Stokastiska processer

Definition. En **stokastisk process** är en mängd (familj) av stokastiska variabler X(t).
Parametern t är oftast (men inte alltid) en tidsvariabel.
Processen kallas diskret om X(t) är en diskret s. v. för varje t.
Processen kallas kontinuerlig om X(t) är en kontinuerlig s. v. för varje t.
Parameter t kan också vara kontinuerlig eller "diskret" (ändligt eller uppräkneligt antal t-värden).

Markovkedjor i diskret tid

Vi ska betrakta en diskret stokastisk process $X(t)$ i diskret tid $t \in \{t_0, t_1\ t_2\ \cdots\}$.
Om en fysisk process kan befinna sig i olika tillstånd som vi betecknar med E_k då $X(t_i) = k$ betyder att processen är i tillstånd E_k vid tidpunkten $t = t_i$.
Mängden av alla möjliga tillstånd { E_k} kallas **tillståndsrummet.**

Uttrycket
$P[X(t_{n+1}) = j \mid X(t_n) = i]$
betecknar övergångssannolikheten (transition probability) att systemet som är i tillståndet E_i vid tidpunkten t_n befinner sig i E_j vid nästa tidpunkt t_{n+1}.

Definition.
Markovkedja i diskret tid är en diskret stokastisk process med diskret tid som uppfyller Markov-villkoret:
För $t_0 < t_1 < t_2 < \cdots < t_n$,

$$P[X(t_{n+1}) = j \mid X(t_n) = i, \ldots, X(t_1) = i_1, X(t_0) = i_0] = P[X(t_{n+1}) = j \mid X(t_n) = i].$$

Minneslösheten: Markovegenskapen (Markov – villkoret) betyder att övergångssannolikheten $P[X(t_{n+1}) = j \mid X(t_n) = i]$ beror endast av ”nu-läge” d.v.s. situationen vid tidpunkten t_n och inte av vägen till detta tillstånd. Vi säger att processen är **minneslös.**

Definition. En Markovkedja är **homogen** om övergångssannolikheten $P[X(t_{n+1}) = j \mid X(t_n)] = i]$ **beror ej** av tiden t_n utan bara av tillstånd E_i, E_j .
d vs
$P[X(t_1) = j \mid X(t_0)] = i] = P[X(t_2) = j \mid X(t_1) = i] = \ldots = P[X(t_{n+1}) = j \mid X(t_n) = i]$.

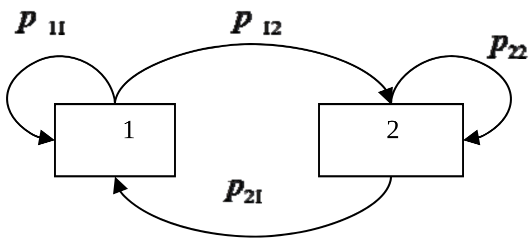
Vi ska i fortsättningen endast betrakta **homogena** Markovkedjor.
För en **homogen** Markovkedja kan vi beteckna

$$p_{ij} = P(E_i \to E_j) = P[X(t_{n+1}) = j \mid X(t_n) = i]$$ (**övergångssannolikheten**)

och definiera **övergångsmatrisen (transition probability matrix):**

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} & \cdots \\ p_{21} & p_{22} & p_{23} & \cdots \\ p_{31} & p_{32} & p_{33} & \cdots \\ \vdots & & & \end{bmatrix}$$

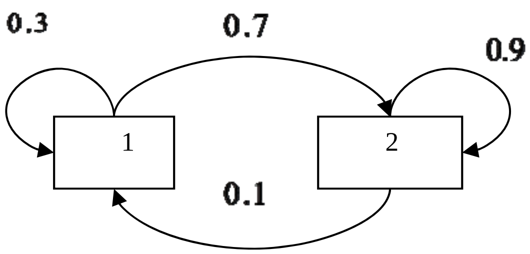
där $\sum_k p_{ik} = 1$ för varje i, d v s summan av alla element i en rad är lika med 1.
Ett sätt att åskådligt beskriva en Markovkedja är att använda en riktad graf med **övergångssannolikheterna .**



