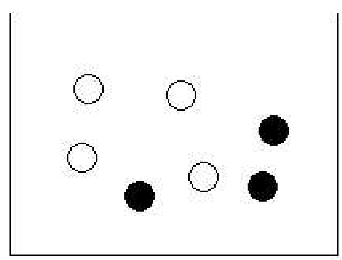
## Oppgaver fra læreboka

4.56

$X_n$ – Antallet hvite baller i urna
Trekk tilf. ball og med sans

- p → Erstatt med en hvit

- 1-p → Erstatt med en svart

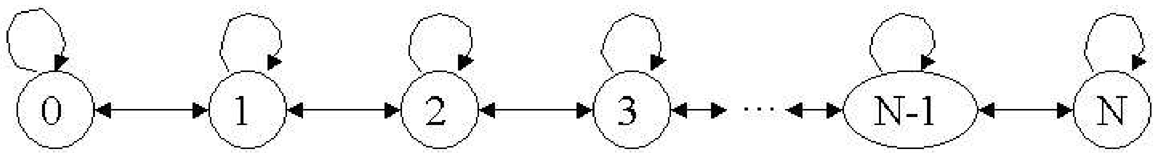


Figur 1: N baller totalt

a) Sannsynlighetsfordelingen til neste tilstand er fullt it bestemt av antallet hvite baller i nåværende tilstand:

$$P_r\{X_n = a_n \mid X_0 = a_0, ..., X_{n-1} = a_{n-1}\} = P_r\{X_{n-1} = \mid X_{n-1} = a_{n-1}\}$$

b)




- Markovkjeden er rekurrent og aperiodisk

c)

$$\begin{aligned}
P_{kk} &= \frac{k}{n}\cdot p + \frac{N-k}{N}\cdot(1-p) \\
P_{k,k-1} &= \frac{k}{n}\cdot(1-p) \\
P_{k,k+1} &= \frac{N-k}{n}\cdot(p)
\end{aligned}$$

$$P = \begin{array}{c} 0 \\ 1 \\ \vdots \\ N \end{array} \begin{bmatrix} 1-p & p & 0 & & \\ \frac{1}{N}(1-p) & \frac{p}{N}+\frac{N-1}{N}(1-p) & \frac{N-1}{N}p & & \\ & & \ddots & \ddots & \\ & & & (1-p) & p \end{bmatrix}$$

d) N = 2

$$P = \begin{bmatrix} 1-p & p & 0 \\ \frac{1}{2}\cdot(1-p) & \frac{1}{2} & \frac{1}{2}\cdot p \\ 0 & 1-p & p \end{bmatrix}$$

Andel tid i hver tilstand = Grensefordeling = Stasjonærfordeling

$$\begin{aligned}
\Pi_0 &= \Pi_0(1-p) + \Pi_1\frac{1}{2}(1-p) \\
\Pi_1 &= \Pi_0 p + \Pi_1\frac{1}{2} + \Pi_2(1-p) \\
\Pi_0 &= \Pi_1 + \Pi_2 = 1
\end{aligned}$$

Det gir

$$\begin{aligned}
\Pi_0 &= \underline{\underline{(1-p)^2}} \\
\Pi_1 &= \underline{\underline{2p(1-p)}} \\
\Pi_2 &= \underline{\underline{p^2}}
\end{aligned}$$

e) Når tiden $\to \infty$ vil hver ball med sannsynlighet P være hvit og sannsynlighet $(1-P)$ være svart, uavhengig av hverandre. Med N baller totalt gir dette binomisk fordeling:

$$\Pi_k = \underline{\underline{\binom{N}{k} P^k (1-P)^{N-k}}}$$

f) Kontrollerer at stasjonærfordelingsbetingelsene holder

$$1) \quad \sum_{k=0}^{N} \Pi_k = 1 \qquad OK$$

$$\begin{aligned}
2) \quad \sum_{i=0}^{N} \Pi_i \cdot P_{ik} \;&=\; \Pi_{k-1} \cdot P_{k-1,k} \;+\; \Pi_k \cdot P_{k,k} \;+\; \Pi_{k+1} \cdot P_{k+1,k} \\
&= \binom{N}{k-1} p^{k-1}(1-p)^{N-(k-1)} \cdot \frac{N-(k-1)}{N} \cdot p \\
&+ \binom{N}{k} p^k (1-p)^{N-k} \left( \frac{k}{N} \cdot p + \frac{N-k}{N} \cdot (1-p) \right) \\
&+ \binom{N}{k+1} p^{k+1}(1-p)^{N-(k+1)} \cdot \frac{(k-1)}{N} \cdot (1-p) \\
&= .....
\end{aligned}$$

$$= \left[ \frac{k}{N} \cdot (1-p) + \frac{k}{N} \cdot p + \frac{N-k}{N} \cdot (1-p) + \frac{N-k}{N} \cdot p \right] \binom{N}{k} p^k (1-p)^{N-k}$$

$$= \underline{\underline{\Pi_k}}$$

g)

$$T = min \, \{ \, n : X_n = N \, \} \qquad P = 1$$

Definerer

$$\mu_i = E \, [ \, T \mid X_0 = i \, ]$$

Har at

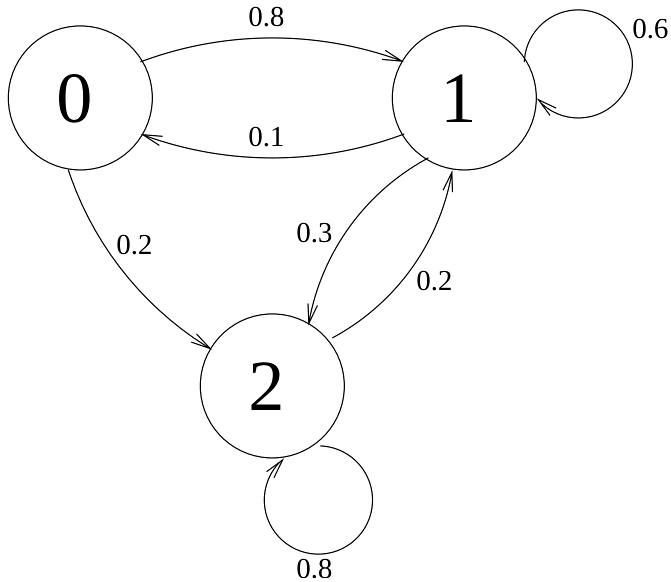
$$T = \sum_{j=i}^{N-1} T_j$$

der $T_j$ = antall steg for å gå fra j til j+1 $\backsim$ geometr. $\left(\frac{j}{N}\right)$ $\quad j \geqslant 1 \;,\; T_0 = 1$
Dermed blir

$$\begin{aligned}
E[T \mid X_0 = i] \;&=\; \sum_{j=i}^{N-1} E[T_i] \\
&=\; \sum_{j=i}^{N-1} \frac{N}{j} = N \sum_{j=i}^{N-1} \frac{1}{j} \qquad i \geqslant 1 \\
E[T \mid X_0 = 0] \;&=\; \underline{\underline{1 + N \sum_{j=i}^{N-1} \frac{1}{j}}}
\end{aligned}$$

Eks. Mai '02, oppg 1

a)




Figur 2: Nodeskjema over markovkjeden

$$
\begin{aligned}
P\{X_2 = 2 | X_0 = 2\} &= P_{21} \cdot P_{12} + P_{22} \cdot P_{22} \\
&= 0.2 \cdot 0.3 + 0.8 \cdot 0.8 \\
&= \underline{\underline{0.7}}
\end{aligned}
$$

$$
\begin{aligned}
P\{X_1 \neq 2, X_2 \neq 2, X_3 \neq 2, X_4 = 2 | X_0 = 2\} &= P_{21}\big[P_{11} \cdot P_{11} \cdot P_{12} + P_{11} \cdot P_{10} \cdot P_{02} + P_{10} \cdot P_{01} \cdot P_{12}\big] \\
&= 0.2\big[0.6 \cdot 0.6 \cdot 0.3 + 0.6 \cdot 0.1 \cdot 0.2 + 0.1 \cdot 0.8 \cdot 0.3\big] \\
&= 0.8\big[0.108 + 0.012 + 0.024\big] \\
&= \underline{\underline{0.0288}}
\end{aligned}
$$

b)

