Using Markov chains

Why not !?

Kévin Cazelles

UQAR/ISEM

August 29, 2014

What we gonna do?

-outline-

- A few notes on probabilities
- Markov chain you've said !
- Few properties of MC
- Applied example : community based approach of MacArthur and Wilson model

A few notes on probabilities

Probability space

- Probability space : $(\Omega, \mathcal{A}, \mathbb{P})$

A few notes on probabilities
Probability space

- Probability space : $(\Omega, \mathcal{A}, \mathbb{P})$
  - Ω : Sample space ("l'univers"), set of possible outcomes

A few notes on probabilities

Probability space

- Probability space : $(\Omega, \mathcal{A}, \mathbb{P})$
  - Ω : Sample space ("l'univers"), set of possible outcomes
  - $\mathcal{A}$: σ-algebra (σ-field), events (particular set of events)

A few notes on probabilities
Probability space

- Probability space : $(\Omega, \mathcal{A}, \mathbb{P})$
  - Ω : Sample space ("l'univers"), set of possible outcomes
  - $\mathcal{A}$: σ-algebra (σ-field), events (particular set of events)
  - $\mathbb{P}$: a probability measure fonction ; $\mathbb{P} : \mathcal{A} \rightarrow [0, 1]$
- Ω : all possibilities, $\mathcal{A}$: how to combine them, $\mathbb{P}$: gives the values in $[0, 1]$ of the possible combination that $\mathcal{A}$ describes ;

A few notes on probabilities

Rules σ-algebra follow

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be our probability space then :

1. $\emptyset \in \mathcal{A}$
2. $A \in \mathcal{A} \Rightarrow \Omega \backslash A \in \mathcal{A}$ (we will denote $\Omega \backslash A$ by $\overline{A}$)
3. $n \in \mathbb{N}$, $(A_1, ..., A_n) \in \mathcal{A}^n \Rightarrow \bigcup_{i=1}^{n} A_i \in \mathcal{A}$

A few notes on probabilities

Rules the probability measure follow

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be our probability space then :

1. $A \in \mathcal{A}, \mathbb{P}(A) \in [0, 1]$
2. $\mathbb{P}(\Omega) = 1$
3. $n \in \mathbb{N},\ (A_1, ..., A_n) \in \mathcal{A}^n \Rightarrow \mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \mathbb{P}(A_i)$

A few notes on probabilities

The famous heads or tails example

1. We toss a coin once and so :

A few notes on probabilities
The famous heads or tails example

1. We toss a coin once and so :
   - $\Omega = \{"heads","tails"\} or \mathbf{\{0,1\}}$

A few notes on probabilities

The famous heads or tails example

1. We toss a coin once and so :
   - $\Omega = \{''\textit{heads}'', ''\textit{tails}''\} \textit{or} \mathbf{\{0,1\}}$
   - $\mathcal{A} = \{\emptyset, 0, 1, \Omega\}$

A few notes on probabilities
The famous heads or tails example

1. We toss a coin once and so :
   - $\Omega = \{"\textit{heads}", "\textit{tails}"\} \textit{or} \mathbf{\{0,1\}}$
   - $\mathcal{A} = \{\emptyset, 0, 1, \Omega\}$
   - $\mathbb{P} : \{\emptyset, 0, 1, \Omega\} \rightarrow [0, 1]$

A few notes on probabilities

The famous heads or tails example

1. We toss a coin once and so :
   - $\Omega = \{"heads","tails"\} or \mathbf{\{0,1\}}$
   - $\mathcal{A} = \{\emptyset, 0, 1, \Omega\}$
   - $\mathbb{P} : \{\emptyset, 0, 1, \Omega\} \rightarrow [0, 1]$

   Based on a wonderful experience, you can write, for instance : $\mathbb{P}(0) = \mathbb{P}(1) = 0.5$ or $\mathbb{P}(0) = 0.25$ and $\mathbb{P}(1) = 0.75$ if the coin is biased
2. We toss a coin twice :