Från grafen kan vi på enkelt sätt definiera matrisen och omvänt, om grafen inte innehåller för många pilar.
För ovanstående graf kan vi ange tillhörande övergångsmatris

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}.$$

Exempel 1. Låt $P = \begin{bmatrix} 0.3 & 0.7 \\ 0.1 & 0.9 \end{bmatrix}$ vara övergångsmatrisen för en Markovkedja med 2 tillstånd.
Rita motsvarande graf med övergångssannolikheter.
Svar:



Absoluta sannolikheter.
Den absoluta sannolikheten att processen befinner sig i tillståndet E_i vid $t = t_n$ betecknas med $p_i(n)$.
Alltså
$p_i(n) = P(X(t_n) = i)$

Därmed t ex $p_2(4) = P(X(t_4) = 2)$ är sannolikheten att processen befinner sig i tillståndet E_2 vid $t = t_4$.
$p_2(0)$ är sannolikheten att processen befinner sig i tillståndet E_2 vid starttiden $t = 0$.

Sannolikhetsvektor :
Vid tiden $t = t_n$ befinner sig processen i ett av tillstånden E_1, E_2, ... med motsvarande sannolikheterna, $p_1(n), p_2(n),\dots$ som vi kan samla i en radvektor. Denna vektor kallar vi sannolikhetsvektor.
Alltså
$\vec{p}(n) = (p_1(n), p_2(n),\dots)$ är en sannolikhetsvektor (probability vector). Notera att **summan av alla koordinater** i en sannolikhetsvektor **är lika med 1.**
$\vec{p}(0) = (p_1(0), p_2(0),\dots)$ är en initial sannolikhetsvektor (**start-sannolikhetsvektor**) som visar att processen startar i E_1 med sannolikheten $p_1(0)$, i E_2 med sannolikheten $p_2(0)$ o s v. Om processen säkert startar i ett visst tillstånd E_k då är motsvarande startsannolikhet lika med 1 och alla andra med 0.
T ex $\vec{p}(0) = (0,1,.0,0\dots)$ visar att processen säkert startar i E_2 (d v s med sannolikheten 1)

Relationer mellan $\vec{p}(n)$, $\vec{p}(0)$ och P:
$\vec{p}(1) = \vec{p}(0)P$
$\vec{p}(2) = \vec{p}(1)P = \vec{p}(0)P^2$
$\vec{p}(3) = \vec{p}(2)P = \vec{p}(0)P^3$
....
$\vec{p}(n) = \vec{p}(n-1)P$
$\vec{p}(n) = \vec{p}(0)P^n$

Med ovanstående relationer kan man studera en Markovkedja med hjälp av startvektorn $\vec{p}(0)$ och övergångsmatrisen P.

Stationära sannolikhetsvektorer

Definition. En sannolikhetsvektor $\vec{q} = (q_1, q_2, \dots.)$ kallas **stationär sannolikhetsvektor** (steady -state probability vector) om den satisfierar ekvationen
$$\vec{q} = \vec{q}P.$$

Om processen startar med en **stationär** sannolikhetsfördelning $\vec{p}(0) = \vec{q}$ då blir det
$\vec{p}(1) = \vec{q}P = \vec{q}$

$\vec{p}(2) = \vec{p}(1)P = \vec{q}P = \vec{q}$

......

$\vec{p}(n) = \vec{q}$ för alla n.

För att bestämma (eventuella) stationära sannolikhetsvektorer skriver vi vektorekvationen $\vec{q}P = \vec{q}$ på komponentform och **lägger till villkoret**

$$q_1 + q_2 + \cdots = 1.$$

Med andra ord löser vi systemet:

$$\vec{q}P = \vec{q}$$

$$q_1 + q_2 + \cdots = 1 \quad (\vec{q} \text{ är en sannolikhetsvektor, summa koordinater } = 1)$$

==

ÖVNINGAR

Uppgift 1. Låt P beteckna övergångsmatrisen för en Markovkedja i diskret tid. Låt $\vec{p}(n)$ beteckna den (transienta) sannolikhetsvektorn, d v s sannolikheten att Markovprocessen vid tidpunkten t=n befinner sig i tillstånd E_1, E_2.