Dersom vi ser på overgangsmatrisen eller node-skjemaet, ser vi at alle tilstandene kommuniserer. Det vil altså si at alle tilstandene er i samme ekvivalensklasse ($0 \to 1 \to 2 \to 0$).

Siden vi har endelig mange tilstander, må minst en tilstand være rekurrent. Videre, siden vi kun har en ekvivalensklasse, må alle tilstandene være rekurrente.

Siden $P_{11} > 0$ må $d(1) = 1$. Vi har kun en ekvivalensklasse, så då må $d(0) = d(1) = d(2) = 1$.

Markovkjeden har en grensefordeling siden den er irredusibel (kun en ekvivalensklasse), positiv rekurrent (rekurrent og endelig antall tilstander) og aperiodisk.

$$
\begin{aligned}
\pi_0 &= 0.1\pi_1 && (1)\\
\pi_1 &= 0.8\pi_0 + 0.6\pi_1 + 0.2\pi_2 && (2)\\
\pi_2 &= 0.2\pi_0 + 0.3\pi_1 + 0.8\pi_2 && (3)\\
\pi_0 &+ \pi_1 + \pi_2 = 1 && (4)
\end{aligned}
$$

Løser vi (1) - (4), får vi

$$\underline{\underline{\pi_0 = 0.037, \pi_1 = 0.370, \pi_2 = 0.593}}$$

c)

La $T = min\{n > 0 : X_n = 0\}$

For å finne tykkelsen av en sone med ren/skifrig sandsten, må vi finne $E[T - 1|X_0 = 0]$. Dette fordi $X_0 = 0$ definerer slutten på en skifersone, mens $X_T = 0$ definerer begynnelsen på en ny.

Vi har at $E[T - 1|X_0 = 0] = E[T|X_0 = 0] - 1$, dermed trenger vi å regne ut $E[T|X_0 = 0]$.

$$
\begin{aligned}
E[T|X_0 = 0] &= E[T|X_0 = 0, X_1 = 1] \cdot P[X_1 = 1|X_0 = 0]\\
&+ E[T|X_0 = 0, X_1 = 2] \cdot P[X_1 = 2|X_0 = 0]\\
&= 0.8 \cdot (E[T|X_0 = 1] + 1) + 0.2 \cdot (E[T|X_0 = 2] + 1)\\
&= 0.8 \cdot \nu_1 + 0.2 \cdot \nu_2 + 1,
\end{aligned}
$$

der $\nu_i = E[T|X_0 = i]$. Vi har altså at

$$\nu_0 = 0.8 \cdot \nu_1 + 0.2 \cdot \nu_2 + 1. \qquad (5)$$

Tilsvarende utregninger for $\nu_1$ og $\nu_2$ gir oss følgende ligninger

$$
\begin{aligned}
\nu_1 &= 0.1 \cdot 1 + 0.6(1 + \nu_1) + 0.3(1 + \nu_2) = 1 + 0.6 \cdot \nu_1 + 0.3 \cdot \nu_2 && (6)\\
\nu_2 &= 0.2(1 + \nu_1) + 0.8(1 + \nu_2) = 1 + 0.2 \cdot \nu_1 + 0.8 \cdot \nu_2 && (7)
\end{aligned}
$$