A few notes on probabilities

The famous heads or tails example

1. We toss a coin once and so :
    - $\Omega = \{"heads", "tails"\} or \mathbf{\{0,1\}}$
    - $\mathcal{A} = \{\emptyset, 0, 1, \Omega\}$
    - $\mathbb{P} : \{\emptyset, 0, 1, \Omega\} \rightarrow [0, 1]$

    Based on a wonderful experience, you can write, for instance : $\mathbb{P}(0) = \mathbb{P}(1) = 0.5$ or $\mathbb{P}(0) = 0.25$ and $\mathbb{P}(1) = 0.75$ if the coin is biased
2. We toss a coin twice :
    - $\Omega = \{0, 1\}^2 = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$

A few notes on probabilities

The famous heads or tails example

1. We toss a coin once and so :
   - $\Omega = \{"heads","tails"\} or \mathbf{\{0,1\}}$
   - $\mathcal{A} = \{\emptyset, 0, 1, \Omega\}$
   - $\mathbb{P} : \{\emptyset, 0, 1, \Omega\} \rightarrow [0, 1]$

   Based on a wonderful experience, you can write, for instance : $\mathbb{P}(0) = \mathbb{P}(1) = 0.5$ or $\mathbb{P}(0) = 0.25$ and $\mathbb{P}(1) = 0.75$ if the coin is biased
2. We toss a coin twice :
   - $\Omega = \{0, 1\}^2 = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$
   - $\mathcal{A} = \mathcal{P}(\Omega) = \mathcal{T}((0, 0), (0, 1), (1, 0), (1, 1))$

A few notes on probabilities

The famous heads or tails example

1. We toss a coin once and so :
   - $\Omega = \{"\textit{heads}","\textit{tails}"\} \textit{or} \mathbf{\{0,1\}}$
   - $\mathcal{A} = \{\emptyset, 0, 1, \Omega\}$
   - $\mathbb{P} : \{\emptyset, 0, 1, \Omega\} \rightarrow [0, 1]$

   Based on a wonderful experience, you can write, for instance : $\mathbb{P}(0) = \mathbb{P}(1) = 0.5$ or $\mathbb{P}(0) = 0.25$ and $\mathbb{P}(1) = 0.75$ if the coin is biased
2. We toss a coin twice :
   - $\Omega = \{0, 1\}^2 = \{(0, 0), (0, 1), (1, 0), (1, 1)\}$
   - $\mathcal{A} = \mathcal{P}(\Omega) = \mathcal{T}((0, 0), (0, 1), (1, 0), (1, 1))$
   - $\mathbb{P} : \mathcal{P}(\Omega) \rightarrow [0, 1]$

A few notes on probabilities
Conditional probability

Let $(\Omega, \mathcal{A}, \mathbb{P})$ be our probability space, $(A, B) \in \mathcal{A}^2$ / $\mathbb{P}(B) > 0$, we define :

- $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{P(B)}$

Now $\mathbb{P}(B) > 0$ and $\mathbb{P}(A) > 0$, A and B are independent if and only if :

- $\mathbb{P}(A|B) = \mathbb{P}(A) \Leftrightarrow \mathbb{P}(B|A) = P(B)$

A few notes on probabilities

Bayes Formula

Bayes formula (extended version): Let $(\Omega, \mathcal{A}, \mathbb{P})$ be our probability space, let $I \subset \mathbb{N}$, $(B_i)_{i \in I}$ be a family of events, such as :

- $(B_i)_{i \in I} \; / \; \forall (i,j) \in IxI \backslash \{i\}, \; B_i \cap B_j = 0$
- $\bigcup_{i \in I} B_i = \Omega$
- $\forall i \in I, \mathbb{P}(B_j) > 0$

then :$\mathbb{P}(A) = \sum_{i \in I} \mathbb{P}(A|B_i)\mathbb{P}(B_i)$

A few notes on probabilities
Random variables

Random variable (or stochastic variable) : Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, let $(E, \mathcal{E})$ be a measurable space, a $(E, \mathcal{E})$-valued random (or stochastic) variable is a function $X : \Omega \rightarrow E$ which is $(\mathcal{A}, \mathcal{E})$-measurable.

