11. Stokastiska processer med diskret tid

11.1. Inledning

Teorin för stokastiska processer (eng: stochastic processes) är en betydelsefull del av sannolikhetsteorin. Den har utvecklats mycket snabbt under senare decennier och har många tillämpningar, bl a inom naturvetenskap och teknik. Man använder den t ex för att studera, hur bakteriepopulationer utvecklas, hur radioaktivt sönderfall sker, hur påkänningen på en bilaxel varierar vid körning, hur trafiken varierar i en vägkorsning eller hur köer uppkommer på ett postkontor.

Teorin är både omfattande och invecklad, och denna bok handlar bara om några grundläggande delar (se litteraturförteckningen för utförligare redogörelser). Av praktiska skäl uppdelas framställningen i två kapitel, 11 och 12, som ägnas åt processer med diskret tid resp kontinuerlig tid.

I § 11.2 ges exempel och allmänna definitioner, § 11.3 och § 11.4 handlar om Markov-kedjor och § 11.5 om förgreningsprocesser.

11.2. Exempel och definitioner

Vad en stokastisk process med diskret tid är, inser man enklast genom att se på några exempel.

Exempel 1. Slumpvandring mellan heltalspunkter

I § 9.2 berättades det om slumpvandring. Vi skall beskriva situationen om igen i något annorlunda ordalag och anlägga nya aspekter på den: Betrakta en 'punkt' som med start i origo vid $t=0$ rör sig vid tidpunkterna $t=1, 2, \ldots$ regelbundet ett steg åt höger samtidigt som den förskjuts uppåt enligt följande regel: Förskjutningen är 1 med sannolikheten p och 0 med sannolikheten $q=1-p$; förskjutningarna vid olika tidpunkter sker oberoende av varandra. Punkten kommer alltså att företa en slumpvandring mellan heltalspunkter, så som den heldragna brutna kurvan i Fig 11:1 anger.

Om man låter punkten starta på nytt, får man en ny kurva; se den streckade kurvan i figuren. Om man flera gånger utför samma procedur, får man kurvor vars utseende slumpmässigt kommer att skilja sig mer eller mindre från varandra.

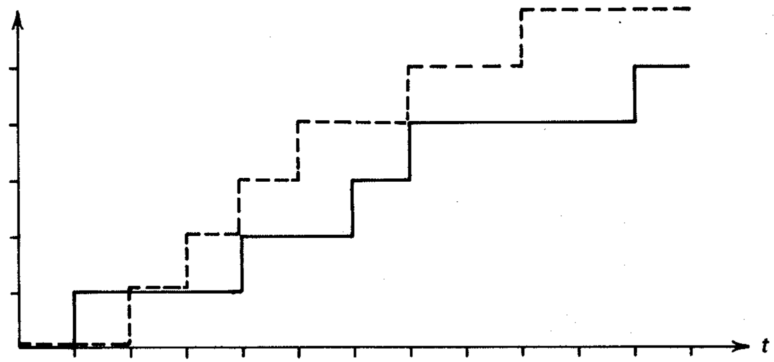

Fig 11:1. Slumpvandring mellan heltalspunkter.

Det är sådana »slumpkurvor» som man studerar inom teorin för stokastiska processer. För att få en modell för hur de uppkommer inför man för varje heltalstidpunkt t en s.v. $X(t)$ som anger den totala förskjutningen i lodrät led räknat från den horisontella axeln.

Så länge man betraktar en enda tidpunkt t, är situationen välbekant. Vi får då en s.v. $X(t)$ som är $Bin(t, p)$; jfr § 9.2. Det nya är att vi, för att kunna studera slumpkurvornas hela förlopp, *samtidigt* måste betrakta flera tidpunkter $t=1, 2, \ldots$. Slumpkurvornas uppförande bestäms tydligen av utfallet av en följd $X(1), X(2), \ldots$ av beroende s.v. Denna följd eller, som man ofta säger, familj av s.v. $\{X(t), t=1, 2, \ldots\}$ är ett exempel på en stokastisk process med diskret tid.

Om vi tänker oss att slumpvandringen slutar vid tidpunkten $t=k$, blir den stokastiska processen något välbekant: familjen $\{X(t), t=1, 2, \ldots, k\}$ blir en k-dimensionell s.v. □

Exempel 2. Allmän slumpvandring

Vi generaliserar nu reglerna för slumpvandringen i Ex 1 så till vida att förskjutningarna i lodrät led vid $t=1, 2, \ldots$ antas vara oberoende s.v. $Y_1, Y_2, \ldots$ med en generell fördelning. Denna fördelning kan vara diskret eller kontinuerlig, och vi tillåter negativa värden på förskjutningen. I Fig 11:2 antyds hur slumpvandringen kan komma att försiggå.

Liksom förut föranleds vi att införa en familj $\{X(t), t=1, 2, \ldots\}$ av s.v. Den stokastiska process med diskret tid som vi då får har många praktiska tillämpningar. Låt oss ta ett ekonomiskt exempel: $X(t)$ är ett företags tillgångar i kronor vid tidpunkten t och Y_t alltså ändringen av tillgångarna från tidpunkten $t-1$ till t. Som exempel på ett problem som då kan vara av intresse kan man nämna följande: Hur stor är sannolikheten att företaget blir ruinerat vid någon av tidpunkterna $t=1, 2, \ldots, k$, dvs annorlunda uttryckt: Hur stor är sannolikheten att $X(t)<0$ för något $t=1, \ldots, k$? (Detta är ett svårt problem, som vi inte kommer att närmare diskutera i denna bok.)

