LTCC:
Stochastic Processes

Maria De Iorio

Department of Statistical Science
University College London
m.deiorio@ucl.ac.uk
Course structure

- 5 × 2 hr lectures
- every Monday: Oct. 9 to Nov. 6 (inclusive) usually at 11:15 - 13:00hrs (except next week at 12!!!!!!)
- assignment:
  - handed out (emailed) in ‘Spring-time’
  - typically, number of questions ∈ {1, 2, 3}
  - you will have at least one week to complete and hand in.
  - should only take ≤ ‘a few’ hours
  - state space of assessment = {0, 1, 2} or {A, B, C}. 


Recommended literature

- Ross (1996) *Stochastic Processes* [Wiley]
What is it about?

Introduction to

- discrete time Markov processes
- continuous time Markov processes
Motivation: Markov chains/processes

- **Physics** thermodynamics & statistical mechanics
- **Chemistry** enzyme activity models
- **Biology** epidemic modelling
- **Sociology** population dynamics
- **Audio** music restoration; hidden Markov models used in speech recognition (and bioinformatics and too many other things to mention)
- **Operational research** queueing theory, game theory, (baseball!? etc
- **Telecommunications** networks etc
- **Internet** search engines, on-line fraud
- **Computational statistics** Markov chain Monte Carlo (Markov random fields)
- **Economics** modelling asset prices, market crashes, etc
- **Text** language can be modelled as a Markov chain; ‘who wrote this text’; spam arms race, etc
- and many, many, more!
Outline

1 preliminaries
   • stochastic processes
   • Markov process

2 Markov chains
   • transition probabilities
   • example
   • notation
   • questions

3 examples
   • gambler’s ruin
   • urn model
   • 1st order weather model
   • 2nd order weather model

4 $n$-step transition probabilities
   • notation
   • Chapman-Kolmogorov
**Definition 1**

A stochastic process is a family of random variables \( \{X(t), t \in \mathcal{T}\} \), with some indexing set \( \mathcal{T} \) (informally ‘time’).

**Remark 2**

Indexing set can be

\[
\mathcal{T} \subseteq \mathbb{Z} \quad \text{discrete} \\
\mathcal{T} = [a, b] \subseteq \mathbb{R} \quad \text{continuous}
\]

**Definition 3**

The set of all values that \( X \) takes, namely \( S:= \{x: X(t) = x, t \in \mathcal{T}\} \) is called the state space of \( X \).

**Remark 4**

For each \( t \), \( X(t) \) is a r.v.:

\[
X(t): \Omega \mapsto S
\]

and we sometimes write \( X: \Omega \times \mathcal{T} \mapsto S \).
Note on terminology

If $S = \{0, 1, 2, \ldots\}$, we refer to the process as integer valued or discrete state process.

If $S =$ real line, we call $X(t)$ a real-valued stochastic process.

If $S$ is Euclidean $k$ space, $X(t)$ is called a $k$-vector process.

If $T = \{0, 1, 2, \ldots\}$, we refer to $X(t)$ as a discrete time stochastic process.

If $T = [0, \infty)$, we refer to $X(t)$ as a continuous time stochastic process.
Example 5

Flip a fair coin \( n \) times. Let \( X(n) = \# \) heads after \( n \) flips. Then
\[
X: \Omega \times T \mapsto S \quad \text{with sample space } \Omega = \{ \text{'heads', 'tails'} \}, \text{ and } T = \mathbb{N}, \text{ with state space } S = \mathbb{N}_0.
\]

Definition 6 (Markov process)

A stochastic process is called a Markov process if it satisfies the Markov property, namely
\[
\mathbb{P}\left( X(t) \leq x \bigg| X(t_n) = x_n, X(t_{n-1}) = x_{n-1}, \ldots, X(t_0) = x(t_0) \right) = \mathbb{P}\left( X(t) \leq x \bigg| X(t_n) = x_n \right)
\]
\[\forall t_0 < t_1 < \ldots < t_n < t \in T \text{ and } \forall x_0, x_1, \ldots, x_n, x \in S. \text{ I.e. the past and future are conditionally independent, given the present.}\]
Example 7 (random walk)