A few notes on probabilities

Random variables

Random variable (or stochastic variable) : Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, let $(E, \mathcal{E})$ be a measurable space, a $(E, \mathcal{E})$-valued random (or stochastic) variable is a function $X : \Omega \rightarrow E$ which is $(\mathcal{A}, \mathcal{E})$-measurable.

To be clear, based on the heads and tails example, we define a random variable X such as :

$$P("heads") = P(X = 1) \qquad P("tails") = P(X = 0)$$

A few notes on probabilities
Random variables

Random variable (or stochastic variable) : Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, let $(E, \mathcal{E})$ be a measurable space, a $(E, \mathcal{E})$-valued random (or stochastic) variable is a function $X : \Omega \rightarrow E$ which is $(\mathcal{A}, \mathcal{E})$-measurable.

To be clear, based on the heads and tails example, we define a random variable X such as :

$$P(''heads'') = P(X = 1) \qquad P(''tails'') = P(X = 0)$$

But this is a shorthand we always use ! Remember this means $X(''heads'') \in \mathcal{A}$, and particularly, $X(''heads'') = 1$
Think about the "tossing twice" experience ....

A few notes on probabilities

Probability distribution of a random variable

Probability distribution : Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, let X be $(E, \mathcal{E})$-valued random variable, the probability distribution of X is the measure $\mathbb{P}_X$ defined on $(E, \mathcal{E})$ such as for any $E_i \in \mathcal{E}$:
$\mathbb{P}_X(E_i) = \mathbb{P}(X^{-1}(E_i)) = \mathbb{P}(X \in E_i)$

A few notes on probabilities
Probability distribution of a random variable

Probability distribution : Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, let X be $(E, \mathcal{E})$-valued random variable, the probability distribution of X is the measure $\mathbb{P}_X$ defined on $(E, \mathcal{E})$ such as for any $E_i \in \mathcal{E}$:
$\mathbb{P}_X(E_i) = \mathbb{P}(X^{-1}(E_i)) = \mathbb{P}(X \in E_i)$

To be clear, a probability distribution associated to each values of the space E a value between $[0, 1]$ ("1 probability") between such as it reflects the situation existing in Ω.

A few notes on probabilities

Stochastic process

Stochastic process (or random process) : Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, let $(E, \mathcal{E})$ be a measurable space, and T be a totally ordered set a stochastic process is collection of random variables ordered by T : $\{X_t : t \in T\}$.

A few notes on probabilities

Stochastic process

Stochastic process (or random process) : Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, let $(E, \mathcal{E})$ be a measurable space, and T be a totally ordered set a stochastic process is collection of random variables ordered by T : $\{X_t : t \in T\}$.

For instance, if we toss a coin, X is then our random values which takes values in $\{0, 1\}$. We can then toss the coin each fucking minute and register each result, we so have to define random process indexed by the time (here discrete). $(X_t, t \in \mathbb{N})$.

Markov chain you've said !

States space

- $\mathcal{S}$ a set of states, $Card(\mathcal{S}) = k$ with $k \in \mathbb{N}^*$,
- $\mathcal{S}$ and $\mathbb{N}_{k-1}$ (or $\mathbb{N}_k^*$) are isomorphic,

Markov chain you've said !

