Introduction to stochastic processes

Eli Gurarie

StatR 301 - Lecture 7
University of Washington - Seattle

May 23, 2013

PROFESSIONAL & CONTINUING EDUCATION
UNIVERSITY of WASHINGTON

A Stochastic Processes Is:

- Any process in which outcomes in some variable (usually time, sometimes space, sometimes something else) are uncertain and best modelled probabilistically.
- **stochastic** is to **deterministic** as **random variable** is to **number**
- Stochastic processes can be continuous or discrete in time (index) and/or state.
  - "time series" (e.g. ARMA models) are usually discrete time-continuous state.
  - "Markov chains" (today's topic) are usually discrete state.

Examples include:

Just about everything.

A Stochastic Processes Is:

- Any process in which outcomes in some variable (usually time, sometimes space, sometimes something else) are uncertain and best modelled probabilistically.
- **stochastic** is to **deterministic** as **random variable** is to **number**
- Stochastic processes can be continuous or discrete in time (index) and/or state.
  - "time series" (e.g. ARMA models) are usually discrete time-continuous state.
  - "Markov chains" (today's topic) are usually discrete state.

Examples include:

Just about everything.

Just about everything includes:

- Weather/Climate
- Population biology
  - Birth/death/reproduction/mortality
  - Migrations and movements
- Evolution
  - Population genetics (Mutation/Selection/Drift)
  - Gene sequences
- Epidemiology
  - Disease spread within a population (SIR models)
  - Disease spread within an organism
  - Development of resistance
- Tools for estimating models
  - MCMC (Markov Chain Monte Carlo)
  - Simulated annealing
- and much, much more

Just about anything ALSO includes:

Your life!

-
-
-

Just about anything ALSO includes:

Your life!

- You drop out of school to make lots of money in the stock market
-
-

Just about anything ALSO includes:

Your life!

- You drop out of school to make lots of money in the stock market
- You lose all your money gambling (Bernoulli and Bernoulli 1713)
-

Just about anything ALSO includes:

Your life!

- You drop out of school to make lots of money in the stock market
- You lose all your money gambling (Bernoulli and Bernoulli 1713)
- You eventually do or don't find your way home from some unknown establishment you've chosed to drink your sorrows away in (Pearson 1905)

Just about anything ALSO includes:

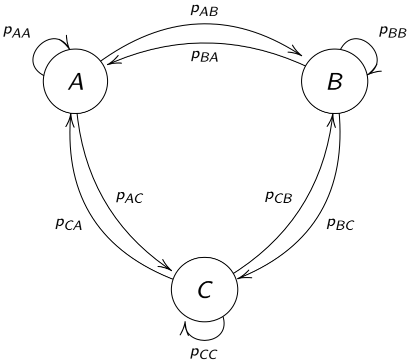
Your life!

- You drop out of school to make lots of money in the stock market
- You lose all your money gambling (Bernoulli and Bernoulli 1713)
- You eventually do or don't find your way home from some unknown establishment you've chosed to drink your sorrows away in (Pearson 1905)

All classic problems in stochastic processes!

Discrete state transitions

Consider $\mathbf{X} = \{X_1, X_2, X_3, ..., X_n\}$ is in some discrete state space $\mathcal{E}$ (here: A, B, C) with fixed probabilities of transitioning from one state to another:



Sample sequence: $\mathbf{X} = CCCBBCACCBABCBA...$
This object is called a **Markov** chain.

Some definitions

$\mathbf{X}_n$ has the **Markov Property** if:

$$\Pr\{X_n = x_n | X_1 = x_1, \cdots, X_{n-1} = x_{n-1}\} = \Pr\{X_n = x_n | X_{n-1} = x_{n-1}\}$$

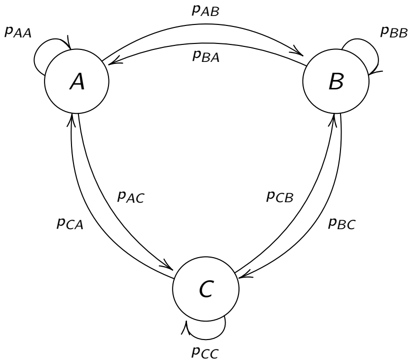
for all n in $x_1, \cdots, x_n$.

In other words, any system whose future depends *only* on the present and not on the past has the *Markov Property* and any $\mathbf{X}_n$ that has the Markov property is a **Markov Chain**.

The $p_{ij}(t)$'s of a Markov chain are transition probabilities. If $p_{ij}(t)$'s are time invariant, $(p_{ij}(t) = p_{ij})$, the chain is called **time homogeneous** or is said to have **stationary transition probabilities**.

Discrete state transitions

We express this process in terms of a **Probability Transition** matrix:



$$\mathbf{M} = \begin{array}{c|ccc|}
{}_{from:}\backslash^{to:} & \text{A} & \text{B} & \text{C} \\
\hline
\text{A} & p_{AA} & p_{AB} & p_{AC} \\
\text{B} & p_{BA} & p_{BB} & p_{BC} \\
\text{C} & p_{CA} & p_{BC} & p_{CC}
\end{array}$$

Such that:

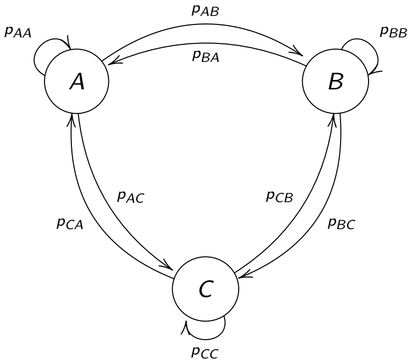
$$M_{ij} = \Pr\left(X_{t+1} = j | X_t = i\right) = p_{ij} \tag{1}$$