Define

\[ X(n + 1) := \begin{cases} 
X(n) + 1, & \text{head} \\
X(n) - 1, & \text{tail} 
\end{cases} \quad X(0) = 0 \]

Then \( T = \mathbb{N}, \ S \subseteq \mathbb{Z}, \) and \( X \) is Markov. [Given value of \( X(n) \), the value of \( X(n + 1) \) is independent of \( X(n - 1), \ldots, X(0) \).]

We assume that the random direction of each jump is independent of all earlier jumps.
Remark 8 (Markov is one-step away from independence)

Consider joint distribution of a stochastic process $X$:

$$
P(X(t_0), \ldots, X(t_n)) = P(X(t_0)) \times P(X(t_1)|X(t_0)) \times P(X(t_2)|X(t_1), X(t_0)) \times P(X(t_3)|X(t_2), X(t_1), X(t_0)) \times \cdots \times P(X(t_n)|X(t_{n-1}), X(t_{n-2}), \ldots, X(t_0))$$

If $X$ is Markov, then this collapses to

$$P(X(t_0), \ldots, X(t_n)) = P(X(t_0)) \prod_{i=1}^{n} P(X(t_i)|X(t_{i-1}))$$

Hence Markov processes are one-step away from independence.
Note on terminology

<table>
<thead>
<tr>
<th></th>
<th>discrete</th>
<th>continuous</th>
</tr>
</thead>
<tbody>
<tr>
<td>discrete, countable</td>
<td>discrete-time Markov chain</td>
<td>continuous-time Markov chain</td>
</tr>
<tr>
<td>continuous</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

**Definition 9 (discrete-time Markov chain)**

A *discrete-time process* \( \{X_n, n \in \mathbb{N}_0\} \) *with countable discrete state space* \( S \) *is a Markov chain* if

\[
\mathbb{P}( X_{n+m} = j \mid X_n = i_n, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0 ) = \mathbb{P}( X_{n+m} = j \mid X_n = i_n ), \forall i_0, i_1, \ldots, i_n, j \in S; \forall m, n \in \mathbb{N}_0
\]
Definition 10 (transition probability)

\[ P(X_{n+1} = j | X_n = i) \] is known as the (one-step) transition probability.

Generally, the one-step transition probability depends on three indexes: \( i, j, \) and \( n \). We will consider the case where this is constant w.r.t. \( n \).

Definition 11 (time-homogeneity)

Markov chain \( X \) is time-homogenous if

\[ P(X_{n+1} = j | X_n = i) = P(X_1 = j | X_0 = i), \forall n \in \mathbb{N}_0 \text{ and } i, j \in S \]
For $S = \{0, 1, \ldots, N - 1\}$, the transition matrix $P \in [0, 1]^{N \times N}$ is

$$
P = \begin{bmatrix}
p_{00} & p_{01} & \cdots & p_{0,N-1} \\
p_{10} & p_{11} & \cdots & p_{1,N-1} \\
\vdots & \vdots & \ddots & \vdots \\
p_{N-1,0} & p_{N-1,1} & \cdots & p_{N-1,N-1}
\end{bmatrix}
$$
Definition 13

A stochastic matrix is a matrix \( P = (p_{ij})_{i,j \in S} \) which satisfies

1. \( p_{ij} \geq 0, \forall i, j \in S \) \[ P \text{ has non-neg. entries} \]

2. \( \sum_{j \in S} p_{ij} = 1, \forall i \in S \) \[ \text{rows of } P \text{ sum to 1} \]

Theorem 14

The transition matrix is a stochastic matrix.