Om startvektorn $\vec{p}(0) = (0.3, 0.7)$ så

i) bestäm $\vec{p}(1)$, $\vec{p}(2)$, $\vec{p}(3)$

ii) bestäm sannolikheten att processen är i tillståndet E_1 i tidpunkten t=3

a) $P = \begin{bmatrix} 0.8 & 0.2 \\ 0.5 & 0.5 \end{bmatrix}$ b) $P = \begin{bmatrix} x & 0.3 \\ 0.4 & y \end{bmatrix}$ c) $P = \begin{bmatrix} x & 0.4 \\ y & 0.11 \end{bmatrix}$

d) $P = \begin{bmatrix} x & 0.2 \\ 0.3 & y \end{bmatrix}$

Svar :

a) i) $\vec{p}(1) = (0.59, 0.41)$, $\vec{p}(2) = (0.677, 0.323)$, $\vec{p}(3) = (0.7031, 0.2969)$

i i) 0.7031

b) i) $\vec{p}(1) = (0.49, 0.51)$, $\vec{p}(2) = (0.547, 0.453)$, $\vec{p}(3) = (0.5641, 0.4359)$

i i) 0.5641

c) i) $\vec{p}(1) = (0.803, 0.197)$, $\vec{p}(2) = (0.657, 0.343)$, $\vec{p}(3) = (0.699, 0.301)$

i i) 0.699

d) i) $\vec{p}(1) = (0.45, 0.55)$, $\vec{p}(2) = (0.525, 0.475)$, $\vec{p}(3) = (0.5625, 0.4375)$

i i) 0.5625

Uppgift 2. Låt P beteckna övergångsmatrisen för en Markovkedja i diskret tid med tre tillstånd E_1, E_2 och E_3.

Låt vidare startvektorn $\vec{p}(0) = (0.3, 0.2, 0.5)$.

i) Bestäm $\vec{p}(1)$

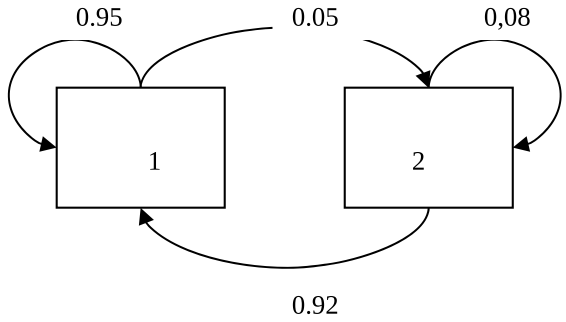
ii) Bestäm sannolikheten att processen är i tillståndet E_2 i tidpunkten t=1.

a) $P=\begin{bmatrix} 0.1 & 0.9 & 0 \\ 0 & 0.5 & 0.5 \\ 0 & 0.3 & 0.7 \end{bmatrix}$ b) $P=\begin{bmatrix} 0.2 & x & 0 \\ 0 & y & 0.4 \\ 0 & z & 0.7 \end{bmatrix}$ c) $P=\begin{bmatrix} 0.2 & x & 0 \\ 0.1 & y & 0 \\ 0 & z & 0 \end{bmatrix}$

Svar:
a) i) $\vec{p}(1)=(0.03,0.52,0.45)$ ii) 0.52
b) i) $\vec{p}(1)=(0.06,0.51,0.43)$ ii) 0.51
c) i) $\vec{p}(1)=(0.08,0.92,0)$ ii) 0.92

Uppgift 3. Ett system kan betraktas som en Markovkedja med två tillstånd E1 och E2 Om systemet under en dag är i tillstånd E1 så övergår systemet nästa dag till E2 med sannolikheten 0.05. Om systemet under en dag befinner sig i E2, är systemet nästa dag i samma tillstånd, E2, med sannolikheten 0.08.
a) Bestäm övergångsmatrisen
b) Rita tillståndsdiagrammet med övergångssannolikheterna.
c) Bestäm sannolikheten att systemet är i tillståndet E1 efter 3 övergångar om systemet startar i E2