Note that:

$$\sum_{j=1}^{n} p_{ij} = 1 \ldots \text{BUT} \ldots \sum_{i=1}^{n} p_{ij} \neq 1 \tag{2}$$

Discrete state transitions

We express this process in terms of a **Probability Transition** matrix:



$$\mathbf{M} = \begin{array}{c|ccc|} {}_{\textit{from:}}\backslash^{\textit{to:}} & \text{A} & \text{B} & \text{C} \\ \hline \text{A} & p_{AA} & p_{AB} & p_{AC} \\ \text{B} & p_{BA} & p_{BB} & p_{BC} \\ \text{C} & p_{CA} & p_{BC} & p_{CC} \end{array}$$

Such that:

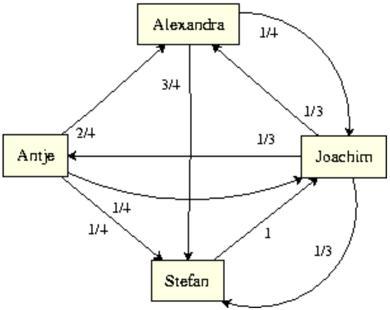
$$\Pr(X_{t+1} = j) = \sum_{i=1}^{N} M_{ij} \Pr(X_t = i) \tag{3}$$

Which can be conveniently rewritten in matrix notation as:

$$\pi_{t+1} = \mathbf{M} \times \pi_t \tag{4}$$

Where π_t is the distribution of the system over all states at time t.

Example 1: Children play catch[1]



$$\mathbf{M} = \begin{array}{c|cccc|} & An & Al & Jo & St \\ \hline An & 0 & 3/4 & 0 & 1/4 \\ Al & 0 & 0 & 1/4 & 3/4 \\ Jo & 1/3 & 1/3 & 0 & 1/3 \\ St & 0 & 0 & 1 & 0 \end{array}$$

Let's give the ball to Antje, and see what happens:

An - Al - St - Jo - St - Jo - Al - St - Jo - St - Jo - Al - Jo - St - Jo - St - Jo - Al - Jo - An - Al - St - Jo - Al - Jo - Al - St - Jo - Al - St - Jo - An - St - Jo - St - Jo - St - Jo - An - St - Jo - An - Al - Jo - St - Jo - An - Al - St - Jo - St - Jo - An - Al - Jo - Al - St - Jo - Al - St - Jo - An - Al - St - Jo - Al - St - Jo - etc ...

This is called a **realization** of a stochastic process.

[1] http://www.leda-tutorial.org/en/unofficial/Pictures/MarkovChain.png

Simulating a Markov chain

```
M <- rbind(c(0, 3/4, 0, 1/4), c(0, 0, 1/4, 3/4), c(1/3, 1/3, 0, 1/3), c(0, 0,
    1, 0))
rowSums(M)

## [1] 1 1 1 1
```

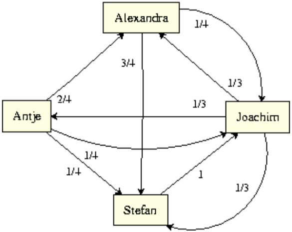
```
State <- 1:4
X <- 1
Xs <- X
for (i in 1:30) {
    X <- sample(State, 1, prob = M[X, ])
    Xs <- c(Xs, X)
}
paste(LETTERS[Xs], collapse = "-")

## [1] "A-B-C-A-B-D-C-D-C-B-D-C-A-B-D-C-D-C-A-B-D-C-A-D-C-D-C-A-B-D-C"
```

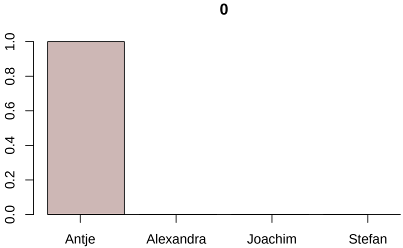
Let's make a subtle change

```
M[4, ] <- c(0, 0, 0, 4)
X <- 1
Xs <- X
for (i in 1:30) {
    X <- sample(State, 1, prob = M[X, ])
    Xs <- c(Xs, X)
}
paste(LETTERS[Xs], collapse = "-")

## [1] "A-B-D-D-D-D-D-D-D-D-D-D-D-D-D-D-D-D-D-D-D-D-D-D-D-D-D-D-D-D-D"
```

What happened?

Consider the process probabilistically

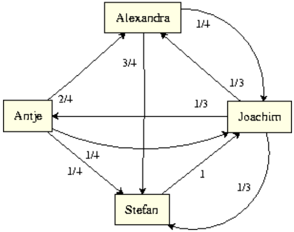


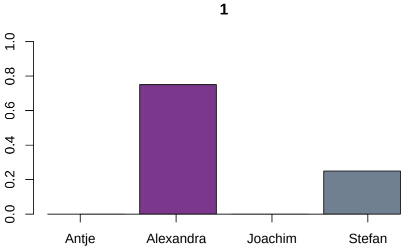


Starting with Antje again:

$$\pi_0 = (1, 0, 0, 0)$$

Consider the process probabilistically

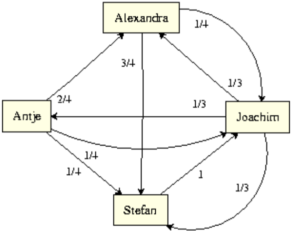


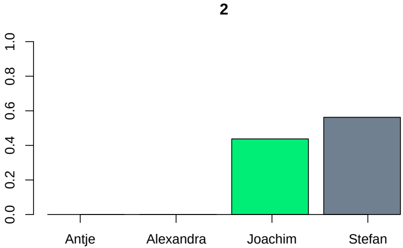


Starting with Antje again:

$$\pi_0 = (1, 0, 0, 0)$$
$$\pi_1 = (0, 0.75, 0, 0.25)$$

Consider the process probabilistically

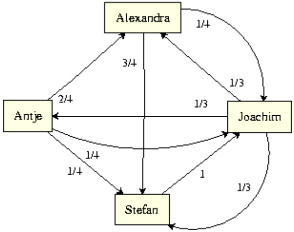


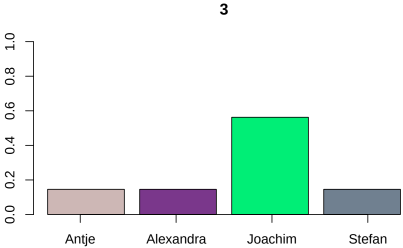


Starting with Antje again:

$$\begin{aligned} \pi_0 &= (1, 0, 0, 0) \\ \pi_1 &= (0, 0.75, 0, 0.25) \\ \pi_2 &= (0, 0, 0.438, 0.562) \end{aligned}$$

Consider the process probabilistically

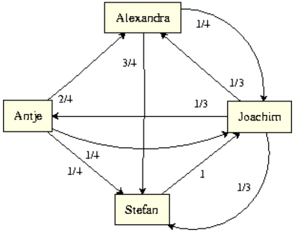


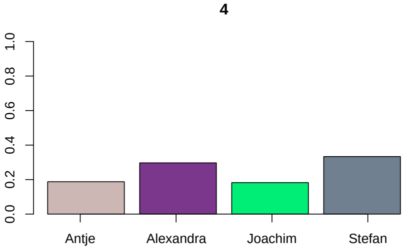


Starting with Antje again:

$$
\begin{aligned}
\pi_0 &= (1, 0, 0, 0) \\
\pi_1 &= (0, 0.75, 0, 0.25) \\
\pi_2 &= (0, 0, 0.438, 0.562) \\
\pi_3 &= (0.146, 0.146, 0.562, 0.146)
\end{aligned}
$$

Consider the process probabilistically

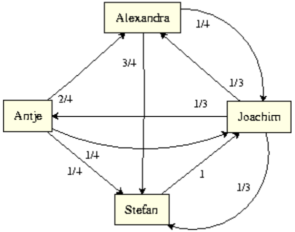


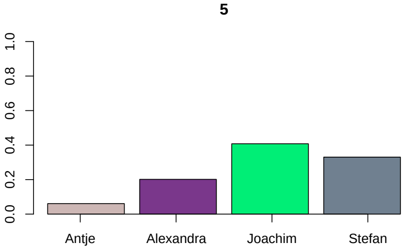


Starting with Antje again:

$$
\begin{aligned}
\pi_0 &= (1, 0, 0, 0) \\
\pi_1 &= (0, 0.75, 0, 0.25) \\
\pi_2 &= (0, 0, 0.438, 0.562) \\
\pi_3 &= (0.146, 0.146, 0.562, 0.146) \\
\pi_4 &= (0.061, 0.201, 0.408, 0.33)
\end{aligned}
$$

Consider the process probabilistically

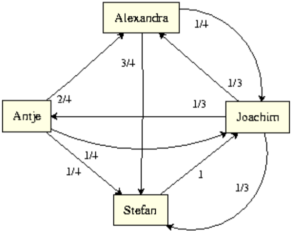




Starting with Antje again:

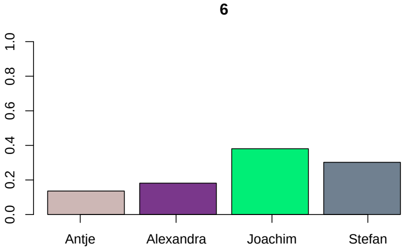
$$
\begin{aligned}
\pi_0 &= (1, 0, 0, 0) \\
\pi_1 &= (0, 0.75, 0, 0.25) \\
\pi_2 &= (0, 0, 0.438, 0.562) \\
\pi_3 &= (0.146, 0.146, 0.562, 0.146) \\
\pi_4 &= (0.061, 0.201, 0.408, 0.33) \\
\pi_5 &= (0.136, 0.181, 0.381, 0.302)
\end{aligned}
$$

Consider the process probabilistically

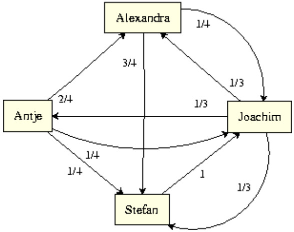


Starting with Antje again:

$$
\begin{aligned}
\pi_0 &= (1, 0, 0, 0) \\
\pi_1 &= (0, 0.75, 0, 0.25) \\
\pi_2 &= (0, 0, 0.438, 0.562) \\
\pi_3 &= (0.146, 0.146, 0.562, 0.146) \\
\pi_4 &= (0.061, 0.201, 0.408, 0.33) \\
\pi_5 &= (0.136, 0.181, 0.381, 0.302) \\
\pi_6 &= (0.127, 0.229, 0.347, 0.297)
\end{aligned}
$$