Proof \( p_{ij} \) is a probability \( \Rightarrow 1 \) and \( \sum_{j \in S} P(X_1 = j | X_0 = i) = 1 \Rightarrow 2 \).
bayer $A$ and player $B$ play a series of games. Now

\[
\begin{align*}
\mathbb{P}(A \text{ wins}) &= p \\
\mathbb{P}(B \text{ wins}) &= 1 - p =: q
\end{align*}
\]

where the outcome of each game is independent. Find the transition probability matrix.

\[
\begin{pmatrix}
p_{-1,-1} & p_{-1,0} & p_{-1,1} \\
p_{0,-1} & p_{0,0} & p_{0,1} \\
p_{1,-1} & p_{1,0} & p_{1,1}
\end{pmatrix}
\]
Remark 15 (marginal distribution notation)

Denote marginal probability that the chain is in state $j$ at time $n$ as:

$$p_j^{(n)} := P(X_n = j)$$

Then the row vector

$$p^{(n)} := (p_j^{(n)})_{j \in S}$$

is the distribution (pmf) of $X_n$ with initial distribution $p^{(0)}$. E.g., for $S = \mathbb{Z}$, the marginal distribution (at time $n$) is $p^{(n)} = [ \ldots, p_{-1}^{(n)}, p_0^{(n)}, p_1^{(n)}, \ldots ]$. 

Example 16

For a random walk, initial distribution is $p^{(0)} = (\delta_j, 0)_{j \in \mathbb{Z}}$. 
Remark 17

Next week or so, we will consider the following:

- $p^{(n)}$ distn. of $A$’s lead over $B$ after $n$ games
- In particular $\lim_{n \to \infty} p_0^{(n)}$, prob. that $A$ and $B$ win equal #games for large $n$ (in the long run)
- distn. of #games played until $X_n = 0$: $[\mathbb{P}(T_0 = n)]_{n \in \mathcal{T}}$ where $T_j := \min\{n > 0: X_n = j\}$
- $\mathbb{E}(T_0)$ mean return time to state 0.
- $X_n \to ?$

Before that, it will be instructive to introduce a couple more examples.
Example 18 (gambler’s ruin)

A and B play for chips; loser pays winner 1 chip.

\[ P(A \text{ wins}) = p \]
\[ P(B \text{ wins}) = 1 - p =: q \]

A starts with a chips, B starts with b chips. Let \( X_n = \#\text{chips } A \text{ has after } n \text{ games} \). Game ends when A or B is bankrupted \((p_{00} = p_{a+b,a+b} = 1)\).

\( X \) is Markov with state space

\[ S = \{ 0, 1, \ldots, a+b \} \]

A is bankrupted \quad \quad B is bankrupted

Transition matrix:

\[
P = \begin{bmatrix}
1 & q & 0 & p \\
q & 0 & p \\
& \ddots & \ddots & \ddots \\
q & 0 & p & 1
\end{bmatrix}
\]
If/when $X_n$ reaches state $0$ or $a + b$, it stays there. The states $0$ and $a + b$ are called absorbing states.

**gambler’s ruin state transition diagram**

$$
\begin{array}{c}
0 & 1 & 2 & \cdots & a+b-2 & a+b-1 & a+b \\
q & p & q & p & q & p & q \\
\end{array}
$$
Remark 19

*Finite state space does not necessarily imply existence of absorbing states.*

Example 20 (urn model)

Consider 2 urns. Urn $A$ contains $N$ white balls. Urn $B$ contains $N$ black balls. At each turn (time index $n = 1, 2, \ldots$) a ball is chosen at random from each urn and the two balls are interchanged. Denote the # of black balls in urn $A$, after $n$th interchange, by $\{X_n, n \in \mathbb{N}_0\}$.

