 lub mniej grach wynosi
- $p^3(1+\lambda)+p^3q$.

Przykład

- Istnieje jedna ścieżka o długości 8: L*R, dla której prawdopodobieństwo wynosi $\lambda\ p^3q$.
- Zatem prawdopodobieństwo wygrania w 8 lub mniej grach wynosi $p^3(1+q)(1+\lambda)$.
- Zauważmy ogólną prawidłowość, że wszystkie dłuższe ścieżki są typu L*…*L*S lub L*…*L*R.
- Ich prawdopodobieństwa przy n pętlach wynoszą
- $\lambda^n p^3$ i $\lambda^n p^3 q$.

Przykład

- Prawdopodobieństwo wygranej przy nieograniczonej liczbie prób wynosi

$$p^3(1+q)(1+\lambda+\lambda^2+\ldots)=\frac{p^3(1+q)}{1-\lambda}$$

Koniec