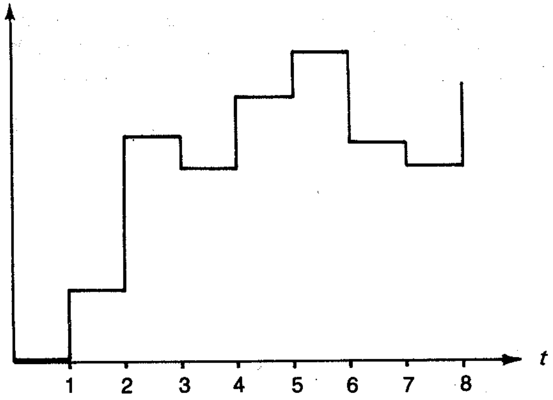

Fig 11:2. Allmän slumpvandring. □

Exemplen inspirerar oss till följande

Definition. En stokastisk process med diskret tid är en familj $\{X(t)\}$ av ändligt eller uppräkneligt oändligt antal s.v. □

I stället för »diskret tid» säger man ibland »diskret parameter». Orsaken 013
härtill är att variabeln t inte alltid är en tidsvariabel:

Exempel 3. Tjockleksvariationer hos garn

Tjockleken hos garn mäts i punkter $t=1, 2, \ldots$ räknat från en lämplig nollpunkt (se 014
Fig 11:3). Om $X(t)$ är tjockleken i punkten t, kan tjockleksvariationerna beskrivas med hjälp av en stokastisk process $\{X(t), t=1, 2, \ldots\}$.

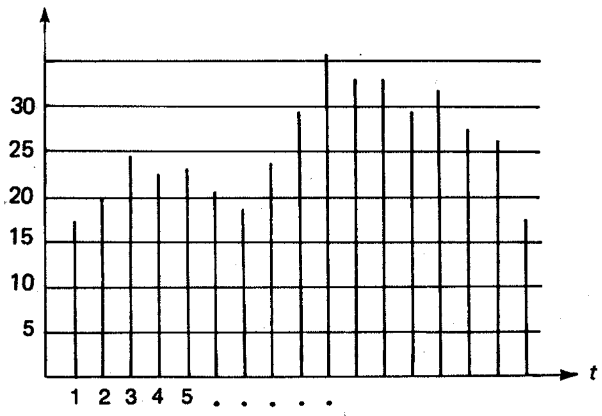

Fig 11:3. Tjockleksvariationer hos garn.

Lägg märke till att beroendet mellan olika medlemmar i familjen i detta fall kommer att avhänga av deras inbördes avstånd, särskilt om det konstanta avståndet mellan punkterna är litet. Närbelägna s.v. $X(t)$ och $X(t+1)$ kommer att vara starkt beroende, medan beroendet mellan s.v. på längre avstånd blir svagare. Vid modellbygget måste man alltså konstruera processen så att den har denna egenskap. □

Man måste klart skilja mellan den stokastiska processen och de värden som den antar då den observeras vid ett visst tillfälle (på samma sätt som man måste skilja på en s.v. och ett av dess utfall). För att poängtera detta inför vi följande terminologi: Den slumpkurva som uppstår, då processen observeras en gång, kallas en *realisering* av processen (eng: realization). I Fig 11:1, 11:2 och 11:3 återges alltså realiseringar av resp processer.

Vi skall införa ytterligare några termer. En stokastisk process kallas *växande* eller *avtagande,* allt eftersom processens realiseringar är växande eller avtagande.

Viktig är följande

Definition. Om alla s.v. som ingår i processen är diskreta, föreligger det en *diskret stokastisk process;* om alla s.v. är kontinuerliga, föreligger det en *kontinuerlig stokastisk process.* □

I Ex 1 har vi alltså ett exempel på en diskret stokastisk process med diskret tid, i Ex 3 ett exempel på en kontinuerlig stokastisk process med diskret tid.

Det finns många intressanta speciella typer av stokastiska processer med diskret tid. Två sådana typer behandlas i fortsättningen, nämligen Markov-kedjor och förgreningsprocesser.

11.3. Markov-kedjor

Teorin för Markov-kedjor är omfattande och användbar i många praktiska sammanhang. De grundläggande egenskaperna behandlas i detta och följande avsnitt.

(a) Definitioner och exempel

En *Markov-kedja* $\{X(t), t=0, 1, 2, \ldots\}$ är en diskret stokastisk process med diskret tid som uppfyller *Markov-villkoret:* För alla $n \geqslant 2$ gäller att

$$
\begin{aligned}
(1) \qquad & P[X(n)=x_n | X(0)=x_0, X(1)=x_1, \ldots, X(n-1)=x_{n-1}]= \\
& =P[X(n)=x_n | X(n-1)=x_{n-1}] \; .
\end{aligned}
$$

Informellt uttryckt innebär detta följande: Om man känner processens värde $X(n-1)$ vid tidpunkten $t=n-1$ och önskar uttala sig om dess värde $X(n)$ vid nästa tidpunkt $t=n$, har man ingen glädje av att dessutom känna processens värden $X(0), \ldots, X(n-2)$ för föregående tidpunkter; detta ger ingen ytterligare information. (Beträffande beteckningssättet i (1) jfr § 2.5.)

