22/10/2019 22/10/2019

CONTINUOUS TIME (DISCRETE STATE) MARKOV CHAIN

$(X_t)_{t\geq 0}$ random variables with values in a set E (countable or finite).

Markov property: $\mathbb{P}(X_{t_{n+1}}=j \mid X_{t_n}=i_n, \dots, X_{t_0}=i_0) = \mathbb{P}(X_{t_{n+1}}=j \mid X_{t_n}=i_n)$

$\forall n, \ \forall j, i_n, \dots, i_0 \in E \qquad \forall \ 0 \leq t_0 < t_1 < \dots < t_{n+1}$

(equivalent to: $\mathbb{P}(X_{t_{n+1}} \in A_{n+1} \mid X_{t_n} \in A_n, \dots, X_{t_0} \in A_0) = \mathbb{P}(X_{t_{n+1}} \in A_{n+1} \mid X_{t_n} \in A_n)$
$A_0, \dots, A_n \in \mathcal{E}, \quad \mathcal{E} = \mathcal{P}(E)$)

(Def.) Time homogeneous if: $\mathbb{P}(X_{t+s}=j \mid X_s=i) = \underbrace{\mathbb{P}(X_t=j \mid X_0=i)}_{p_{ij}(t)} \quad \forall s$

(Def.) $P_t = (p_{ij}(t))_{i,j\in E}$ transition matrix (at time t) (analog of P^n discrete time)

Properties of P_t: 1. $0 \leq p_{ij}(t) \leq 1$ | P_t is a stochastic matrix

2. $\sum_{j\in E} p_{ij}(t) = 1$

Semi-group property: $P_{t+s} = P_t P_s \quad \forall t,s \geq 0$

or, with indexes: Chapman-Kolmogorov equation:

$$p_{ij}(t+s) = \sum_{k\in E} p_{ik}(t)\, p_{kj}(s)$$

This follows from the Markov property, because:

$$p_{ij}(t+s) = \mathbb{P}(X_{t+s}=j \mid X_0=i) = \sum_{k\in E} \mathbb{P}(X_{t+s}=j, X_t=k \mid X_0=i)$$

$$= \sum_{k\in E} \mathbb{P}(X_{t+s}=j \mid X_t=k, \cancel{X_0=i}) \cdot \underbrace{\frac{\mathbb{P}(X_t=k, X_0=i)}{\mathbb{P}(X_0=i)}}_{=p_{ik}(t)}$$

because of the Markov property

$$= \sum_{k\in E} \underbrace{\mathbb{P}(X_{t+s}=j \mid X_t=k)}_{p_{kj}(s)}\, p_{ik}(t)$$

$$= \sum_{k\in E} p_{kj}(s) \cdot p_{ik}(t)$$

In "regular situations" (unspecified for the moment) the following limits $(t \to 0^+)$ exist:

$$\lim_{t\to 0^+} \frac{p_{ij}(t)}{t} \ (\neq j), \qquad \lim_{t\to 0^+} \frac{-1+p_{ii}(t)}{t}$$

$$\lim_{t\to 0^+} \frac{p_{ij}(t)}{t} := \underline{q_{ij}} \geq 0, \qquad \lim_{t\to 0^+} \frac{p_{ii}(t)-1}{t} := \underline{q_{ii}} \leq 0$$

$\underbrace{q_{ij}, q_{ii}}_{\text{called TRANSITION RATES}} \in \mathbb{R}$ and

$p_{ij}'(t) = \sum_{k\in E} p_{ik}(t)\, q_{kj}$ Forward Kolmogorov equation **FKE**

$p_{ij}'(t) = \sum_{k\in E} q_{ik}\, p_{kj}(t)$ Backward Kolmogorov equation **BKE**

(Initial condition: $p_{ij}(0) = \delta_{ij}$)

(Def.) $Q = (q_{ij})_{i,j\in E}$ is called Transition rate matrix

Properties: 1. $q_{ij} \geq 0 \quad \forall i \neq j \qquad q_{ii} \leq 0$

$\left(\lim_{t\to 0^+} \frac{p_{ii}(t)-1}{t} \leq 0\right)$

2. $\sum_{j\in E} q_{ij} = 0 \qquad \left(\sum_{j\in E} q_{ij} = \lim_{t\to 0^+} \sum_{j\in E} \frac{p_{ij}(t)-\delta_{ij}}{t} = 0\right)$

(Q is almost a stochastic matrix if not that:
1. the diagonal is negative
2. the sum of every row is 0, not 1)

Proof of the Forward Kolmogorov equation

$$p_{ij}'(t) = \lim_{s\to 0^+} \frac{p_{ij}(t-s) - p_{ij}(t)}{-s} \qquad t > 0$$

by Chapman Kolmogorov

$$= \lim_{s\to 0^+} \frac{p_{ij}(t-s) - \sum_k p_{ik}(t-s)\, p_{kj}(s)}{-s}$$

$$= \underbrace{\lim_{s\to 0^+} p_{ij}(t-s) \frac{1-p_{jj}(s)}{-s}}_{p_{ij}(t)\, q_{jj}} - \underbrace{\lim_{s\to 0^+} \sum_{k\neq j} p_{ik}(t-s) \frac{p_{kj}(s)}{-s}}_{\sum_{k\neq j} p_{ik}(t)(-q_{kj})}$$

$$= p_{ij}(t)\, q_{jj} + \sum_{k\neq j} p_{ik}(t)\, q_{kj} \qquad \blacksquare$$

Proof of the Backward Kolmogorov equation:

If we write: $\frac{p_{ij}(t-s) - p_{ij}(t)}{-s} = \frac{p_{ij}(t-s) - \sum_k p_{ik}(s)\, p_{kj}(t-s)}{-s}$

$$= \underbrace{\frac{1-p_{ii}(s)}{-s}}\, p_{ij}(t) - \sum_{k\neq i} \frac{p_{ik}(s)}{-s}\, p_{kj}(t-s)$$

$$\xrightarrow{s\to 0^+} \; = q_{ii}\, p_{ij}(t) + \sum_{k\neq i} q_{ik}\, p_{kj}(t) \qquad \blacksquare$$

$$p_{ij}'(t) = \sum_k p_{ik}(t)\, q_{kj} \implies P_t' = P_t \cdot Q$$

NOTIONS

$$\begin{cases} x' = Ax \\ x(0) = 0 \end{cases} \quad A \text{ matrix (finite, square)} \longrightarrow x(t) = e^{tA} + x_0$$

$$\underbrace{\sum_{n \geq 0} \frac{t^n A^n}{n}}\ ? \qquad \text{We know: } \|A^n\| \leq \|A\|^n$$

$$(\|A^n u\| = \|A(A^{n-1})u\| \leq \|A\| \cdot \|A^{n-1} u\| \ldots$$

Absolutely convergent and it defines e^{tA} in $x(t)$

some properties: ① $e^{(t+s)A} = e^{tA} e^{sA}$

② $\frac{d}{dt} e^{tA} = Ae^{tA} = e^{tA} A$

dim: $\frac{d}{dt} e^{tA} = e^{tA} A$

$$\lim_{s \to 0} \frac{e^{(t+s)A} - e^{tA}}{s} \overset{\text{(def)}}{=} \lim_{s\to 0} \left(\sum_{n=0}^{\infty} \frac{(t+s)^n A^n - t^n A^n}{n! \cdot s} \right)$$

term with $n = 0$ is 0

$$= \lim_{s \to 0^+} \frac{(t+s)A - tA}{1 \cdot s} + \underbrace{\lim_{s \to 0} \sum_{n \geq 2} (\ldots)}$$

$$\lim \sum_{n \geq 2} \frac{(t+s)^n A^n - t^n A^n}{n!\, s}$$

$$\longrightarrow = A + \sum_{n \geq 2} \left(\lim_{s \to 0} \frac{(t+s)^n - t^n}{s} \right) \frac{A^n}{n!}$$

$$= A + \sum_{n \geq 2} n t^{n-1} \frac{A^n}{n! (n-1)!}$$

$$= A \left(Id + \sum_{n \geq 2} \frac{t^n A^n}{n!} \right) = Ae^{tA} \quad \blacksquare$$

③ $\frac{dx(t)}{dt} = A(e^{tA} x) = Ax(t)$

Example: Two state continous time MC:

$$Q = \begin{bmatrix} -a & a \\ b & -b \end{bmatrix} \qquad a, b > 0, \qquad P_t\ ?$$

all'esame non tipo una 8x8 ma conosci già la soluzione di 6/7 eq. e devi calcolare un paio

Let's use the forward Kolmogorov equation.

$$p_{ij}'(t) = \sum_k p_{ik}(t)\, q_{kj}$$

$j = 1$: $\quad p_{i1}'(t) = -p_{i1}(t)\, a + p_{i2}(t)\, b$

$j = 2$: $\quad p_{i2}'(t) = p_{i1}(t)\, a - p_{i2}(t)\, b$

(Example.) $\quad x(t) = p_{i1}(t), \quad y(t) = p_{i2}(t)$

$$\begin{cases} x' = -ax + by \\ y' = ax - by \end{cases}$$

$$p_{i1}(t) + p_{i2}(t) = 1 \implies y(t) = 1 - x(t)$$

$$\Longrightarrow \begin{cases} x'(t) = -(a+b)\, x(t) + b \\ x(0) = \delta_{i1} \end{cases}$$

(formula standard) $\Longrightarrow$

$$x(t) = \delta_{i1} e^{-t(a+b)} + \int_0^t e^{-(t-s)(a+b)}\, b\, ds$$

$$= \delta_{i1} e^{-t(a+b)} + \frac{b}{b+a}\left(1 - e^{-t(a+b)}\right)$$

$i = 1$: $\quad x(t) = e^{-t(a+b)} + \frac{b}{b+a}\left(1 - e^{-t(a+b)}\right)$

$$P_t = \begin{bmatrix} \frac{b + ae^{-t(a+b)}}{a+b} & \frac{a(1 - e^{-t(a+b)})}{a+b} \\ \frac{b(1 - e^{-t(a+b)})}{a+b} & \frac{a + b(1 - e^{-t(a+b)})}{a+b} \end{bmatrix}$$