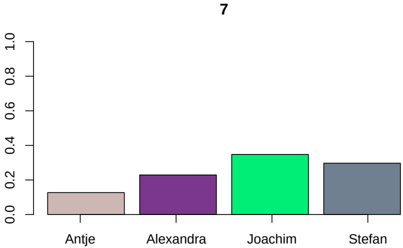


Consider the process probabilistically

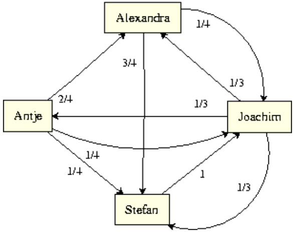


Starting with Antje again:

$$
\begin{aligned}
\pi_0 &= (1, 0, 0, 0) \\
\pi_1 &= (0, 0.75, 0, 0.25) \\
\pi_2 &= (0, 0, 0.438, 0.562) \\
\pi_3 &= (0.146, 0.146, 0.562, 0.146) \\
\pi_4 &= (0.061, 0.201, 0.408, 0.33) \\
\pi_5 &= (0.136, 0.181, 0.381, 0.302) \\
\pi_6 &= (0.127, 0.229, 0.347, 0.297) \\
\pi_7 &= (0.116, 0.211, 0.354, 0.319)
\end{aligned}
$$

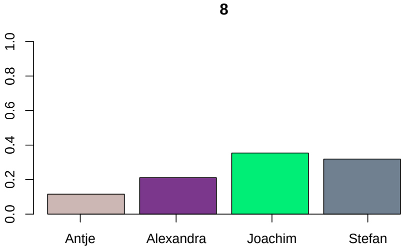


Consider the process probabilistically

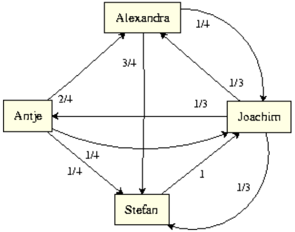


Starting with Antje again:

$$
\begin{aligned}
\pi_0 &= (1, 0, 0, 0) \\
\pi_1 &= (0, 0.75, 0, 0.25) \\
\pi_2 &= (0, 0, 0.438, 0.562) \\
\pi_3 &= (0.146, 0.146, 0.562, 0.146) \\
\pi_4 &= (0.061, 0.201, 0.408, 0.33) \\
\pi_5 &= (0.136, 0.181, 0.381, 0.302) \\
\pi_6 &= (0.127, 0.229, 0.347, 0.297) \\
\pi_7 &= (0.116, 0.211, 0.354, 0.319) \\
\pi_8 &= (0.118, 0.205, 0.372, 0.305)
\end{aligned}
$$

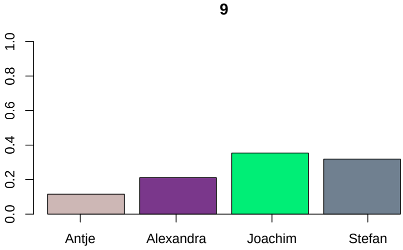


Consider the process probabilistically

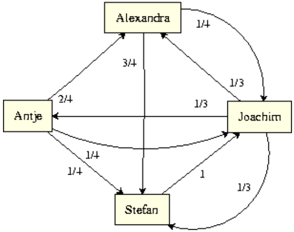


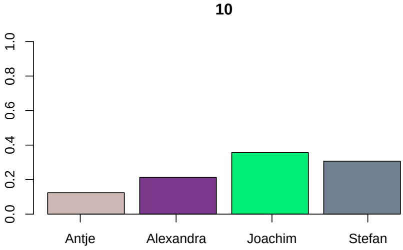
Starting with Antje again:

$$
\begin{aligned}
\pi_0 &= (1, 0, 0, 0) \\
\pi_1 &= (0, 0.75, 0, 0.25) \\
\pi_2 &= (0, 0, 0.438, 0.562) \\
\pi_3 &= (0.146, 0.146, 0.562, 0.146) \\
\pi_4 &= (0.061, 0.201, 0.408, 0.33) \\
\pi_5 &= (0.136, 0.181, 0.381, 0.302) \\
\pi_6 &= (0.127, 0.229, 0.347, 0.297) \\
\pi_7 &= (0.116, 0.211, 0.354, 0.319) \\
\pi_8 &= (0.118, 0.205, 0.372, 0.305) \\
\pi_9 &= (0.124, 0.212, 0.356, 0.307)
\end{aligned}
$$



Consider the process probabilistically

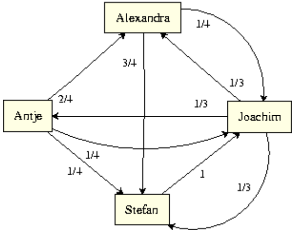


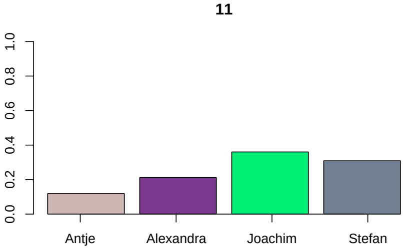


Starting with Antje again:

$$\begin{aligned}
\pi_0 &= (1, 0, 0, 0)\\
\pi_1 &= (0, 0.75, 0, 0.25)\\
\pi_2 &= (0, 0, 0.438, 0.562)\\
\pi_3 &= (0.146, 0.146, 0.562, 0.146)\\
\pi_4 &= (0.061, 0.201, 0.408, 0.33)\\
\pi_5 &= (0.136, 0.181, 0.381, 0.302)\\
\pi_6 &= (0.127, 0.229, 0.347, 0.297)\\
\pi_7 &= (0.116, 0.211, 0.354, 0.319)\\
\pi_8 &= (0.118, 0.205, 0.372, 0.305)\\
\pi_9 &= (0.124, 0.212, 0.356, 0.307)\\
\pi_{10} &= (0.119, 0.212, 0.36, 0.309)
\end{aligned}$$