De olika värden som $X(t)$ kan anta kallas *tillstånd* och betecknas E_1, E_2, $E_3, \ldots$. Denna terminologi är praktisk, ty man kan då inkludera situationer då processen inte antar kvantitativa värden. (Exempel: E_1=elev sover, E_2=elev äter, E_3=elev på föreläsning, . . .). Om antalet tillstånd är ändligt, säges kedjan vara *ändlig*.

En realisering av en Markov-kedja kan se ut som i Fig 11:4:

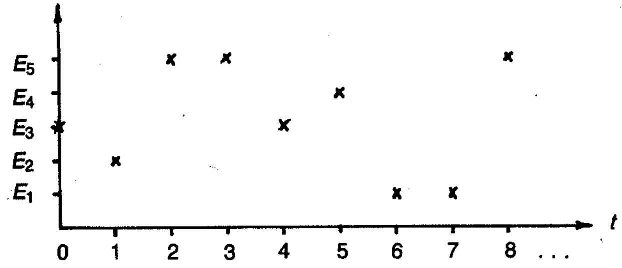

Fig 11:4. Realisering av Markov-kedja.

Sannolikheten att processen vid $t=n$ är i ett visst tillstånd E_i skrivs

(2) $$p_i^{(n)}=P(X(n)=E_i)$$

och kallas en *absolut sannolikhet*. Givetvis är $\sum_i p_i^{(n)}=1$. Av stor betydelse är, som framgår av Markov-villkoret, sannolikheter av typen

(3) $$p_{ij}(n)=P[(X(n)=E_j|X(n-1)=E_i] \ .$$

Detta kallas en *övergångssannolikhet* och är den betingade sannolikheten att processen skall vara i E_j vid $t=n$, om den var i E_i vid $t=n-1$. I fortsättningen antages att övergångssannolikheterna inte beror av n, så att man kan skriva $p_{ij}(n)=p_{ij}$. Matrisen

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

kallas *övergångsmatrisen*. Varje radsumma $\sum_{j=1}^{\infty} p_{ij}$ i en sådan matris är 1, ty denna summa anger sannolikheten att, om processen vid en tidpunkt befinner sig i E_i, den vid nästa tidpunkt övergår till något av tillstånden E_1, $E_2, \ldots$, och det är ju en säker händelse.

Om speciellt kedjan är ändlig och omfattar, säg, N olika tillstånd, är P en $N\times N$ matris.

Exempel 4. Partibyte

I ett land finns två partier, V och H. En person, som deltar i val vid tidpunkterna $t=0, 1, 2, \ldots$, kan varje gång befinna sig i ettdera av två tillstånd: E_1=rösta på V, E_2=rösta på H. Om han vid ett visst val röstar på V, antas sannolikheten vara $1-\beta$ att han även nästa gång röstar på V. Om han däremot röstar på H, är sannolikheten $1-\alpha$ att han även nästa gång röstar på H. Tidigare val antas inte påverka hans uppförande. Under dessa förutsättningar regleras personens partival av en Markov-kedja med övergångsmatrisen

$$P=\begin{bmatrix} 1-\beta & \beta \\ \alpha & 1-\alpha \end{bmatrix}.$$

Vi antar att $0<\alpha<1$, $0<\beta<1$. □

Exempel 5. Hasardspel med begränsade kassor

Spelarna A och B kastar krona och klave med ett välgjort mynt. Om krona kommer upp, får A 1 krona av B. Om klave kommer upp får B 1 krona av A. Spelarna har från början 2 kronor vardera. Spelet slutar då någon av spelarna är ruinerad. Spelets förlopp kan beskrivas med en Markov-kedja med fem tillstånd E_0, $E_1, \ldots, E_4$, där E_i=»A:s kassa är i kronor». Övergångsmatrisen är

$$P=\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Då processen befinner sig i E_0 eller E_4 är A resp B ruinerad. □

När $p_{jj}=1$ kallas E_j ett *absorberande* tillstånd. I Ex 5 är E_0 och E_4 absorberande. När processen hamnat i ett sådant tillstånd kan den inte komma ur det igen.

Exempel 6. Ehrenfests diffusionsmodell

I två behållare A och B finns det inalles a molekyler. Vid var och en av tidpunkterna $t=1, 2, \ldots$ väljer man en molekyl slumpmässigt och flyttar över den till den andra behållaren. Låt E_i $(i=0, 1, \ldots, a)$ betyda att det finns i stycken molekyler i A och alltså

$a-i$ stycken i B. Antag att systemet vid $t=n-1$ befinner sig i E_i. Nästa gång kan det befinna sig i E_{i+1} eller i E_{i-1}. Den förstnämnda förändringen inträffar om man väljer en molekyl i B, vilket tydligen inträffar med sannolikheten $(a-i)/a$, och den sistnämnda om man väljer en i A, vilket inträffar med sannolikheten i/a. Övergångsmatrisen blir (om vi för tydlighets skull sätter ut index för E i kanterna)

$$P=\begin{array}{c|ccccccccc|}
 & 0 & 1 & 2 & 3 & 4 & \dots & a-2 & a-1 & a \\
\hline
0 & 0 & 1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 \\
1 & \frac{1}{a} & 0 & \frac{a-1}{a} & 0 & 0 & \dots & 0 & 0 & 0 \\
2 & 0 & \frac{2}{a} & 0 & \frac{a-2}{a} & 0 & \dots & 0 & 0 & 0 \\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\
a-1 & 0 & 0 & 0 & 0 & 0 & \dots & \frac{a-1}{a} & 0 & \frac{1}{a} \\
a & 0 & 0 & 0 & 0 & 0 & \dots & 0 & 1 & 0
\end{array}$$

