Continuous time Markov Chains

Thomas Mountford

EPFL

February 22, 2021

WELCOME

The course will follow the book of Norris closely, though some proofs will be different.
If time permits, we will do some renewal theory at the end.

I will use the Moodle page for discussions, answering questions.
I will have an office hours Friday 14:15-15:00

Probability Theory: A review

A *Probability Space* is a triple $(\Omega, \mathcal{F}, P)$ where

- Ω is a nonempty set.
- $\mathcal{F}$ is a sigma algebra of subsets of Ω. An element of $\mathcal{F}$ Is called an *event*
- P is a probability measure on $\mathcal{F}$. (NOT on Ω!)

Probability Theory:continued

A sigma algebra of subsets of $\Omega, \mathcal{F}$ Is a collection of subsets of Ω such that

- $\Omega \in \mathcal{F}$.
- $A \in \mathcal{F} \implies A^c \in \mathcal{F}$
- $A_1, A_2, \cdots A_k, \cdots \in \mathcal{F} \implies \cup_k A_k \in \mathcal{F}$

Think of events in $\mathcal{F}$ as things we can ask whether they occur (or not). I.e. whether $\omega \in A$ (or not) If we can ask whether A occurs or not, then we can ask whether A^c occurs or not. If we can ask whether each of the A_k occurs or not then we can ask whether at least one of them occurred.

Probability Theory:continued

A probability measure P on $\mathcal{F}$ is a function $P: \mathcal{F} \to [0,1]$ such that

- $P(\Omega) = 1.$
- If $A_1, A_2, \cdots A_k, \cdots \in \mathcal{F}$ are disjoint (i.e. $i \neq j \implies A_i \cap A_j = \emptyset$), then $P(\cup_k A_k) = \sum_k P(A_k)$

We can think of P as a kind of proportion. If we think of $(\Omega, \mathcal{F}, P)$ s a model of an experiment in which an outcome ω is "picked", then if the experiment is repeated *independently* endlessly $P(A)$ is the proportion of times A occurs.

Examples

(1) Consider $\Omega = \{a, b, c\}$, $\mathcal{F} = 2^\Omega$ (i.e. all subsets of Ω) and

$$P(A) = \frac{1}{2}\mathbb{1}_{a\in A} + \frac{1}{3}\mathbb{1}_{b\in A} + \frac{1}{6}\mathbb{1}_{c\in A}$$

(2) $\Omega = [0, 1]$, $\mathcal{F} = \mathcal{B}$, the Borellian subsets of Ω and

$$P(A) = \lambda(A)$$

where λ is Lebesgue measure on Ω

Random Variables

Given a probability space $(\Omega, \mathcal{F}, P)$, a random variable

$$X : \Omega \ \rightarrow \ \mathbb{R}$$

so that

$$\forall B \subset \mathbb{R}, \text{ Borellian, } X^{-1}(B) \in \ \mathcal{F}.$$

(Recall $X^{-1}(B) \ = \ \{\omega : X(\omega) \in B\}$). Or for every Borellian subset B of $\mathbb{R}$, the subset $\{\omega : X(\omega) \in B\}$ is an event.
Note: P plays no role in decding whether X is a random variable. If X is a function to a different space (say $\mathbb{R}^n$), then whave the same definition and use the phrase *random element*.

Given $(\Omega, \mathcal{F}, P)$ and X a random variable on this space, we can create the probability space $(\mathbb{R}, \mathcal{B}, P_X)$ where (CHECK it is a probability)

$$P_X(B) \ = \ P(X^{-1}(B))$$

P_X is the *law* of X.

Random Variables contd

Given a random variable X, its *distribution function*, $F_X(t)$ is the function on $\mathbb{R}$

$$F_X(t) = P_X((-\infty, t])$$

Note F_X has the properties

a F_X is increasing

b F_X is right continuous (with left limits)

c $\lim_{t\to\infty} F_X(t) = 1, \quad \lim_{t\to-\infty} F_X(t) = 0$

In fact if two variables X and Y (not necessarily defined on the same probabiliy space) have the same distribution function, then they have the same distribution.
If a function on $\mathbb{R}$ has the properties a)-c) above, then it is the distribution function of some random variable.