Consider the process probabilistically

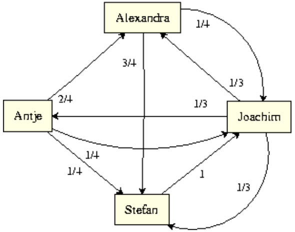


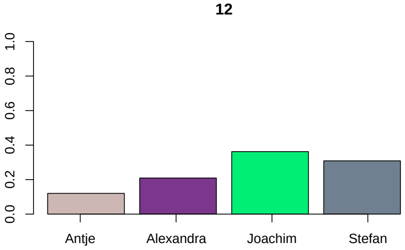


Starting with Antje again:

$$
\begin{aligned}
\pi_0 &= (1, 0, 0, 0) \\
\pi_1 &= (0, 0.75, 0, 0.25) \\
\pi_2 &= (0, 0, 0.438, 0.562) \\
\pi_3 &= (0.146, 0.146, 0.562, 0.146) \\
\pi_4 &= (0.061, 0.201, 0.408, 0.33) \\
\pi_5 &= (0.136, 0.181, 0.381, 0.302) \\
\pi_6 &= (0.127, 0.229, 0.347, 0.297) \\
\pi_7 &= (0.116, 0.211, 0.354, 0.319) \\
\pi_8 &= (0.118, 0.205, 0.372, 0.305) \\
\pi_9 &= (0.124, 0.212, 0.356, 0.307) \\
\pi_{10} &= (0.119, 0.212, 0.36, 0.309) \\
\pi_{11} &= (0.12, 0.209, 0.362, 0.309)
\end{aligned}
$$

Consider the process probabilistically

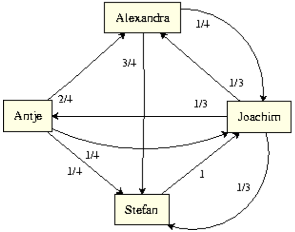


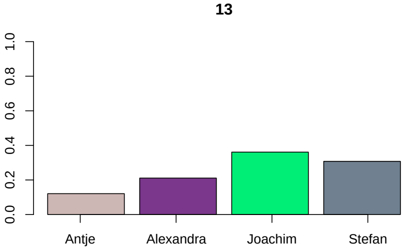


Starting with Antje again:

$$
\begin{aligned}
\pi_0 &= (1, 0, 0, 0)\\
\pi_1 &= (0, 0.75, 0, 0.25)\\
\pi_2 &= (0, 0, 0.438, 0.562)\\
\pi_3 &= (0.146, 0.146, 0.562, 0.146)\\
\pi_4 &= (0.061, 0.201, 0.408, 0.33)\\
\pi_5 &= (0.136, 0.181, 0.381, 0.302)\\
\pi_6 &= (0.127, 0.229, 0.347, 0.297)\\
\pi_7 &= (0.116, 0.211, 0.354, 0.319)\\
\pi_8 &= (0.118, 0.205, 0.372, 0.305)\\
\pi_9 &= (0.124, 0.212, 0.356, 0.307)\\
\pi_{10} &= (0.119, 0.212, 0.36, 0.309)\\
\pi_{11} &= (0.12, 0.209, 0.362, 0.309)\\
\pi_{12} &= (0.121, 0.211, 0.361, 0.308)
\end{aligned}
$$

Consider the process probabilistically

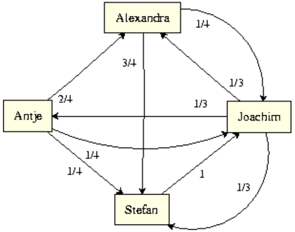


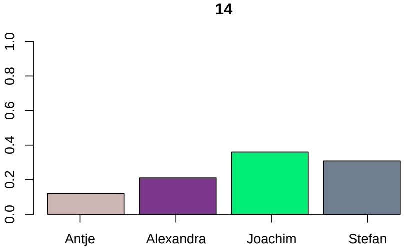


Starting with Antje again:

$$
\begin{aligned}
\pi_0 &= (1, 0, 0, 0)\\
\pi_1 &= (0, 0.75, 0, 0.25)\\
\pi_2 &= (0, 0, 0.438, 0.562)\\
\pi_3 &= (0.146, 0.146, 0.562, 0.146)\\
\pi_4 &= (0.061, 0.201, 0.408, 0.33)\\
\pi_5 &= (0.136, 0.181, 0.381, 0.302)\\
\pi_6 &= (0.127, 0.229, 0.347, 0.297)\\
\pi_7 &= (0.116, 0.211, 0.354, 0.319)\\
\pi_8 &= (0.118, 0.205, 0.372, 0.305)\\
\pi_9 &= (0.124, 0.212, 0.356, 0.307)\\
\pi_{10} &= (0.119, 0.212, 0.36, 0.309)\\
\pi_{11} &= (0.12, 0.209, 0.362, 0.309)\\
\pi_{12} &= (0.121, 0.211, 0.361, 0.308)\\
\pi_{13} &= (0.12, 0.211, 0.36, 0.309)
\end{aligned}
$$