□

(b) Övergångssannolikheter av r:te ordningen

Storheterna p_{ij} i övergångsmatrisen anger sannolikheterna för övergång i ett steg från ett tillstånd till ett annat. Vi kan uttrycka detta genom att skriva $p_{ij}=P(E_i \to E_j \text{ i } 1 \text{ steg})$. Allmännare definierar vi *övergångssannolikheter av r:te ordningen*

(4) $$p_{ij}^{(r)}=P(E_i \to E_j \text{ i } r \text{ steg}) \qquad (r=1, 2, \dots).$$

Låt oss först beräkna övergångssannolikheterna av 2:a ordningen. För att processen i två steg skall komma från E_i till E_j måste den passera något tillstånd E_v efter ett steg, dvs vi måste ha $E_i \to E_v \to E_j$. För varje fixt v är, enligt definitionen på en Markov-kedja, tillhörande sannolikhet lika med $p_{iv}p_{vj}$. Summering över alla tänkbara v ger

$$p_{ij}^{(2)}=\sum_v p_{iv}p_{vj}\,.$$

Detta är ju regeln för matrismultiplikation, och det är tydligen praktiskt att införa matrisen

$$P^{(2)}=(p_{ij}^{(2)}).$$

Vi har i själva verket visat att

$$P^{(2)}=P^2.$$

Allmänt inför vi matrisen

(5) $$P^{(r)}=(p_{ij}^{(r)}).$$

Då gäller

Sats 1. $P^{(r)}=P^r$. □

Man kan bevisa satsen t ex med induktion.

Vidare gäller

Sats 2. (»Chapman–Kolmogorovs sats»)

$$P^{(r+s)}=P^{(r)}P^{(s)}.$$ □

Bevis: Man kan använda Sats 1 eller resonera så här: $E_i \to E_j$ i $r+s$ steg inträffar om $E_i \to E_v$ i r steg och därefter $E_v \to E_j$ i s steg. För fixt v är tillhörande sannolikhet $p_{iv}^{(r)}p_{vj}^{(s)}$. Summation över v ger

$$p_{ij}^{(r+s)}=\sum_v p_{iv}^{(r)}p_{vj}^{(s)}$$

och därmed är satsen bevisad. □

Exempel 4. Partibyte (forts)

Vi har enligt Sats 1

$$P^{(2)}=\begin{bmatrix}1-\beta & \beta \\ \alpha & 1-\alpha\end{bmatrix}\begin{bmatrix}1-\beta & \beta \\ \alpha & 1-\alpha\end{bmatrix}=\begin{bmatrix}(1-\beta)^2+\alpha\beta & \beta(1-\beta)+\beta(1-\alpha) \\ \alpha(1-\beta)+\alpha(1-\alpha) & \alpha\beta+(1-\alpha)^2\end{bmatrix}$$

varur övergångssannolikheterna av 2:a ordningen kan avläsas. □

(c) Absoluta sannolikheter

Vi antar att processen vid $t=0$ startar i E_i med en given sannolikhet $p_i^{(0)}$. Radvektorn (observera: litet p)

(6) $$p^{(0)}=(p_1^{(0)}, p_2^{(0)}, \dots)$$

kallas *startfördelningen* eller *startvektorn*. Summan av sannolikheterna i denna vektor är givetvis 1. (Om processen säkert startar i ett visst tillstånd, säg i E_1, är $p_1^{(0)}$ lika med 1 och övriga sannolikheter 0.)

Med hjälp av startvektorn och övergångsmatrisen kan man i princip studera processens uppförande. Låt, liksom i (2), $p_i^{(n)}$ betyda sannolikheten att processen är i E_i vid $t=n$. Inför, analogt med startvektorn, radvektorn

(7) $\quad p^{(n)}=(p_1^{(n)}, p_2^{(n)}, \ldots).$

Med samma resonemang som vid beviset av Sats 2 inses att

(8) $\quad p^{(n)}=p^{(0)}P^n.$

Exempel 4. Partibyte (forts)

Antag att den okunnige väljaren vid första val efter uppnådd »valmyndighetsålder», kastar krona och klave med ett välgjort mynt men sedan håller fast vid V-partiet med sannolikheten 0.90 och vid H-partiet med sannolikheten 0.70. Vid $t=1$ får vi

$$p^{(1)}=(1/2\ \ 1/2)\begin{bmatrix}0.90 & 0.10\\ 0.30 & 0.70\end{bmatrix}=(0.60\quad 0.40).$$

Med sannolikheten 0.60 väljer personen alltså V-partiet, med 0.40 H-partiet. □

Vi skall nu införa begreppet stationär fördelning. Låt

$$\pi=(\pi_1, \pi_2, \ldots)$$

vara en sannolikhetsvektor. Starta processen genom att välja $p^{(0)}=\pi$. Om $p^{(n)}=\pi$ för alla $n=1, 2, \ldots$, säges π vara en *stationär fördelning*. Sannolikheten att processen är i ett visst tillstånd E_i är då alltid densamma när den än observeras (nämligen lika med π_i). Om man tar $n=1$ i (8) ser man att π är lösning till ekvationen

$$\pi=\pi P.$$

Denna lösning är ofta, men inte alltid, entydigt bestämd. Vi återkommer till denna ekvation senare.