23/10/2019

Theorem: If $\sup |q_{ii}| < +\infty$ then BKE and FKE have the same unique solution:

$$P_t = e^{tQ}$$

$$\mathbb{P}(X_{t+h} = j \mid X_t = i) \underset{\uparrow}{=} \delta_{ij} + q_{ij} h + o(h)$$

taylor expansion of $p_{ij}(t)$

If $j \neq i$ the probability of a jump $i \rightsquigarrow j$ in an interval $[t, t+h]$ is:

$$q_{ij} h + o(h)$$

in particular the probability of 2 or more jumps is $o(h)$

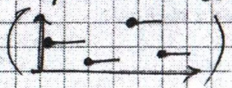
$$\mathbb{P}(X_{t_n} = i_n, X_{t_{n-1}} = i_{n-1}, \ldots, X_0 = i_0) = \qquad (0 < t_1 < \ldots < t_n)$$

$$= \underset{\substack{\uparrow \\ \text{(JD)} \\ \text{Joined distribution}}}{\mathbb{P}(X_n = i_n)\, p_{i_0 i_1}(t_1)\, p_{i_1 i_2}(t_2 - t_1) \cdots p_{i_{n-1} i_n}(t_n - t_{n-1})}$$

$(X_t)_{t \geq 0}$ random variables on $(\Omega, \mathcal{F}, \mathbb{P})$ $\quad \forall \omega$ fixed

$t \longrightarrow X_t(\omega)$ trajectory

One could show that it is always possible, given $(P_t)_{t\geq 0}$ transition semigroup, construct a family of random variables $(X_t)_{t\geq 0}$ with values in E such that (JD) holds and trajectories $t \to X_t(\omega)$ piecewise costant and continuous from the right



Remark: $t \mapsto p_{ij}(t)$ continuous $\Leftarrow$ $t \mapsto X_t$ continuous in probability

proof.

$$|p_{ij}(t+s) - p_{ij}(t)| = |\mathbb{P}_i(X_{t+s}=j) - \mathbb{P}_i(X_t=j)| \qquad t,s>0$$

$$= |\cancel{\mathbb{P}_i(X_{t+s}=j, X_t=j)} + \mathbb{P}_i(X_{t+s}=j, X_t\neq j) -$$

$$- \mathbb{P}_i(X_{t+s}\neq j, X_t=j) - \cancel{\mathbb{P}_i(X_{t+s}=j, X_t=j)}|$$

$$= |\mathbb{P}_i(X_{t+s}=j, X_t\neq j) - \mathbb{P}_i(X_{t+s}\neq j, X_t=j)| \Rightarrow \circledast$$

as $s\to 0$ $\quad |p_{ij}(t+s) - p_{ij}(t)| \to 0 \quad \to \circledast \to 0$ (and then convergence in probability)

Note: $\circledast \leq 2\,\mathbb{P}_i(X_{t+s}\neq X_t) \longrightarrow 0$ if $s \mapsto X_{t+s}$ is continuous in probability ∎

Exponential distribution of exit times $\qquad ((\Omega, \mathcal{F}, \mathbb{P}))$

$i\in E$: $\quad T_i = \inf\{t>0 \mid X_t \neq i\}$

Remark: $\{T_i > t\}$ is an event i.e. $\{T_i>t\}\in\mathcal{F}$ because:

$$\{T_i > t\} = \bigcap_{0\leq s\leq t} \{X_s = i\}$$

$$= \bigcap_{\substack{0\leq s\leq t \,\cap\, \mathbb{Q} \\ \text{or } s=t}} \{X_s = i\}$$

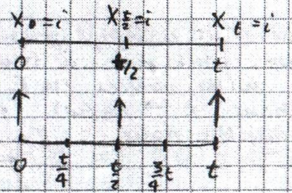
by right continuity of trajection (devo trovare qualcosa di contabile per assicurarmi che sia un evento)

Theorem if $-\infty < q_{ii} < 0$ then $T_i \sim \mathcal{E}(-q_{ii})$

proof.

$$\{T_i > t\} \subseteq \{X_{t\frac{2^n}{2^n}} = i, X_{t\frac{2^n-1}{2^n}} = i, \ldots, X_0 = i\}$$

Indeed $\{T_i > t\} = \bigcap_{n\geq 0} \underbrace{\{X_t = i, X_{t\frac{2^n-1}{2^n}} = i, \ldots, X_0 = i\}}_{\text{decreasing}}$



$n=1$ conditions:

$n=2$ conditions:

$$\mathbb{P}_i(T_i > t) = \lim_{n\to\infty} \mathbb{P}_i(X_t = i, X_{t\frac{2^n-1}{2^n}} = i, \ldots, X_0 = i)$$

$$= \lim_{n\to\infty} \left(p_{ii}\left(\frac{t}{2^n}\right)\right)^{2^n}$$

by Taylor: $\quad p_{ii}\left(\frac{t}{2^n}\right) = 1 + q_{ii}\frac{t}{2^n} + o\left(\frac{t}{2^n}\right)$

$$\lim_{n\to\infty}\left(1 + \frac{q_{ii}t}{2^n} + o\left(\frac{t}{2^n}\right)\right)^{2^n} = e^{q_{ii}t} = e^{-(-q_{ii})t}$$

density $\quad \frac{d}{dt}\mathbb{P}_i(T_i \leq t) = \frac{d}{dt}\left(1 - \mathbb{P}_i(T_i > t)\right) = -q_{ii}e^{q_{ii}t} \sim \mathcal{E}(-q_{ii})$