$X$ is Markov. $X_0 = 0$ (urn $A$ starts out with 0 black balls.) State space: $S = \{0, \ldots, N\}$. Transition probabilities:

- one more black ball in $A$: $p_{i,i+1} = \mathbb{P}(X_{n+1} = i + 1 | X_n = i)$
- one less black ball in $A$: $p_{i,i-1} = \mathbb{P}(X_{n+1} = i - 1 | X_n = i)$
- no change in # black balls in $A$: $p_{i,i} = \mathbb{P}(X_{n+1} = i | X_n = i)$
**one more black ball in** \( A \): 
\[
p_{i,i+1} = \mathbb{P}(X_{n+1} = i + 1|X_n = i)
\]
\[
= \mathbb{P}(\text{white from } A, \text{ black from } B|N - i \text{ white in } A, N - i \text{ blacks in } B)
\]
\[
= \left( \frac{N - i}{N} \right)^2 = \left( 1 - \frac{i}{N} \right)^2 =: q_i^*
\]

**one less black ball in** \( A \): 
\[
p_{i,i-1} = \mathbb{P}(X_{n+1} = i - 1|X_n = i)
\]
\[
= \mathbb{P}(\text{black from } A, \text{ white from } B|i \text{ black in } A, i \text{ white in } B)
\]
\[
= \left( \frac{i}{N} \right)^2 =: q_i
\]

**no change in # black balls in** \( A \): 
\[
p_{i,i} = 1 - q_i - q_i^*
\]
\[ P = \begin{bmatrix}
0 & 1 & & & \\
q_1 & 1 - q_1 - q_1^* & q_1^* & & \\
q_2 & 1 - q_2 - q_2^* & q_2^* & & \\
& \ddots & \ddots & \ddots & \\
q_{N-1} & 1 - q_{N-1} - q_{N-1}^* & q_{N-1}^* & & 0
\end{bmatrix} \]

See also state transition diagram. I.e., if/when you are in state 0 (resp. \( N \)), you never stay there, you always go back to state 1 (resp. \( N - 1 \)). States 0 and \( N \) are called reflecting barriers.
Example 21 (weather ‘forecasting’)

1. 2 possible weather conditions on any single day: \{rainy, sunny\}.

2. Tomorrow’s weather depends **only** on today’s weather.

3. \( P(\text{rain tomorrow} | \text{rain today}) = \alpha \)

4. \( P(\text{sunny tomorrow} | \text{sunny today}) = \beta \)

Find transition probabilities.

Solution

\[
Y_n := \begin{cases} 
1, & \text{sunny on } n\text{th day} \\
0, & \text{rains on } n\text{th day} 
\end{cases}
\]

\( S = \{0, 1\} \)

\[
\Rightarrow Y_n \text{ Markov, and}
\]

\[
\Rightarrow p_{00} = \alpha, \quad p_{01} = 1 - \alpha, \quad \Rightarrow p_{10} = 1 - \beta, \quad p_{11} = \beta
\]

I.e.

\[
P = \begin{bmatrix}
\alpha & 1 - \alpha \\
1 - \beta & \beta
\end{bmatrix}
\]
Example 22 (weather ‘forecasting’ II)

Now assume tomorrow’s weather depends only on weather of today and yesterday

Find transition probabilities.

Solution

\[ \mathbb{P}(Y_{n+1} \mid Y_n, Y_{n-1}, \ldots) = \mathbb{P}(Y_{n+1} \mid Y_n, Y_{n-1}) \neq \mathbb{P}(Y_{n+1} \mid Y_n) \]

Hence, \( Y_n \) not Markov (at least the way we have defined it). However...

Remark 23

Define \( \{X_n\} \) s.t.

\[ X_n := Y_{n-1} + 2Y_n, \quad X_0 = 0 \]

i.e. for \( Y_{n-1} = i \) and \( Y_n = j \), we have \( X_n := i + 2j \). then \( X \) is Markov.