Exempel 4. Partibyte (forts)

Låt oss se på övergångsmatrisen

$$P=\begin{bmatrix}1-\beta & \beta\\ \alpha & 1-\alpha\end{bmatrix}.$$

Ekvationen $\pi=\pi P$ blir utskriven

$$\pi_1=\pi_1(1-\beta)+\pi_2\alpha$$
$$\pi_2=\pi_1\beta+\pi_2(1-\alpha).$$

Den satisfieras av sannolikhetsvektorn

$$\pi=(\alpha/(\alpha+\beta),\ \beta/(\alpha+\beta))$$

som alltså är en stationär fördelning. Lösningen är entydig. □

11.4. Mer om Markov-kedjor

(a) Klassificering av en Markov-kedja och dess tillstånd

Vi skall berätta om beständiga och obeständiga tillstånd samt om irreducibla Markov-kedjor.

Låt processen starta i ett visst tillstånd E_i. Kommer den att återvända dit någon gång? Svaret är trefaldigt: Ibland sker det säkert, ibland med en viss sannolikhet mellan 0 och 1, ibland inte alls.

Sannolikheten att processen vid tidpunkten $t=n$ för första gången efter starten i E_i på nytt befinner sig i E_i betecknas $f_{ii}^{(n)}$. Summan

$$f_{ii}=f_{ii}^{(1)}+f_{ii}^{(2)}+\ldots$$

anger sannolikheten att processen förr eller senare på nytt befinner sig i E_i.

Definition. Om $f_{ii}=1$ säges tillståndet E_i vara *beständigt*. Om $f_{ii}<1$ säges E_i vara *obeständigt*. □

(De engelska termerna är recurrent, ibland persistent, resp transient.) Om E_i är beständigt, återkommer processen säkert till E_i. När den återvänder, startar processen på nytt från E_i och kommer alltså att återvända ännu en gång, osv. Den kommer följaktligen att om och om igen återvända, oändligt många gånger.

Om däremot E_i är ett obeständigt tillstånd, återvänder processen till E_i med en sannolikhet f_{ii} som är mindre än 1. Det finns alltså en positiv sannolikhet $1-f_{ii}$ att den aldrig kommer till E_i igen.

Exempel 7.

Låt övergångsmatrisen vara

$$P=\begin{bmatrix}1 & 0 & 0\\ 2/3 & 0 & 1/3\\ 0 & 3/4 & 1/4\end{bmatrix}.$$

Vi skall undersöka vilka tillstånd som är beständiga och vilka som är obeständiga. Vi ser att E_1 är absorberande. För ett sådant tillstånd är $f_{11}^{(1)}=1$ och alltså $f_{11}=1$; således är E_1 ett beständigt tillstånd. Låt oss därnäst undersöka E_2. Om processen startar i E_2, kan den inte vara där även vid $t=1$, ty $p_{22}=0$; alltså blir $f_{22}^{(1)}=0$. Däremot kan den vara i E_2 vid $t=2$; detta sker under förutsättning att först E_3 och sedan E_2 inträffar; sannolikheten för detta är $(1/3)(3/4)=1/4$; alltså blir $f_{22}^{(2)}=1/4$. Allmänt återvänder den till E_2 vid $t=n$, där $n\geqslant 2$, med sannolikheten $f_{22}^{(n)}=(1/3)(1/4)^{n-2}(3/4)$. Följaktligen är

$$f_{22}=\frac{1}{4}\sum_{n=2}^{\infty}(1/4)^{n-2}=\frac{1}{4}/(1-\frac{1}{4})=1/3<1.$$

Tillståndet E_2 är alltså obeständigt. På liknande sätt visas att E_3 är obeständigt. □

Vi skall nu se på två tillstånd samtidigt. Tillstånden E_i och E_j säges *kommunicera tvåsidigt* med varandra, om det med positiv sannolikhet går att komma från E_i till E_j och vice versa, dvs om $p_{ij}^{(r)}>0$ för åtminstone något $r=1, 2, \ldots$, och likaså $p_{ji}^{(r)}>0$ för åtminstone något r. Om det är möjligt att komma från E_i till E_j men inte från E_j till E_i säges E_i *kommunicera ensidigt* med E_j.

Exempel 7 (forts)

Se på övergångsmatrisen. Man kan komma från E_2 till E_3 och vice versa; alltså kommunicerar dessa tillstånd tvåsidigt med varandra. Vidare inses att både E_2 och E_3 kommunicerar ensidigt med E_1. □

Vi anger utan bevis en värdefull sats:

Sats 3. Om två tillstånd kommunicerar tvåsidigt med varandra, är antingen båda tillstånden beständiga eller båda obeständiga. □

Vi skall nu se på hela Markov-kedjan på en gång. En Markov-kedja säges vara *irreducibel* om alla dess tillstånd kommunicerar tvåsidigt med varandra. En kedja som inte är irreducibel kallas *reducibel.*

Exempel 8.

$$P=\begin{bmatrix}1/3 & 0 & 2/3\\ 1 & 0 & 0\\ 3/4 & 1/4 & 0\end{bmatrix};\qquad P=\begin{bmatrix}1/2 & 0 & 1/2 & 0\\ 0 & 1/4 & 0 & 3/4\\ 2/3 & 0 & 1/3 & 0\\ 0 & 4/5 & 0 & 1/5\end{bmatrix}.$$

Den vänstra kedjan är irreducibel. Den högra är reducibel, ty man kan inte komma från E_1 eller E_3 till E_2 eller E_4 (eller vice versa); den kan delas i två irreducibla delkedjor med tillstånden (E_1, E_3) och (E_2, E_4) och tillhörande övergångsmatriser är

$$P=\begin{bmatrix}1/2 & 1/2\\ 2/3 & 1/3\end{bmatrix};\qquad P=\begin{bmatrix}1/4 & 3/4\\ 4/5 & 1/5\end{bmatrix}.$$ □

Sats 3 visar att, i en irreducibel kedja, så är antingen alla tillstånd beständiga eller alla tillstånd obeständiga.