$-\infty < q_{ii} < 0 \qquad \underbrace{\mathbb{E}_i[T_i]}_{\substack{\text{average} \\ \text{time spent} \\ \text{in } i \text{ starting from } i}} = \frac{1}{-q_{ii}}$ ∎

Note:
- if $q_{ii} = 0$ then i is absorbing: $p_{ii}(t) = 1 \quad \forall t$
- if $q_{ii} = -\infty$ then i is an istantaneous state

Theorem $\quad \forall j \neq i \qquad \mathbb{P}_i(X_{T_i} = j) = \frac{q_{ij}}{-q_{ii}} \qquad (-\infty < q_{ii} < 0)$

where:

X_{T_i} is the random variable "state visited when leaving i"

Remark: X_{T_i} has values in $E - \{i\}$,

$\left(\frac{q_{ij}}{-q_{ii}}\right)_{j\in E-\{i\}}$ is its distribution $\quad \frac{\sum_{j\neq i} q_{ij}}{q_{ii}} = \frac{-q_{ii}}{-q_{ii}} = 1$

Come colleghiamo continuous MC a discrete MC?
($\exists$ una discrete MC associata alla continuous):

$$\hat{p}_{ij} = \begin{cases} \frac{q_{ij}}{-q_{ii}} & \text{if } q_{ii}\neq 0 \ (-\infty < q_{ii} < 0) \\ 1 & \text{if } q_{ii} = 0, i=j \\ 0 & (\text{if } q_{ii}=0 \text{ and } i\neq j) \text{ or } (q_{ii}\neq 0, i=j) \end{cases} \longrightarrow (\hat{p}_{ij})$$

stochastic matrix (discrete) $(i,j\in E)$ called DISCRETE SKELETON

29/10/2019

$(X_t)_{t\geq 0}$ Markov chain with a discrete state space E.

Recap

$T_i = \inf\{t \mid X_t \neq i\}$ exit time from i

$(X_{T_i})(\omega) = \begin{cases} X_{T_i(\omega)} & \text{if } T_i(\omega) < +\infty \\ \text{arbitrary} & \text{if } T_i(\omega) = +\infty \end{cases}$

Theorem: $\forall i$ st. $-\infty < q_{ii} < 0$: $\mathbb{P}_i(X_{T_i} = j) = \dfrac{q_{ij}}{-q_{ii}}$ $\forall j$

proof.

1st discrete approximation of T_i $(T_i \sim \mathcal{E}(-q_{ii}))$

$S_n = \sum_{k=0}^{\infty} (k+1)2^{-n} \chi_{\{k2^{-n} < T_i \leq (k+1)2^{-n}\}}$

$S_n \geq T_i$ $\forall n$ and $S_n \leq T_i + 2^{-n}$ $\forall n$

Consequence $S_n \to T_i$ as $n \to \infty$ almost surely

Moreover $S_n \downarrow$ (increasing sequence). On $\{k2^{-n} < T_i \leq (k+1)2^{-n}\}$ we have:

$S_n = (k+1)2^{-n}$, $n \to n+1$ the set becomes:

$\{2k \cdot 2^{-(n+1)} < T_i \leq (2k+1)2^{-(n+1)}\} \cup \{(2k+1)2^{-(n+1)} < T_i \leq (2k+2)2^{-(n+1)}\}$

$S_{n+1} = \underbrace{(2k+1)2^{-(n+1)}}$ smaller then $S_{n+1} = \underbrace{(2k+2)\cdot 2^{-(n+1)}} = (k+1)2^{-n}$

S_n are stopping times. $\{S_n = (k+1)2^{-n}\} = \{k2^{-n} < T_i \leq (k+1)2^{-n}\}$

$= \underbrace{\{T_i \leq (k+1)2^{-n}\}}_{\in \mathcal{F}_{(k+1)2^{-n}}} \cap \underbrace{\{T_i \leq k2^{-n}\}^c}_{\in \mathcal{F}_{k2^{-n}} \subset \mathcal{F}_{(k+1)2^{-n}}}$ $\Big\}\in \mathcal{F}_{(k+1)2^{-n}}$

Since $S_n \searrow T_i$ (by right continuity of trajectives)

$X_{S_n} \to X_{T_i}$ almost surely

$\mathbb{P}_i(X_{T_i} = j) = \lim_{n\to\infty} \mathbb{P}_i(X_{S_n} = j)$ ⊛

by dominated convergence

$\mathbb{E}_i[\chi_{\{X_{T_i}=j\}}]$ $\mathbb{E}[\chi_{\{X_{S_n}=j\}}]$

$\mathbb{P}_i(X_{S_n} = j) = \sum_{k=0}^{\infty} \mathbb{P}_i(X_{(k+1)2^{-n}} = j, \underbrace{S_n = (k+1)2^{-n}}_{k2^{-n} < T_i \leq (k+1)2^{-n}})$

Remark: if $X_{(k+1)2^{-n}} = j \longrightarrow T_i \leq (k+1)2^{-n}$

$\Rightarrow \{X_{(k+1)2^{-n}} = j\} \subseteq \{T_i \leq (k+1)2^{-n}\}$

22

$\Longrightarrow \{X_{(k+1)2^{-n}} = j, S_n = (k+1)2^{-n}\} = \{X_{(k+1)2^{-n}} = j, T_i > k2^{-n}\}$