Sub sigma algebras

Given a collection of events $\{A_i\}_{i \in I}$, the sigma field *generated* by $\{A_i\}_{i \in I}$is the smallest sigma field containing $A_i \forall i$. It is written $\sigma\{A_i i \in I\}$. Given random variables $X_1, X_2, \cdots X_N$ the sigma field generated by the $X_i, i \leq N$, $\sigma(X_1, X_2, \cdots X_N)$ is the smallest sigma field with respect to which X_i are measurable. I.e it contains

$$\{X_i \in B\} \ \forall \ i, \forall B \in \mathcal{B}$$

In fact $\sigma(X_1, X_2, \cdots X_N) \ = \ \{(X_1, X_2, X_N) \in A)\}_{A \in \mathcal{B}^N}$

Conditional probability

Given an event B of nonzero probability in a probability space$(\Omega, \mathcal{F}, P)$, the conditional probability of event $A \in \mathcal{F}$ *given* B, $P(A \mid B)$ is defined to be

$$P(A \mid B) \; = \; \frac{P(A \cap B)}{P(B)}$$

You can think of the information that B occured effectively

- replaces Ω by B
- (more generally) replaces events A by $A \cap B$
- replaces P by $P(. \mid B)$

Conditional probability continued

We will use the following repeatedly

Lemma

Law of Total Probability.: Given a partition of Ω into events $B_1, B_2, \cdots B_M$ and any event A

$$P(A) = \sum_{i=1}^{M} P(B_i)P(A \mid B_i)$$

(if $P(B_i) = 0$, we can give $P(A \mid B_i)$ any value in $[0, 1]$ and the formula will still be true.) It is also true for countable partitions.
By induction we obtain

Lemma

Given events $A_1, A_2, \cdots A_N$

$$P(\cap_i A_i) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1 \cap A_2) \cdots P(A_N \mid A_1 \cap A_1 \cap \cdots \cap A_{N-1})$$

For a random variable X (or a random element) taking countably many values $\{i_1, i_2, \cdots\}$, $P(A \mid X)$ is the random variable that is equal to

$$P(A \mid \{X = i_j\}) \text{ on event } \{X = i_j\}$$

Independence

Definition: Two events A and B are said to be *independent* if

$$P(A \cap B) = P(A)P(B)$$

(So if $P(B) > 0, A$ and B are independent if and only if $P(A) = P(A \mid B)$.) If $P(B) = 0$, then B is independent of every other event A

Independence continued

The definition of independence for more than 2 events is a little less straightforward:
Events $A_1, A_2, \cdots A_N$ are independent if for every choice of $D_1, D_2, \cdots D_N$ where $D_i = A_i$ or Ω, we have

$$P(\cap_i D_i) = \prod_i P(D_i)$$

(So A_1, A_2 and A_3 are independent if and only if

$$P(A_1 \cap A_2) = P(A_1)P(A_2),\; P(A_1 \cap A_3) = P(A_1)P(A_3),\; P(A_2 \cap A_3) = P(A_2)P(A_3$$

and $P(A_1 \cap A_2 \cap A_3) = P(A_1)P(A_2)P(A_3)$)

Independence of R.V.s

Random variables $X_1, X_2, \cdots X_N$ are said to be independent if

$$\forall B_i \in \mathcal{B} \quad P\left(X_1 \in B_1, \cdots X_N \in \ B_N\right) \ = \ \prod_i P(X_i \in B_i)$$

This generalizes. Sub sigma fields $\mathcal{G}_I, i = 1, 2 \cdots N$ are independent if for every choice of $A_i \in \mathcal{G}_I$,

$$P\left(\cap_i A_i\right) \ = \ \prod_i P(A_i).$$

(So random variables X_i are independent if and only if the sigma fields $\sigma(X_i)$ are independent.)

Conditional Independence

Definition: Two events A and B are said to be *conditionally independent* given a third event C if

$$P(A \cap B \mid C) \;=\; P(A \mid C)P(B \mid C).$$

Two random variables X, Y are said to be conditionally independent given random variable Z is for each Borels B_1 and B_2

$$P(X \in B_1, Y \in B_2 \mid Z) \;=\; P(X \in B_1 \mid Z)P(Y \in B_2 \mid Z)$$