States space

- $\mathcal{S}$ a set of states, $Card(\mathcal{S}) = k$ with $k \in \mathbb{N}^*$,
- $\mathcal{S}$ and $\mathbb{N}_{k-1}$ (or $\mathbb{N}_k^*$) are isomorphic,

Examples :

- Traffic lights : $\mathcal{S} = \{''red'', ''orange'', ''green''\}$, $\mathcal{S}$ and $\{0, 1, 2\}$ are isomorphic.
- Biogeography: $\mathcal{S} = \{$"Sp1 on the island", "Sp1 not on the island"$\}$
- Population dynamics : $\mathcal{S} = \mathbb{N}$

Markov chain you've said !

Stochastic process (random process)

- Let $\mathbb{T}$ be a totally ordered set and $t \in \mathbb{T}$,
- $(X_t)_{t\in\mathbb{T}}$ is a ($\mathcal{S}$-values) stochastic process then : $\forall t \in \mathbb{T}$, X_t is a random variables of states space $\mathcal{S}$.

Examples :

- Tossing a coin each 10 seconds $(X_t)_{t\in\mathbb{N}}$ records the results
- Population dynamics : $(X_t)_{t\in\mathbb{R}^+}$ records the number of individuals

Markov chain you've said !

Markov process

Based on the work of Andreï Andreïevitch Markov (1856-1922) we define :

- A finite markov process $(X_t)_{t\in\mathbb{N}}$ is a finite stochastic process such that, $\forall t \geqslant 0$:

Markov chain you've said !

Markov process

Based on the work of Andreï Andreïevitch Markov (1856-1922) we define :

- A finite markov process $(X_t)_{t\in\mathbb{N}}$ is a finite stochastic process such that, $\forall t \geqslant 0$:

$$\mathbb{P}\left(\bigcap_{j=0}^{t} X_j\right) > 0$$

$$\mathbb{P}\left(X_{t+1} = i | \bigcap_{j=0}^{t} X_j\right) = \mathbb{P}(X_{t+1} = i | X_t)$$

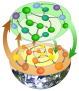
The latter conditional probability is called a **transition probability**.

Markov chain you've said !

Markov Chain

A **Finite Markov Chain** is a finite Markov process for which the transition probabilities do not depend on t.

- A finite Markov process can also be referred as a "no memory finite stochastic process"
- You can also find "homogenous finite Markov chain"
- I wrote the definitions above according to Kemeny, J. G., and Snell, J. L. (1960). Finite markov chains (Springer., Vol. 40, p. 210).



Markov chain you've said !

Extended definition : Markov chain of order $m > 1$

$m \in \mathbb{N}^*$, a finite Markov chain of order (a m-memory process) is a finite stochastic process such that :

$$\forall t \geqslant m$$

$$\mathbb{P}\left(\bigcap_{j=0}^{t} X_p\right) > 0$$

$$\mathbb{P}\left(X_{t+1} = i | \bigcap_{j=0}^{n} X_p\right) = \mathbb{P}\left(X_{t+1} = i | \bigcap_{j=t-m+1}^{n} X_j\right)$$

the latter does not depend on t

Markov chain you've said !

About the next sections

- MC stands for Markov Chain
- we consider solely finite MC (order 1)
- the increment is then 1 but can be regarded as *dt*
- main question : $\forall (i,t) \in \mathbb{N}_{k-1} \times \mathbb{N}, \mathbb{P}(X_t = i) = ?$

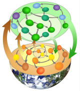
Markov chain you've said !

About the next sections

- MC stands for Markov Chain
- we consider solely finite MC (order 1)
- the increment is then 1 but can be regarded as dt
- main question : $\forall (i,t) \in \mathbb{N}_{k-1} \times \mathbb{N}, \mathbb{P}(X_t = i) = ?$

For such MC :

- a MC can be regraded as random walk on a graph
- All we know about the MC is given by the transition probabilities



Markov chain you've said !

Transition matrix

- The **transition matrix** of a Markov chain is defined by :

$$\mathbf{P} = (p_{i,j}, (i,j) \in \mathbb{N}^2_{k-1});\ \ p_{i,j} = \mathbb{P}(X_{t+1} = j | X_t = i)$$

Markov chain you've said !