$\Longrightarrow$ They're probability is the same (since they're the same):

$\mathbb{P}_i(X_{(k+1)2^{-n}} = j, S_n = (k+1)2^{-n}) = \mathbb{P}_i(X_{(k+1)2^{-n}} = j, T_i > k2^{-n})$

$= \mathbb{P}_i(X_{(k+1)2^{-n}} = j \mid T_i > k2^{-n}) \underbrace{\mathbb{P}(T_i > k2^{-n})}_{e^{q_{ii}k2^{-n}}}$

because T_i for $\mathbb{P}_i \sim \mathcal{E}(-q_{ii})$

$= \mathbb{P}_i(X_{(k+1)2^{-n}} \mid \underbrace{X_{k2^{-n}} = i}_{\text{present}}, \underbrace{\cancel{T_i > k2^{-n}}}_{\text{past}}) e^{q_{ii}k2^{-n}}$ Markov:

$\left(\{T_i > k2^{-n}\} = \bigcap_{0\leq j \leq k2^{-n}} \{X_j = i\}\right)$

Sum up: $\mathbb{P}_i(X_{S_n} = j) = \sum_{k=0}^{\infty} \mathbb{P}(X_{(k+1)2^{-n}} = j \mid X_{k2^{-n}} = i) e^{q_{ii}k2^{(-n)}}$

$= \sum_{k=0}^{\infty} p_{ij}(2^{-n})(e^{q_{ii}2^{-n}})^k$

$= \dfrac{p_{ij}(2^{-n})}{1 - e^{q_{ii}2^{-n}}}$

As $n \to \infty$ $2^{-n} \to 0$, $p_{ij}(2^{-n}) = q_{ii}2^{-n} + o(2^{-n})$

$\lim_{n\to\infty} \mathbb{P}_i(X_{S_n} = j) = \lim_{n\to\infty} \dfrac{q_{ij}2^{-n} + o(2^{-n})}{1 + e^{q_{ii}2^{-n}}}$

$= \lim_{n\to\infty} \dfrac{q_{ij}2^{-n} + o(2^{-n})}{-q_{ii}2^{-n} + o(2^{-n})} = \dfrac{q_{ij}}{-q_{ii}}$ ∎

Now, $(X_t)_{t\geq 0}$ with $Q = (q_{ij})_{ij\in E}$ $\rightsquigarrow$ discrete skeleton $(Y_n)_{n\geq 0}$ with:

$p_{ij} = \begin{cases} \dfrac{q_{ij}}{-q_{ii}} & q_{ii} \neq 0,\ i \neq j \\ 1 & q_{ii} = 0,\ i = j \\ 0 & (q_{ii} = 0,\ i \neq j) \vee (q_{ii} \neq 0,\ i = j) \end{cases}$ $q_{ii} \in [-\infty, 0]$

Example: **POISSON PROCESS**

Markov chain with:

$Q = \begin{bmatrix} -\lambda & \lambda & 0 & 0 & 0 & \\ 0 & -\lambda & \lambda & 0 & 0 & \cdots \\ 0 & 0 & -\lambda & \lambda & 0 & \\ & & \vdots & & & \end{bmatrix}$ $\lambda > 0$

States: $N = \{0, 1, 2, 3, \dots\}$

$i \rightsquigarrow i+1$ after a time $\sim \mathcal{E}(\lambda)$

(example) Discrete skeleton : $\begin{bmatrix} 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ & & \cdots & & \end{bmatrix}$

Proposition : $(N_t)_{t\geq 0}$ MC with the **Q** as in the example.

$$\mathbb{P}_i(N_t = n \mid N_0 = 0) = e^{-\lambda t}\frac{(\lambda t)^n}{n!} \quad (\text{i.e. } N_t \sim P(\lambda t))$$

proof.

Forward Kolmogorov equation: $p_{00}'(t) = \sum_k p_{0k}(t)\, q_{k0}$

$$= -\lambda p_{00}(t) + \lambda p_{01}(t)$$

Backward Kolmogorov equation: $p_{00}'(t) = \sum_k q_{0k}\, p_{k0}(t)$

$$\underbrace{= -\lambda p_{00}(t)}_{(*)} + \cancel{\lambda p_{10}(t)}$$

$p_{ij}(t) = 0 \ \forall j < i$ because in this process we jump always to higher values

From (*):

$\Rightarrow p_{00}(t) = e^{-\lambda t}$, with the condition $p_{00}(0) = 1$

Now we go back by induction.
Suppose that the statement holds for a given n, then

$\mathbb{P}_i(N_t = n+1 \mid N_0 = 0) = p_{0\,n+1}(t)$

$$p_{0\,n+1}'(t) = -\lambda p_{0\,n+1}(t) + \lambda p_{01}(t)$$
$$= -\lambda p_{0\,n+1}(t) + \lambda\frac{(\lambda t)^n}{n!}e^{-\lambda t}$$

solve $(+\, p_{00}(0) = 0)$ and find: $p_{0\,n+1}(t) = \dfrac{(\lambda t)^{n+1}}{(n+1)!}e^{-\lambda t}$ ∎