Consider the process probabilistically

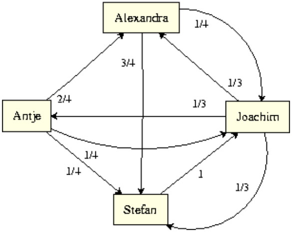


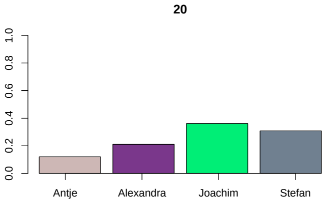


Starting with Antje again:

$$
\begin{aligned}
\pi_0 &= (1, 0, 0, 0) \\
\pi_1 &= (0, 0.75, 0, 0.25) \\
\pi_2 &= (0, 0, 0.438, 0.562) \\
\pi_3 &= (0.146, 0.146, 0.562, 0.146) \\
\pi_4 &= (0.061, 0.201, 0.408, 0.33) \\
\pi_5 &= (0.136, 0.181, 0.381, 0.302) \\
\pi_6 &= (0.127, 0.229, 0.347, 0.297) \\
\pi_7 &= (0.116, 0.211, 0.354, 0.319) \\
\pi_8 &= (0.118, 0.205, 0.372, 0.305) \\
\pi_9 &= (0.124, 0.212, 0.356, 0.307) \\
\pi_{10} &= (0.119, 0.212, 0.36, 0.309) \\
\pi_{11} &= (0.12, 0.209, 0.362, 0.309) \\
\pi_{12} &= (0.121, 0.211, 0.361, 0.308) \\
\pi_{13} &= (0.12, 0.211, 0.36, 0.309) \\
\pi_{14} &= (0.12, 0.21, 0.361, 0.308)
\end{aligned}
$$

The stationary state





The state: $\pi^* = (0.12, 0.21, 0.361, 0.308)$ is referred to as **stationary**. Note:

1. The name can be a bit confusing: the ball is not stationary, it is always moving around.
2. The state can be solved for mathematically:

$$\pi^* = \pi^* \mathbf{M} \tag{5}$$

   This is a straighforward linear algebra problem, and is usually easy to obtain

3. All states have a value between 0 and 1 and have finite probablity of being revisited forever and ever until the children's arms fall off. Such states are termed **recurrent**, **persistent** or **ergodic**.

Finding stationary state probabilities

Via matrix multiplication:

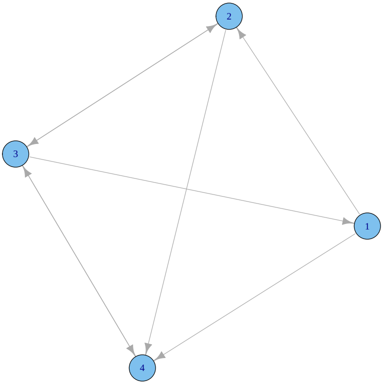
```
powermat = function(M, t) {
    Mt = M
    if (t > 1)
        for (i in 2:t) Mt = Mt %*% M
    Mt
}
powermat(M, 100)

##        [,1]   [,2]   [,3]   [,4]
## [1,] 0.1203 0.2105 0.3609 0.3083
## [2,] 0.1203 0.2105 0.3609 0.3083
## [3,] 0.1203 0.2105 0.3609 0.3083
## [4,] 0.1203 0.2105 0.3609 0.3083
```

Via matrix algebra:

```
n <- nrow(M)
a <- rbind((t(M) - diag(n))[-n, ], rep(1, n))
b <- c(rep(0, n - 1), 1)
solve(a, b)

## [1] 0.1203 0.2105 0.3609 0.3083
```

Graphing a Markov matrix

```
require(igraph)
M <- rbind(c(0, 0.75, 0, 0.25), c(0, 0, 0.25, 0.75), c(1/3, 1/3, 0, 1/3), c(0,
    0, 1, 0))
M.graph <- graph.adjacency(M, weighted = TRUE)
plot(M.graph)
```



(Check out the very cool: `tkplot(M.graph)` and `rglplot()`functions).

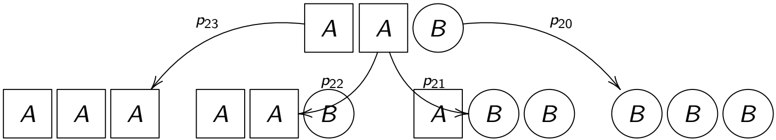
Example 2: Genetic Drift

Consider a population of size N that features 2 **alleles** (A and B) at a given **locus**.

$$\boxed{A}\ \boxed{A}\ B$$

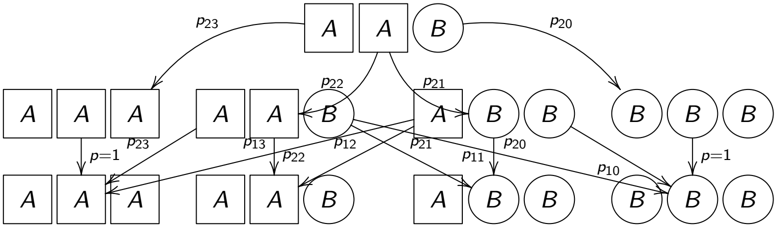
Example 2: Genetic Drift

Consider a population of size N that features 2 **alleles** (A and B) at a given **locus**.