Svar:
a) $P=\begin{bmatrix} 0.95 & 0.05 \\ 0.92 & 0.08 \end{bmatrix}$
b) se figuren



c)

Startvektorn är $\vec{p}(0)=(0, 1)$ (eftersom systemet startar i E2).
Vi beräknar

$$\vec{p}(1)=\vec{p}(0)P=(0,1)\begin{bmatrix} 0.95 & 0.05 \\ 0.92 & 0.08 \end{bmatrix}=(0.92,0.08)$$

$$\vec{p}(2)=\vec{p}(1)P=(0.92,0.08)\begin{bmatrix} 0.95 & 0.05 \\ 0.92 & 0.08 \end{bmatrix}=(0.9476, \quad 0.0524)$$

$$\vec{p}(3)=\vec{p}(2)P=(0.9476, \quad 0.0524)\begin{bmatrix} 0.95 & 0.05 \\ 0.92 & 0.08 \end{bmatrix}=(0.948428, \quad 0.051572)$$

Sannolikheten att systemet är i tillståndet E1 efter 3 övergångar är 0.948428
Svar c: 0.948428

Exempel. En Markovkedja har övergångsmatrisen P= $\begin{bmatrix} 0.3 & 0.7 \\ 0.5 & 0.5 \end{bmatrix}$

Bestäm alla stationära sannolikhetsvektorer.

Lösning:

Låt $\vec{q} = (x, y)$ vara en stationär sannolikhetsvektor.
Då gäller

$$\vec{q}P = \vec{q} \qquad \text{och}$$
$$x + y = 1$$

Vi skriver $\vec{q}P = \vec{q}$ på komponentform:

$$(x, y)\begin{bmatrix} 0.3 & 0.7 \\ 0.5 & 0.5 \end{bmatrix} = (x, y) \Rightarrow \begin{array}{l} 0.3x + 0.5y = x \\ 0.7x + 0.5y = y \end{array}$$

och lägger till ekvationen

$$x + y = 1 \qquad (\vec{q} \text{ är en sannolikhetsvektor})$$

Därmed har vi systemet:

$$\begin{cases} 0.3x + 0.5y = x \\ 0.7x + 0.5y = y \\ x + y = 1 \end{cases} \Rightarrow \begin{cases} -0.7x + 0.5y = 0 \\ 0.7x - 0.5y = 0 \\ x + y = 1 \end{cases}$$

Andra ekvationen är samma som första.
Från första ekvationen har vi $y = \frac{7x}{5}$ som vi substituerar i tredje ekvationen och får

$$x + \frac{7x}{5} = 1 \Rightarrow \frac{12x}{5} = 1 \Rightarrow x = \frac{5}{12}.$$

Svar: $\vec{q} = (5/12, \quad 7/12)$

Uppgift 4. En Markovkedja har övergångsmatrisen

a) P= $\begin{bmatrix} 0.2 & 0.8 \\ 0.6 & 0.4 \end{bmatrix}$ b) P= $\begin{bmatrix} 0.1 & 0.9 \\ 0.6 & 0.4 \end{bmatrix}$ c) P= $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

Bestäm alla stationära sannolikhetsvektorer.

Svar: a) $\vec{q} = (3/7, \quad 4/7)$ b) $\vec{q} = (2/5, \quad 3/5)$ c) oändligt många lösningar
$\vec{q} = (t, \quad 1-t)$ där $0 \le t \le 1$.

Uppgift 5. En Markovkedja har övergångsmatrisen.

$$P = \begin{bmatrix} \frac{1}{10} & \frac{1}{5} & \frac{7}{10} \\ \frac{2}{5} & \frac{3}{10} & \frac{3}{10} \\ \frac{1}{5} & \frac{1}{2} & \frac{3}{10} \end{bmatrix}$$

Bestäm alla stationära sannolikhetsvektorer.

Svar: $\vec{q} = (\frac{17}{69}, \frac{49}{138}, \frac{55}{138})$