It's Poisson distribution if we start from 0.
Another property of the Poisson process

Theorem : $(N_t)_{t\geq 0}$ Poisson process $\forall n \ \forall t_1 < t_2 < \ldots < t_n$

$N_{t_1}, N_{t_2} - N_{t_1}, N_{t_3} - N_{t_2}, \ldots, N_{t_k} - N_{t_{k-1}}$

are independent (with respect to $\mathbb{P}_0$) and

$$N_{t_k} - N_{t_{k-1}} \sim P(\lambda(t_k - t_{k-1}))$$

PURE BIRTH PROCESSES

30/10/2019

MC with state space $\mathbb{N} = \{0,1,2,\ldots\}$.
The state is number of individuals

$(n) \xrightarrow{\mathcal{E}(\lambda_n)} (n+1) \quad \lambda_n > 0$

Special cases:

1. $\lambda_n = \lambda > 0 \quad \forall n$ Poisson process
2. $\lambda_n = \lambda(n+1)$ (waiting time for a new arrival $\sim \mathcal{E}(\lambda(n+1))$, expectation is: $\frac{1}{\lambda(n+1)}$)

Remark. If there are n individuals and after a random time $R_i \sim \mathcal{E}(\lambda)$, the individual i gives birth to a new individual is $M = \min_{1\leq i\leq n} R_i$

Lemma: $R_1, \ldots, R_n$ independent, $R_i \sim \mathcal{E}(\theta_i)$
then $M = \min_{1\leq i\leq n} R_i$, $M \sim \mathcal{E}(\theta_1 + \ldots + \theta_n)$

proof.

$$\mathbb{P}(M > t) = \mathbb{P}\left(\cap_{1\leq i\leq n}\{R_i > t\}\right)$$
$$= \prod_{1\leq i\leq n}\mathbb{P}(R_i > t) = \prod_{1\leq i\leq n} e^{-\theta_i t} = e^{-(\sum_{i=1}^n \theta_i)t} \quad ∎$$

3. $\lambda_n = \lambda n$: $Q = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots \\ 0 & -\lambda & \lambda & 0 & \cdots \\ 0 & 0 & -2\lambda & 2\lambda & 0 \cdots \\ & & \cdots & & \end{bmatrix}$ Yule process

$\lambda_n = n\lambda$: one can compute transition probabilities in the case $n = 1,2,3$:

$$P_i(X_t = k) = p_{ik}(t) = \underbrace{\binom{k-1}{k-i} e^{-\lambda i t}(1 - e^{-\lambda t})^{k-i}}_{\text{Pascal distribution with parameters } i \text{ and } e^{-\lambda t}}$$

for $k = i, i+1, \ldots$
$= 0$ for $k < 1$

Exercise : Case 1 $\lambda_n = \lambda$ fixed, compute $P_i(X_t = k)$: $X_t \sim i + N_t$, $N_t \sim P(\lambda t$

PURE DEATH PROCESSES

$(X_t)_{t\geq 0}$, states $\mathbb{N}$, X_t is number of individuals at time t

$(n-1) \xleftarrow{\mathcal{E}(\mu_n)} (n)$: $Q = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & \cdots \\ \mu_1 & -\mu_1 & 0 & 0 & 0 & \cdots \\ 0 & \mu_2 & -\mu_2 & 0 & 0 & 0 \cdots \end{bmatrix}$

Exact formulae for $p_{ij}(t)$ in special cases like:

$\mu_n = \mu > 0$ constant , $\mu_n = n\mu$ linear

BIRTH AND DEATH PROCESSES

$(X_t)_{t\geq 0}$ states $\mathbb{N}$, X_t is number of individuals at time t

$$(n-1) \xleftarrow{\mathcal{E}(\mu_n)} (n) \xrightarrow{\mathcal{E}(\lambda_n)} (n+1) \quad : \quad Q = \begin{bmatrix} -\lambda_n & \lambda_n & 0 & 0 & 0 & \dots \\ \mu_n & -(\lambda_n+\mu_n) & \lambda_n & 0 & 0 & \dots \\ 0 & \mu_n & -(\lambda_n+\mu_n) & \lambda_n & 0 & \dots \\ & & \vdots & & & \end{bmatrix}$$

Consider that all the times are independent , then:
(if not it's not a Markov chain)

- if there are i individuals a death occurs after a random time $\mathcal{E}(\mu_n)$
- if there are i individuals a birth occurs after a random time $\mathcal{E}(\lambda_n)$

You leave the state i after a random time which is the minimum of $B_n \sim \mathcal{E}(\lambda_n)$ and $D_n \sim \mathcal{E}(\mu_n)$
$\Longrightarrow \mathcal{E}(\mu_n + \lambda_n)$.

When leaving i you jump in $j \neq i$ with prob: $\dfrac{q_{ij}}{-q_{ii}}$.