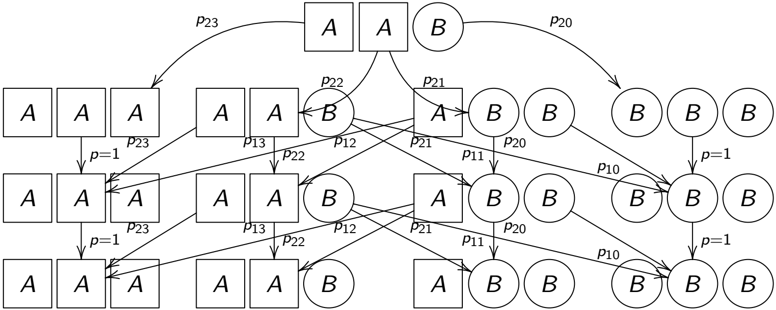
Example 2: Genetic Drift

Consider a population of size N that features 2 **alleles** (A and B) at a given **locus**.



Example 2: Genetic Drift

Consider a population of size N that features 2 **alleles** (A and B) at a given **locus**.



If the State X is defined as number of A alleles in the population, then:

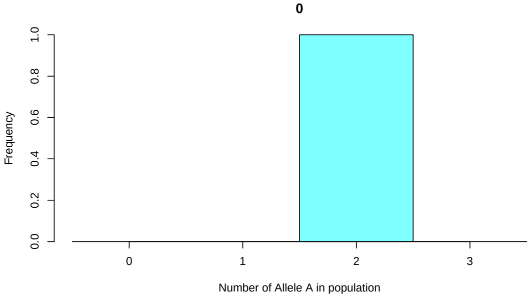
$$\mathbf{M} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ \left(\frac{2}{3}\right)^3 & 3\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)^2 & 3\left(\frac{2}{3}\right)\left(\frac{1}{3}\right)^2 & \left(\frac{1}{3}\right)^3 \\ \left(\frac{1}{3}\right)^3 & 3\left(\frac{2}{3}\right)\left(\frac{1}{3}\right)^2 & 3\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)^2 & \left(\frac{2}{3}\right)^3 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

$$\mathbf{M} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{pmatrix} 1.000 & 0.000 & 0.000 & 0.000 \\ 0.296 & 0.444 & 0.222 & 0.037 \\ 0.037 & 0.222 & 0.444 & 0.296 \\ 0.000 & 0.000 & 0.000 & 1.000 \end{pmatrix} \end{array}$$

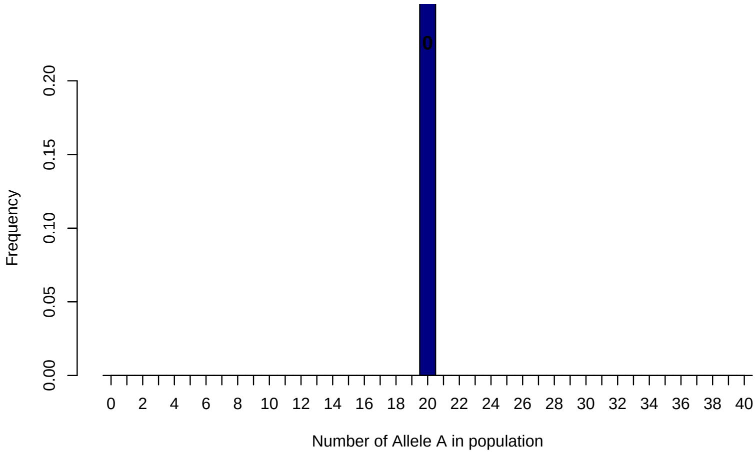
N=4

```
MakeFW <- function(N) {
    M <- matrix(NA, N + 1, N + 1)
    for (i in 1:(N + 1)) M[i, ] <- dbinom(0:N, N, (i
    return(M)
}
```

```
N <- 4
M <- MakeFW(N)
X <- rep(0, N + 1)
X[round(N/2) + 1] <- 1
for (i in 1:20) {
    barplot(as.vector(X), main = i - 1)
    X <- X %*% M
}
```



N=40



0
Frequency
0.00
0.05
0.10
0.15
0.20
0 2 4 6 8 10 12 14 16 18 20 22 24 26 28 30 32 34 36 38 40
Number of Allele A in population

Fixation and transience

General Fisher-Wright matrix:

$$p_{ij} = Pr\{A_{t+1} = j | A_t = i\} = \binom{2N}{j} \left(\frac{i}{2N}\right)^j \left(1 - \frac{i}{2N}\right)^{2N-j} \qquad (6)$$

The Fisher-Wright matrix is an **absorbing matrix**, and the states 0% and 100% are **absorbing states**. Other states are called **transient** (contrast with **recurrent**), because the process does not necessarily return to them.
Some properties of genetic drift:

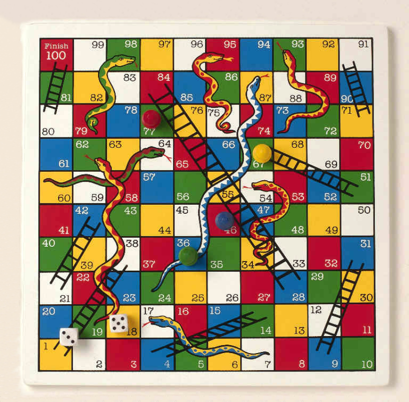
- Always eventually fixates at 0 or N.
- Proportion of fixation depends on initial proportion of a given allele.
- Rate of fixation depends inversely on N

The final moral:

- **Genetic drift** is a Markovian fluctuation in allele frequencies that leads inexorably to fixation for small populations, but is counteracted by **mutation** and **migration** for large populations.