Proof ...
\[ P(X_{n+1} \mid X_n, X_{n-1}, \ldots) = P(Y_{n+1}, Y_n \mid Y_n, Y_{n-1}, \ldots) \]
\[ = P(Y_{n+1} \mid Y_n, Y_{n-1}, \ldots) \]
\[ = P(Y_{n+1} \mid Y_n, Y_{n-1}) \]
\[ = P(Y_{n+1}, Y_n \mid Y_n, Y_{n-1}) \]
\[ = P(X_{n+1} \mid X_n) \]

Homework: compute transition matrix of \( X \) and draw state transition diagram
Recall notation from Remark 15:

\[
\begin{align*}
    p_j^{(n)} &= \mathbb{P}(X_n = j) \\
    p_{ij} &= \mathbb{P}(X_{n+1} = j | X_n = i)
\end{align*}
\]

**Definition 24 (n-step transition probabilities)**

Define the **n-step transition probabilities** by

\[
p_j^{(n)} := \mathbb{P}(X_n = j | X_0 = i)
\]

and the **n-step transition matrix** by

\[
P^{(n)} := (p_{ij}^{(n)})_{i,j \in S}
\]

**Remark 25**

\[
P^{(0)} = I, \text{ (the identity matrix): } p_{ij}^{(0)} = \mathbb{P}(X_0 = j | X_0 = i) = \delta_{ij}
\]

**Remark 26**

\[X\text{ time-homogeneous}
\]

\[
\mathbb{P}(X_{n+m} = j | X_m = i) = \mathbb{P}(X_n = j | X_0 = i) = p_{ij}^{(n)} \quad \forall n, m \in \mathcal{T}; \ i, j \in S
\]
Theorem 27 (Chapman-Kolmogorov (cf p215, G&S))

\[ P^{(m+n)} = P^{(m)}P^{(n)}, \text{ for all } m, n \in \mathbb{N}_0 \]

**Proof** I.e., we want to show that:

\[ p_{ij}^{(m+n)} = \sum_{k \in S} p_{ik}^{(m)} p_{kj}^{(n)} \]

for all \( m, n \in \mathbb{N}_0 \) and \( i, j \in S \). LHS is

\[
p_{ij}^{(m+n)} = \mathbb{P}(X_{n+m} = j \mid X_0 = i) \\
= \sum_{k \in S} \mathbb{P}(X_{n+m} = j, X_m = k \mid X_0 = i) \\
= \sum_{k \in S} \mathbb{P}(X_{n+m} = j \mid X_m = k, X_0 = i) \mathbb{P}(X_m = k \mid X_0 = i) \quad [\text{Markov}] \\
= \sum_{k \in S} \mathbb{P}(X_n = j \mid X_0 = k) \mathbb{P}(X_m = k \mid X_0 = i) \quad [t - \text{homog.}] \\
= \sum_{k \in S} p_{ik}^{(m)} p_{kj}^{(n)} \]

Corollary 28

\[ P^{(n)} = P^n \]

**Proof** (i.e., this says the \( n \)th step transition matrix is the \( n \)th power of the one-step transition matrix.) From Chapman-Kolmogorov

\[ P^{(m+n)} = P^{(m)}P^{(n)} \]

But

\[
\begin{align*}
P^{(n)} &= P^{(n-1+1)} \\
&= P^{(n-1)}P(1) \\
&= P^{(n-1)}P \\
&= P^{(n-2)}P(1)P \\
&= P^{(n-2)}PP \\
&= P^{(n-2)}P^2 \\
& \vdots \\
& \vdots \\
& = P^n \quad \blacksquare
\end{align*}
\]
Corollary 29

\[ p^{(n)} = p^{(0)} P^n \]

Proof I.e. this says distn. at time \( n \) is row-matrix product of initial distn. with \( n \)-step transition matrix. We have

\[
p^{(n)}_j = \mathbb{P}(X_n = j) = \sum_{i \in S} \mathbb{P}(X_n = j, X_0 = i)
\]

\[
= \sum_{i \in S} \mathbb{P}(X_0 = i) \mathbb{P}(X_n = j | X_0 = i)
\]

\[
= \sum_{i \in S} p^{(0)}_i p^{(n)}_{ij}
\]

i.e. \( p^{(n)} = p^{(0)} P^{(n)} \). Chapman-Kolmogorov (Corollary 28) completes the proof. ■