Om speciellt kedjan är ändlig, är det omöjligt att alla tillstånd är obeständiga. (Vi bevisar inte detta.) Härav följer i sin tur att i en *ändlig irreducibel kedja är alla tillstånd beständiga.* (Det finns ju bara två fall: alla tillstånd beständiga, alla tillstånd obeständiga, och det senare är omöjligt, eftersom kedjan är ändlig.) När man har en sådan kedja behöver man alltså inte företa någon undersökning av tillståndens karaktär, t ex via förstagångssannolikheter: man vet utan närmare undersökning att tillstånden är beständiga. I Ex 4 ser vi ett exempel på en ändlig irreducibel kedja.

Vi skall nu berätta något om periodiska och aperiodiska tillstånd. Låt d vara största gemensamma divisorn för alla heltal $n\geqslant 1$ för vilka $p_{ii}^{(n)}>0$. Om $d\geqslant 2$ säges tillståndet E_i vara *periodiskt* och ha perioden d. Om $d=1$ säges E_i vara *aperiodiskt.*

Om t ex processen kan återvända till E_i endast efter 3, 6, 9, ... steg har E_i perioden 3.

Två tillstånd, som kommunicerar tvåsidigt, har alltid samma period; likaså är de aperiodiska samtidigt. Detta medför att, i en irreducibel kedja, antingen alla tillstånd är aperiodiska eller också har de samma period.

Exempel 9.

Antag att

$$P=\begin{bmatrix}0 & 1/2 & 1/2 & 0\\ 1/2 & 0 & 0 & 1/2\\ 1/2 & 0 & 0 & 1/2\\ 0 & 1/2 & 1/2 & 0\end{bmatrix}.$$

Kedjan är irreducibel. Om den är i E_1 kan den återvända dit endast efter 2, 4, 6, ... steg och E_1 är därför ett periodiskt tillstånd med perioden $d=2$; på samma sätt är det med de övriga tillstånden. □

(b) Om asymptotiska fördelningar

Vi konstaterade i (8) att de absoluta sannolikheterna $p_i^{(n)}$ kan bestämmas med hjälp av relationen

$$p^{(n)}=p^{(0)}P^n.$$

Vi skall undersöka vad som händer med dessa sannolikheter då $n \to \infty$, dvs då processen avlägsnar sig mer och mer från startpunkten.

Exempel 10. Kedja med två tillstånd

Vi betraktar en kedja med samma övergångsmatris som i Ex 4, alltså

$$P=\begin{bmatrix} 1-\beta & \beta \\ \alpha & 1-\alpha \end{bmatrix}.$$

Man kan på olika vägar, t ex med induktion, visa att diagonalelementen i

$$P^n=\begin{bmatrix} p_{11}^{(n)} & p_{12}^{(n)} \\ p_{21}^{(n)} & p_{22}^{(n)} \end{bmatrix}$$

är

$$p_{11}^{(n)}=\frac{\alpha}{\alpha+\beta}+\frac{\beta}{\alpha+\beta}(1-\alpha-\beta)^n$$

$$p_{22}^{(n)}=\frac{\beta}{\alpha+\beta}+\frac{\alpha}{\alpha+\beta}(1-\alpha-\beta)^n.$$

Därmed känner vi hela P^n. Om $0<\alpha<1$ och $0<\beta<1$ ser man att

$$P^n \to \begin{bmatrix} \dfrac{\alpha}{\alpha+\beta} & \dfrac{\beta}{\alpha+\beta} \\ \dfrac{\alpha}{\alpha+\beta} & \dfrac{\beta}{\alpha+\beta} \end{bmatrix} \quad \text{då } n \to \infty.$$

Raderna är alltså lika och har radsumman 1. Likaså följer att då $n \to \infty$

$$p^{(n)}=p^{(0)}P^n \to \left(\frac{\alpha}{\alpha+\beta} \quad \frac{\beta}{\alpha+\beta}\right)$$

och detta sker oberoende av startvektorn $p^{(0)}$. Om man sätter igång kedjan med en godtycklig startvektor och observerar den efter lång tid, är alltså sannolikheten att den befinner sig i E_1 eller E_2 ungefär $\alpha/(\alpha+\beta)$ resp $\beta/(\alpha+\beta)$. □

Nu inför vi följande

Definition. Om

$$p^{(n)}=(p_1^{(n)}, p_2^{(n)}, \ldots) \text{ går mot } \pi=(\pi_1, \pi_2, \ldots)$$

då $n \to \infty$, $\pi_j \geq 0$, $\Sigma\pi_j=1$ och π ej beror av startvektorn, säges kedjan ha en *asymptotisk fördelning* (ibland: *jämviktsfördelning*). □

Talen $\pi_1, \pi_2, \ldots$ kallas *jämviktssannolikheter.*

Det är inte alltid enkelt att avgöra om det finns en asymptotisk fördelning, och flera kriterier har utarbetats. Vi anger ett sådant utan bevis:

Sats 4. Om man i en ändlig kedja kan finna ett $r>0$ så beskaffat att alla element i någon kolonn i matrisen P^r är positiva, existerar det en asymptotisk fördelning. □

(Speciellt gäller alltså detta om övergångsmatrisen själv har denna egenskap.) I Anm 2 i slutet av avsnittet anges ytterligare ett kriterium.