Snakes and Ladders

An adorable example, from:
http://www.r-bloggers.com/basics-on-markov-chain-for-parents/



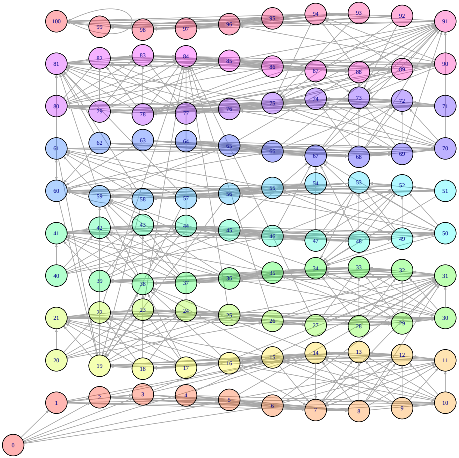
Snakes and Ladders

The Transition Matrix:

```
n = 100

M = matrix(0, n + 1, n + 1 + 6)
rownames(M) = 0:n
colnames(M) = 0:(n + 6)

for (i in 1:6) {
    diag(M[, (i + 1):(i + 1 + n)]) = 1/6
}

M[, n + 1] = apply(M[, (n + 1):(n + 1 + 6)], 1, sum)

M = M[, 1:(n + 1)]

starting = c(4, 9, 17, 20, 28, 40, 51, 54, 62, 64, 63, 71, 93, 95, 92)
ending = c(14, 31, 7, 38, 84, 59, 67, 34, 19, 60, 81, 91, 73, 75, 78)

for (i in 1:length(starting)) {
    v = M[, starting[i] + 1]
    ind = which(v > 0)
    M[ind, starting[i] + 1] = 0
    M[ind, ending[i] + 1] = M[ind, ending[i] + 1] + v[ind]
}

row.names(M) <- 0:100
M.graph <- graph.adjacency(M, weighted = TRUE)
```

Snakes and Ladders

```
xs <- c(0, rep(c(1:10, 10:1), 5))
ys <- c(0, rep(1:10, each = 10) + sin(seq(0, 2 * pi, length = 10))/5)

plot(M.graph, layout = cbind(xs, ys), edge.arrow.size = 0.2, vertex.size = 10,
    vertex.color = rainbow(101, alpha = 0.3), vertex.label.cex = 0.5)
```



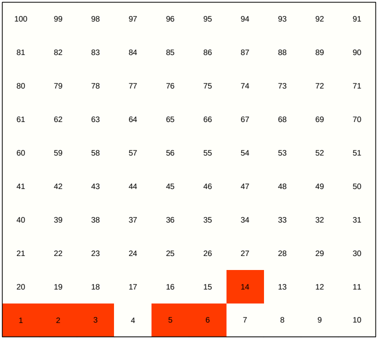
What can the Transition Matrix tell us?

The distribution after t turns:

$$X_t = X_0 \mathbf{M}^t$$

(where X_0 is $\{1, 0, 0, \ldots\}$).

Position after 1 turns



What can the Transition Matrix tell us?

Your daughter loses her piece on the 10th move, and insists it was on square 60. Here are the conditional probabilities that it is one of 58, 59 or 60:

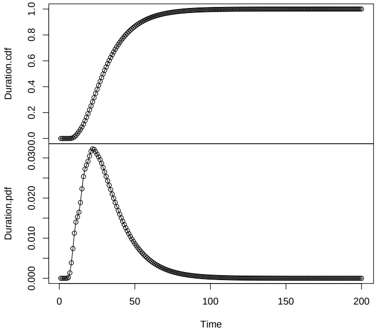
```
pi.0 <- c(1, rep(0, 100))
(pi.0 %*% powermat(M, 10))[59:61]/sum((pi.0 %*% powermat(M, 10))[59:61])

## [1] 0.1597 0.5168 0.3235
```

What can the Transition Matrix tell us?

How long can you expect to play?

```r
pi.t = pi.0 %*% M

Duration.cdf <- rep(NA, 200)
for (h in 1:200) {
    Duration.cdf[h] = pi.t[n + 1]
    pi.t = pi.t %*% M
}

Duration.pdf <- diff(Duration.cdf)
sum(Duration.pdf * 1:length(Duration.pdf))

## [1] 32.16

sum(1 - Duration.cdf)

## [1] 32.16
```

```r
plot.ts(cbind(Duration.cdf, Duration.pdf), type

## Warning: number of rows of result is
not a multiple of vector length (arg 2)
```



These are “extinction” (or “gambler’s ruin”) problems.

What can the Transition Matrix tell us?

How long will you be playing with two or three players?

```
sum((1 - Duration.cdf)^2)

## [1] 23.4

sum((1 - Duration.cdf)^3)

## [1] 20.02
```

With M, everything is Quick and Easy!

Historical aside on stochastic processes

Andrei Andreevich Markov (1856-1922) was a Russian mathematician who came up with the most widely used formalism and much of the theory for

A passionate pedagogue, he was a strong proponent of problem-solving over seminar-style lectures. A politically principled activist, he refused tsarist honors, requested that he be excommunicated from the Russian Orthodox Church out of solidarity with the recently excommunicated **Leo Tolstoy**, publicly renounced his "membership in the electorate" when Parliament was dissolved, and eventually left his teaching post when the government demanded that teachers spy on students.

He said this of his most famous English colleague: *"I can judge all work only from a strictly mathematical point of view and from this viewpoint it is clear to me that ... Pearson has done nothing of any note."*[2]

---

[2] from: Basharin et al. (2005) The Life and Work of A.A. Markov