Løser vi (**??**) og (**??**) får vi at

$$\nu_2 = 30$$

og

$$\nu_1 = 25.$$

Ved å sette $\nu_1$ og $\nu_2$ inn i (**??**) får vi

$$\nu_0 = E[T|X_0 = 0] = 1 + 0.8 \cdot 25 + 0.2 \cdot 30 = 27,$$

som betyr at tykkelsen er

$$\underline{\underline{E[T-1|X_0 = 0] = 27 - 1 = 26}}$$

d)

Algoritme:

- Sett initialverdi $X_0 = i_o$
- For $i = 1, 2, 3, ...$
  1. Trekk $u \sim U[0,1]$
  2. Dersom $u < P_{X_{i-1},0}$ sett $X_i = 0$
  3. Dersom $P_{X_{i-1},0} \leq u < P_{X_{i-1},0} + P_{X_{i-1},1}$ sett $X_i = 1$
  4. Dersom $u \geq P_{X_{i-1},0} + P_{X_{i-1},1}$ sett $X_i = 2$

Fra realisasjonen $x_0, x_1, x_2, \ldots$ finner man først tykkelsen til hvert av de simulerte lagene mellom skifer. Det vil si å finne $T_1, T_2, T_3, \ldots$. En realisasjon kan se f.eks. slik ut:

$$1\ 2\ 0\ \underbrace{1\ 1\ 2\ 2\ 0}_{T_1}\ \underbrace{2\ 1\ 0}_{T_2}\ \underbrace{2\ 2\ 1\ 0}_{T_3}\ \underbrace{1\ 2\ 1\ 0}_{T_4}\ \underbrace{2\ 1\ 2\ 2\ 0}_{T_5}\ \cdots$$

Et estimat av $P\{T - 1 \geq d\}$ vil då være gitt som:

$$\underline{\underline{P\{\widehat{T - 1 \geq} d\} = \frac{1}{N}\sum_{i=1}^{N} I\{T_i - 1 \geq d\},}} \tag{8}$$

der $N$ er antall simulerte lag, og $I\{\cdot\}$ er indikatorfunksjonen.

En alternativ algoritme kan vere å starte mange simuleringer på $X_0 = 0$ og simulere kun fram til neste skifer (altså neste $X_i = 0$). Ved å ta vare på de $T_i$-ene man da får, kan man regne ut estimatet (**??**).

$\boxed{\text{Eks. Aug '01, oppg 1}}$

**a)** Sannsynsfordelinga til antal kuler i urna i neste tilstand er fullt ut bestemt av antal kuler i urna i noverande tilstand. Altså:

$$P\{X_n = a_n | X_0 = a_0, X_1 = a_1, \ldots, X_{n-1} = a_{n-1}\} = P\{X_n = a_n | X_{n-1} = a_{n-1}\}$$

**b)** Ei markovkjede er irredusibel dersom den kun har ei ekvivalensklasse, det vil seie at alle tilstandane kommuniserer med kvarandre. Det er tilfelle for markovkjeda i denne oppgåva.

Dersom markovkjeda er i tilstand $n$ ved tidspunkt $t_0$, kan den kun komme tilbake til tilstanden $n$ ved tidspunkta $t_0 + 2$, $t_0 + 4$, $t_0 + 6$,... Det betyr at alle tilstandane i markovkjeda har periode 2. Den er altså ikkje aperiodisk.

**c)** Sidan alle tilstandane har periode 2, er det kun tilstandane 1 og 3 som er tilgjengelige på 3 steg frå tilstand 0. Vi får dermed at:

$$P_{0j}^3 = \begin{cases} 1 \cdot \frac{1}{M} \cdot +1 \cdot \frac{M-1}{M} \cdot \frac{2}{M} = \frac{3M-2}{M^2} & \text{for } j = 1 \\ 1 \cdot \frac{M-1}{M} \frac{M-2}{M} = \frac{M^2-3M+2}{M^2} & \text{for } j = 3 \\ 0 & \text{ellers} \end{cases}$$

**d)** Venstre side:

$$\begin{aligned} \binom{M}{i} P_{i,i+1} &= \tfrac{M!}{i!(M-i)!} \cdot (1 - i/M) = \\ &= \tfrac{M!}{i!(M-i)!} \cdot \tfrac{M-i}{M} = \tfrac{(M-1)!}{i!(M-i-1)!} \end{aligned}$$