Om det finns en asymptotisk fördelning är det lätt att finna den, ty man har

Sats 5. Om det finns en asymptotisk fördelning π, bestäms den ur relationen

$$\pi=\pi P.$$ □

Gränsfördelningen π är alltså en vänsteregenvektor till matrisen P med egenvärdet 1.

Bevis: Vi bevisar satsen bara för en ändlig kedja. Vi vet att

$$p^{(0)}P^n \to \pi.$$

Men

$$p^{(0)}P^{n+1}=[p^{(0)}P^n]P.$$

Gränsövergången $n \to \infty$ ger satsen. □

Exempel 10. Kedja med två tillstånd (forts)

Om både α och β ligger mellan 0 och 1 (gränserna exkluderade) existerar enligt Sats 4 en asymptotisk fördelning. Sats 5 ger, om vi sätter $\pi=(\pi_1, \pi_2)$,

$$(\pi_1, \pi_2)=(\pi_1, \pi_2)\begin{bmatrix} 1-\beta & \beta \\ \alpha & 1-\alpha \end{bmatrix}$$

dvs utskrivet

$$\pi_1=(1-\beta)\pi_1+\alpha\pi_2$$

$$\pi_2=\beta\pi_1+(1-\alpha)\pi_2.$$

Eftersom $\pi_1+\pi_2=1$, följer härav att

$$\pi_1=\frac{\alpha}{\alpha+\beta}; \quad \pi_2=\frac{\beta}{\alpha+\beta}$$

i överensstämmelse med vad vi förut fann genom direkt räkning. □

Exempel 11. Bilmärken

I ett land finns det bara tre bilmärken E_1, E_2 och E_3. Vid köp av ny bil påverkas ägaren vid valet endast av det senast innehavda bilmärket. Övergångsmatrisen antas vara

$$P=\begin{bmatrix}0.4 & 0.4 & 0.2\\ 0 & 0.7 & 0.3\\ 0 & 0.1 & 0.9\end{bmatrix}.$$

Märkestroheten är alltså störst för E_3 ($p_{33}=0.9$), minst för E_1 ($p_{11}=0.4$). Kedjan är reducibel, ty E_1 kommunicerar endast ensidigt med E_2 och E_3 (den som köper E_2 eller E_3 köper sedan aldrig E_1).

Sats 4 visar att kedjan har en asymptotisk fördelning $\pi=(\pi_1, \pi_2, \pi_3)$. Sats 5 ger ekvationssystemet

$$\pi_1=0.4\pi_1$$

$$\pi_2=0.4\pi_1+0.7\pi_2+0.1\pi_3$$

$$\pi_3=0.2\pi_1+0.3\pi_2+0.9\pi_3$$

varav $\pi=(0, 1/4, 3/4)$.

Hur än fördelningen på olika märken ser ut från början kommer alltså i limes sannolikheten att vara 3/4 att en person har märket E_3 och 1/4 att han har E_2; märket E_1 kommer helt att försvinna. □

Anmärkning 1. Jämförelse av asymptotisk fördelning och stationär fördelning

Blanda inte ihop begreppen asymptotisk fördelning och stationär fördelning! Båda fördelningarna satisfierar ekvationen $\pi=\pi P$, men deras innebörd är olika. Läs igenom definitionerna! *Om* det finns en asymptotisk fördelning π, så är π också stationär fördelning, dvs om processen startar med startvektorn π så blir $p^{(n)}=\pi$. Däremot behöver inte en stationär fördelning vara en asymptotisk fördelning. □

Exempel 12.

En Markov-kedja har övergångsmatrisen

$$P=\begin{bmatrix}1/2 & 1/2 & 0\\ 1/2 & 1/2 & 0\\ 0 & 0 & 1\end{bmatrix}.$$

Varje sannolikhetsvektor π som satisfierar ekvationen $\pi=\pi P$ kan skrivas $\pi=(c/2, c/2, 1-c)$ där c är en godtycklig konstant, sådan att $0\leq c\leq 1$. Det finns alltså en mängd stationära fördelningar. Däremot finns det inte någon asymptotisk fördelning (ty om processen startar i E_1 eller E_2 kan den aldrig hamna i E_3, men om den startar i E_3 förblir den i E_3; något gränsvärde av $p^{(n)}$ som är oberoende av startvektorn existerar alltså inte). □

Exempel 13. Periodisk kedja

Antag att

$$P=\begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 1 & 0 & 0\end{bmatrix}.$$

Om kedjan startar t ex i E_1 blir utfallet helt bestämt: $E_1E_2E_3E_1E_2E_3$ osv. Kedjan oscillerar alltså periodiskt och har perioden $d=3$. Någon asymptotisk fördelning finns inte. Däremot finns det en (och endast en) stationär fördelning, nämligen $\pi=(1/3, 1/3, 1/3)$. □

Anmärkning 2. Ett kriterium på asymptotisk fördelning

I Sats 4 angavs ett kriterium som garanterar att en Markov-kedja har en asymptotisk fördelning. Nu skall vi ange ett mer raffinerat sådant. Jämför två rader i en matris genom att se efter om båda raderna har positiva element i minst en kolonn. Om så är fallet säger vi att detta radpar är *kolonnpositivt.* □

Sats 6. Om man kan finna ett $r>0$ sådant att alla radpar i P^r är kolonnpositiva, så har Markov-kedjan en asymptotisk fördelning. □

Exempel 14.