You leave i to go $i+1$ if $B_i < D_i$

If you compute : $\mathbb{P}(B_i < D_i) = \dfrac{\overset{q_{ii+1}}{\lambda_i}}{\underset{-q_{ii}}{\mu_i + \lambda_i}}$

(A pathological example) : Pure birth processes with explosion in finite time :
if transition rates λ_n satisfy $\sum_{n\geq 0} \frac{1}{\lambda_n} < +\infty$

$$(0) \xrightarrow[T_0]{\mathcal{E}(\lambda_0)} (1) \xrightarrow[T_1]{\mathcal{E}(\lambda_1)} (2) \quad \dots \quad (n)$$

occurs a time of : $T_0 + T_1 + \dots + T_{n-1}$: $\mathbb{E}[\dots] = \frac{1}{\lambda_0} + \dots + \frac{1}{\lambda_{n-1}}$

Then in a finite time X_t diverges

CLASSES OF STATES, RECURRENCE AND TRANSIENCE

Communication classes:
j is accessible from i if (discrete: $\exists n : p_{ij}^{(n)} > 0$) :
$\exists t$ s.t. $p_{ij}(t) > 0$

j and i are communicate if each one is accessible from the other.

Classes of states, recurrence and transient are the same def. as in discrete time.

Prop j is accessible from i if and only if $\exists n$ $i_1, \dots, i_n$ such that $i \neq i_1$

i.e. $\dfrac{q_{ii_1}\, q_{i_1 i_2} \cdots q_{i_n j} > 0}{q_{i i_1} > 0, \dots, q_{i_n j} > 0}$

$$p_{ii_1}(t) = q_{ii_1} t + o(t)$$

then $q_{ii_1} > 0 \Longrightarrow p_{ii_1}(t) > 0$ for small t

Communication class are the same as those of the discrete skeleton $\left(\frac{q_{ij}}{-q_{ii}}\right)_{ij}$

Def. A state i is recurrent if :

$$\mathbb{P}_i\left(\omega \mid \{t \geq 0 \mid X_t(\omega) = i\} \text{ is unbounded}\right) = 1$$

(i.e. you visit i for times $t \geq 0$ in an unbounded subset of $[0, +\infty)$ with probability 1)

(we won't say "visit i infinite times" perchè è tempo continuo, se sono in i ci sono già per un periodo di tempo $\longrightarrow$ già infinito)

Def. A state i is transient if :

$$\mathbb{P}_i\left(\omega \mid \{t \geq 0 \mid X_t(\omega) = i\} \text{ is unbounded}\right) = 0$$

there is a theorem that says that this event can be only 1 or 0
(the prob. of this event)

Theorem: The following hold:
1. A state i is recurrent (or transient) for the MC continuous $(X_t)_{t\geq 0}$ $\Longleftrightarrow$ it is recurrent (or transient) for the discrete skeleton
2. Every state is either recurrent or transient and all the states of the same communication class are all recurrent or all transient

Continuous time MC with $Q = (q_{ij})_{ij}$ 12/11/2019

$\rightsquigarrow$ discrete skeleton (discrete time MC) : $P = (p_{ij})_{ij}$

$$p_{ij} = \begin{cases} \frac{q_{ij}}{-q_{ii}} & \text{if } i \neq j,\ q_{ii} \neq 0 \\ 0 & \text{if } i = j,\ q_{ii} \neq 0 \\ 1 & \text{if } i = j,\ q_{ii} = 0 \end{cases}$$

Theorem: i is recurrent (resp transient) for a continuous MC $(X_t)_{t \geq 0}$ if and only if it is recurrent (resp transient) for the discrete skeleton

INVARIANT DENSITIES

$(X_t)_{t\geq 0}$ continuous time MC, $(P_t)_{t \geq 0}$ transition semigroup.
If $X_0 \sim \mu$ i.e. $\mathbb{P}(X_0 = i) = \mu_i \ \forall i$ then:
$\mathbb{P}(X_t = j) = \sum_i \mathbb{P}(X_t = j \mid X_0 = i)\, \mathbb{P}(X_0 = i) = \sum_i p_{ij}(t)\, \mu_i$

Def $(\mu_i)_{i \in E}$ is an invariant density if and only if
$$\mu_j = \sum_{i \in E} \mu_i p_{ij}(t) \quad \forall j\ \forall t$$

Moreover: $\frac{d}{dt}\mu_j = \frac{d}{dt}\sum_{i\in E} \mu_i p_{ij}(t) \overset{\text{under some (regularity)}}{=\!=} \sum_{i \in E} \mu_i \frac{d}{dt} p_{ij}(t)$

at $t = 0$: $\boxed{0 = \sum_{i \in E} \mu_i q_{ij}}$ eq to solve to find the invariant density of a continuous MC (ID)

Prop. Suppose $q_{ii} > -\infty$ and $(p_{ij}(t))_{ij}$ are the unique solutions of FKE and BFE then $(\mu_i)_{i\in E}$ s.t. $\mu_i \geq 0$ and $\sum_{i \in E} \mu_i = 1$ is an invariant density if and only if (ID) holds $\forall j \in E$

proof.