Høyre side:

$$\binom{M}{i+1} P_{i+1,i} = \frac{M!}{(i+1)!(M-i-1)!} \cdot \frac{i+1}{M} = \frac{(M-1)!}{i!(M-i-1)!}$$

Vi ser at de to sidene er like. For at en markovkjede skal være tidsreversibel, må likheten

$$\pi_i P_{i\,j} = \pi_j P_{j\,i} \text{ for alle } i, j$$

gjelde. La $(x_0, x_1, \ldots, x_M)$ være positive slik at

$$\pi_i = \frac{x_i}{\sum_{k=0}^M x_k} \text{ for alle } i$$

Da vil kravet om reversibilitet bli redusert til

$$\begin{gathered} \pi_i P_{i\,j} = \pi_j P_{j\,i} \text{ for alle } i, j \\ \Updownarrow \\ \frac{x_i}{\sum_{k=0}^M x_k} P_{i\,j} = \frac{x_j}{\sum_{k=0}^M x_k} P_{j\,i} \\ \Updownarrow \\ x_i P_{i\,j} = x_j P_{j\,i} \text{ for alle } i, j \end{gathered}$$

I vårt tilfelle er det nok at

$$x_i P_{i,i+1} = x_{i+1} P_{i+1,i} \text{ for alle } i \tag{9}$$

siden $P_{ij} = 0$ ellers. Vi ser at $x_i = \binom{M}{i}$ oppfyller ligning (**??**).

**e)** P.g.a. tidsreversibilitet må vi ha

$$\pi_i = \frac{x_i}{\sum_{k=0}^{M} x_k} = \frac{\binom{M}{i}}{\sum_{k=0}^{M}\binom{M}{k}} = \binom{M}{i} 2^{-M}$$

dvs binomisk$(M, 1/2)$ fordeling. Intuitiv tolkning: Etter et stort antall trekninger vil hver kule med sannsynlighet 1/2 være i en av de to urnene, uavhengig av hverandre.

**f)** Iflg. vink ser vi først på

$$\begin{aligned}
m_i^{n+1} &= E(X_{n+1}|X_0 = i) \\
&= \sum_{j=0}^{M} E(X_{n+1}|X_0 = i, X_n = j)P(X_n = j|X_0 = i) \\
&= \sum_{j=0}^{M} E(X_{n+1}|X_n = j)P(X_n = j|X_0 = i) \\
&= \sum_{j=0}^{M} [(j-1)\frac{j}{M} + (j+1)(1 - \frac{j}{M})]P_{ij}^n \\
&= \sum_{j=0}^{M} [1 + (1 - \frac{2}{M})j]P_{ij}^n \\
&= 1 + (1 - \frac{2}{M})E(X_n|X_0 = i) \\
&= 1 + (1 - \frac{2}{M})m_i^n
\end{aligned}$$

Løser ved hjelp av rekursjon. For $n = 0$:

$$m_i^0 = \frac{M}{2} + (1 - \frac{2}{M})^0(i - \frac{M}{2}) = \frac{M}{2} + (i - \frac{M}{2}) = i, \tag{10}$$

som stemmer med det vi forventer. Antar nå at formelen gjelder for $n$. Skal vise at den også gjelder for $n + 1$.

$$\begin{aligned}
m_i^{n+1} &= 1 + (1 - \frac{2}{M})m_i^n \\
&= 1 + (1 - \frac{2}{M})\frac{M}{2} + (1 - \frac{2}{M})^n(i - \frac{M}{2}) \\
&= 1 + (\frac{M}{2} - 1) + (1 - \frac{2}{M})^{n+1}(i - \frac{M}{2}) \\
&= \frac{M}{2} + (1 - \frac{2}{M})^{n+1}(i - \frac{M}{2})
\end{aligned}$$