Antag att

$$P=\begin{bmatrix}0.5 & 0.5 & 0\\ 0.6 & 0 & 0.4\\ 0 & 0.1 & 0.9\end{bmatrix}.$$

Rad 1 och 2 har positiva element i kolonn 1, rad 1 och 3 i kolonn 2 och rad 2 och 3 i kolonn 3.

Alla radpar i P är alltså kolonnpositiva, och det finns enligt Sats 6 en asymptotisk fördelning. □

Övningar

1101 Per, Pål och Petter kastar en boll mellan sig. Per kastar med sannolikheten 0.3 till Pål och med sannolikheten 0.7 till Petter. Pål kastar med sannolikheten 0.6 till Per och med sannolikheten 0.4 till Petter, som i sin tur med lika sannolikhet kastar till sina båda vänner. Alla kast utföres oberoende av varandra. Detta kan tydligen uppfattas som en Markov-kedja. Inför lämpliga tillstånd E_1, E_2, E_3 och ställ upp övergångsmatrisen. (§ 11.3)

1102 En Markov-kedja har övergångsmatrisen

$$\begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{bmatrix}.$$

Beräkna övergångssannolikheterna av andra ordningen. (§ 11.3)

1103 Antag att Markov-kedjan i föregående exempel har startvektorn $(1/2, 1/2, 0)$. Bestäm de absoluta sannolikheterna $p_i^{(2)}$, $i = 1, 2, 3$. (§ 11.3)

1104 En Markov-kedja har övergångsmatrisen

$$\begin{bmatrix} 1/2 & 0 & 1/2 \\ 0 & 1/2 & 1/2 \\ 1/2 & 1/2 & 0 \end{bmatrix}.$$

Bestäm alla stationära fördelningar till kedjan. (§ 11.3)

1105 En Markov-kedja har övergångsmatrisen

$$\begin{bmatrix} 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \\ 1 & 0 & 0 \end{bmatrix}.$$

Undersök vilka tillstånd som är beständiga och vilka som är obeständiga. (§ 11.4)

1106 En Markov-kedja har övergångsmatrisen

$$\begin{bmatrix} 1/2 & 0 & 1/2 & 0 \\ 1 & 0 & 0 & 0 \\ 1/3 & 0 & 2/3 & 0 \\ 1/3 & 1/3 & 1/3 & 0 \end{bmatrix}.$$

Undersök vilka tillstånd som är beständiga och vilka som är obeständiga. (§ 11.4)

1107 Låt E_i vara ett obeständigt tillstånd i en Markov-kedja och f_{ii} sannolikheten att processen, vid start i E_i, återvänder till E_i. Sätt X_i lika med antalet gånger som processen återvänder. Bestäm fördelningen för X_i samt dess väntevärde. (§ 11.4)

1108 En Markov-kedja har övergångsmatrisen

$$\begin{bmatrix} 1/2 & 0 & 1/2 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 1/2 & 0 & 0 & 1/2 \\ 1/2 & 1/2 & 0 & 0 \end{bmatrix}.$$

Undersök om kedjan är irreducibel. (§ 11.4)

1109 En Markov-kedja har övergångsmatrisen

$$\begin{bmatrix} 0 & 1 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 1/2 & 1/2 \end{bmatrix}.$$

Undersök om kedjan har någon asymptotisk fördelning och bestäm i så fall denna. (§ 11.4)

1110 En Markov-kedja har övergångsmatrisen

$$\begin{bmatrix} 0 & 0 & 1 \\ 1/2 & 1/2 & 0 \\ 1/3 & 2/3 & 0 \end{bmatrix}.$$

Undersök om kedjan har någon asymptotisk fördelning och bestäm i så fall denna. (§ 11.4)

1111 En Markov-kedja har övergångsmatrisen

$$\begin{bmatrix} 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \end{bmatrix}.$$

Undersök om kedjan har någon asymptotisk fördelning och bestäm i så fall denna. (§ 11.4)

1112 En ändlig Markov-kedja har en övergångsmatris som är dubbelstokastisk. Därmed menas att inte bara alla radsummor utan även alla kolonnsummor är lika med ett. Visa att om en sådan kedja har en asymptotisk fördelning så är denna fördelning likformig. (§ 11.4)

1113 Följande modell är ett specialfall av Bernoulli-Laplace's diffusionsmodell: I Urna 1 och 2 finns det två vita resp två svarta kulor. Man låter vid varje ögonblick $t = 1, 2, \ldots$ en slumpmässigt vald kula från varje urna byta plats. Inför en lämplig Markov-kedja och undersök om den har någon asymptotisk fördelning. Ange i så fall denna. (§ 11.4)

1114 Vid en kärnreaktion passerar en neutron genom ett radioaktivt preparat. Neutronen träffar med sannolikheten p en atomkärna, som genast klyvs och därvid utsänder k nya neutroner (den »gamla» absorberas). Med sannolikheten $1-p$ passerar neutronen ut ur det radioaktiva preparatet utan att ha träffat någon kärna. Antag att de nya neutronerna var för sig uppträder som den första och oberoende av varandra. Vilket samband måste råda mellan k och p för att kärnreaktionen skall kunna utvecklas till en explosion och ej stanna av? (§ 11.5)

1115 I en förgreningsprocess har antalet avkomlingar till en individ sannolikhetsfunktionen $p_Y(j) = e^{-1}/j!$, $j = 0, 1, \ldots$. Bestäm sgf $P_Y(s)$ och extinktionssannolikheten π. (§ 11.5)