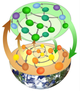
Transition matrix

- The **transition matrix** of a Markov chain is defined by :

$$\mathbf{P} = (p_{i,j}, (i,j) \in \mathbb{N}^2_{k-1}); \;\; p_{i,j} = \mathbb{P}(X_{t+1} = j | X_t = i)$$

- Also called a stochastic matrix, the sum of each line provides 1.
- This is directly provided by the total probability formula :

$$\sum_{j=0}^{k-1} p_{i,j} = \sum_{j=0}^{k-1} \mathbb{P}(X_{t+1} = j | X_t = i) = 1$$



Markov chain you've said !

describing a Markov Chain

A finite stochastic process $X_{t, t \in \mathbb{N}}$ is a Markov Chain with an initial distribution Λ_0 and a transition matrix $\mathbf{P}$: $p_{i,j}, (i,j) \in \mathbb{N}_{k-1}^2$ if :

- $\forall i \in \mathbb{N}_{k-1},\ \mathbb{P}(X_0 = i) = \lambda_{i,0}$
- $\forall t \geqslant 0,\ \mathbb{P}(X_{t+1} = j | X_t = i) = p_{i,j}$

Markov chain you've said !

describing a Markov Chain

A finite stochastic process $X_{t, t \in \mathbb{N}}$ is a Markov Chain with an initial distribution Λ_0 and a transition matrix **P** : $p_{i,j}, (i,j) \in \mathbb{N}^2_{k-1}$ if :

- $\forall i \in \mathbb{N}_{k-1},\ \mathbb{P}(X_0 = i) = \lambda_{i,0}$
- $\forall t \geqslant 0,\ \mathbb{P}(X_{t+1} = j | X_t = i) = p_{i,j}$

Then we describe the transition between n and $n+1$ as follows :

$$\mathbb{P}(X_{t+1} = i) \quad = \quad \sum_{i=0}^{k-1} p_{i,j} \mathbb{P}(X_t = j)$$

Markov chain you've said !
describing a Markov Chain

A finite stochastic process $X_{t,t\in\mathbb{N}}$ is a Markov Chain with an initial distribution Λ_0 and a transition matrix $\mathbf{P}$: $p_{i,j}, (i,j) \in \mathbb{N}_{k-1}^2$ if :

- $\forall i \in \mathbb{N}_{k-1},\ \mathbb{P}(X_0 = i) = \lambda_{i,0}$
- $\forall t \geqslant 0,\ \mathbb{P}(X_{t+1} = j | X_t = i) = p_{i,j}$

Then we describe the transition between n and $n+1$ as follows :

$$\mathbb{P}(X_{t+1} = i) \quad = \quad \sum_{i=0}^{k-1} p_{i,j}\mathbb{P}(X_t = j)$$

We can show :

$$\forall n \geqslant 0, \Lambda_t \quad = \quad \Lambda_0 \mathbf{P}^t$$

Few properties of MC

Ergotic Markov chain

- An irreducible (or ergotic) finite MC is a finite MC with only 1 closed communicative class, it is possible to go from every state to every state
- A regular finite MC is an ergotic MC such that : $\exists k \mid \mathbf{P}^k > 0$

Few properties of MC

Ergotic Markov chain

- An irreducible (or ergotic) finite MC is a finite MC with only 1 closed communicative class, it is possible to go from every state to every state
- A regular finite MC is an ergotic MC such that : $\exists k \mid \mathbf{P}^k > 0$

<u>Theorem</u> :

- A MC with a **P** transition matrix is ergotic iff the eigen value 1 is simple and the only eigen value whose module is one
- Its limiting distribution π is given by the unique normalized left side eigen vector associated
- This also provides the portion of time spent in each states

Continuous Markov Chain

- The ordered set is now continuous
- ....
- Work in progress