We checked that ID is necessary.
sufficient (for E finite):

$$\frac{d}{dt}\left(\mu_j - \sum_{i \in E} \mu_i p_{ij}(t)\right)$$

$$\text{BKE} \longrightarrow \ = -\sum_{i\in E}\sum_{k \in E} \mu_i q_{ik} p_{kj}(t)$$

$$= -\sum_{k \in E} \underbrace{\left(\sum_{i \in E} \mu_i q_{ik}\right)}_{=0 \text{ by (ID)}} p_{kj}(t) = 0$$

If $\frac{d}{dt}(\cdot) = 0 \Rightarrow (\cdot) = \text{constant}$. So: $t \mapsto \mu_j - \sum_{i \in E} \mu_i p_{ij}(t) = \text{constant}$

$t \mapsto \mu_j - \sum_{i \in E} \mu_i p_{ij}(t) = \mu_j - \sum_{i\in E} \mu_i p_{ij}(0) = \mu_j - \mu_j = 0$

In the discrete time we have: $\pi = \pi P$
continuous : $0 = \mu Q$

Remark: $0 = \mu Q \iff \sum_{i\in E} \mu_i q_{ij} = 0 \quad \forall j$
$\iff \sum_{i \neq j} \mu_i q_{ij} = \mu_j(-q_{jj})$

* If $q_{jj} \neq 0 \iff \mu_j = \sum_{i \neq j} \mu_i \frac{q_{ij}}{-q_{ii}}$

we recognize here the discrete skeleton to the equation to find the invariant densities

Theorem: Invariant densities of a continuous time MC are invariant densities of the discrete skeleton.

Example : Continuous time birth and death chain

$$Q = \begin{bmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & 0 \\ \mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & 0 & 0 \\ 0 & \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 & 0 \end{bmatrix} \quad \lambda_n, \mu_n > 0$$

Irreducible. Invariant densities (if they exist) are unique.
It exists iff $\sum_{n \geq 1} \frac{\lambda_0 \cdots \lambda_{n-1}}{\mu_1 \cdots \mu_n} < +\infty$ and is given by

$\pi_0 = \frac{1}{Z}$, $\pi_n = \frac{\lambda_0 \cdots \lambda_{n-1}}{\mu_1 \cdots \mu_n} \cdot \frac{1}{Z}$ for $n > 0$

where $Z = 1 + \sum_{n \geq 1} \frac{\lambda_0 \cdots \lambda_{n-1}}{\mu_1 \cdots \mu_n}$ is the normalization constant

$$0 = \pi Q : \begin{cases} -\pi_0 \lambda_0 + \pi_1 \mu_1 = 0 \\ \pi_{n-1}\lambda_{n-1} - (\lambda_n + \mu_n)\pi_n + \pi_{n+1}\mu_{n+1} = 0 \end{cases}$$

$$\Rightarrow \pi_1 = \frac{\lambda_0}{\mu_1} \pi_0$$

Inductively: $\pi_n = \frac{\lambda_0 \cdots \lambda_{n-1}}{\mu_1 \cdots \mu_n} \pi_0$

Turns out that $\sum_{n \geq 0} \pi_n = 1 \iff \pi_0 \Big(1 + \underbrace{\sum_{n\geq 1} \frac{\lambda_0 \cdots \lambda_{n-1}}{\mu_1 \cdots \mu_n}}\Big) = 1$

if the series is divergent $\rightarrow \nexists$ solution
if the series is convergent $\rightarrow \exists!$ solution

Theorem (Ergodic theorem)

In the case of discrete MC irreducible positive recurrent we have:

$$\frac{1}{n}\sum_{k=1}^{n} \chi_{\{X_k = i\}} \longrightarrow \pi_i \quad \text{invariant density}$$

We have the same result for continuous positive recurrent MC:

$$\frac{1}{t}\int_0^t \chi_{\{X_s = i\}}\, ds \longrightarrow \pi_i \quad \text{invariant density}$$

17

M/M/K QUEUE MODEL

(another birth and death model)

- There are **k** counters (servers).
- Service times at any counter are $\sim \mathcal{E}(\mu)$
- The times between the arrivals of two consecutive customers in the system is $\sim \mathcal{E}(\lambda)$
- All the exponentionals random variables are independent

M / M / k

exponential interarrivals times — exponential service times — number of counters

1. Q? (for X_t = number of customers in the system at time t)

N_t = number of customers arriving in the system from time 0 to time t

N_t is a Poisson process : $N_t \sim P(\lambda t)$

We look for :

$n \to n+1$ after $\mathcal{E}(\lambda)$

$n \to n-1$ $\begin{cases} \text{if } n \geq k \text{ after min of } (k \text{ indep. } \mathcal{E}(\mu)) = \mathcal{E}(k\mu) \\ \text{if } n \leq k \text{ after min of } (n \text{ indep } \mathcal{E}(\mu)) = \mathcal{E}(n\mu) \end{cases}$

You leave **n** after $\underbrace{\min\{\mathcal{E}(\lambda), \mathcal{E}(\mu \cdot \min\{n,k\})\}}$

$$\equiv \mathcal{E}(\lambda + \mu \min\{n,k\})$$

$$\longrightarrow q_{nn} = -(\lambda + \mu \min\{k,n\})$$

$$\underbrace{\frac{q_{n\,n+1}}{-q_{nn}}} = \mathbb{P}(A<D) = \frac{\lambda}{\lambda + \mu \cdot \min\{k,n\}}$$

$\hookrightarrow \begin{cases} A \sim \mathcal{E}(\lambda) \\ D \sim \mathcal{E}(\lambda + \mu \cdot \min\{n,k\}) \end{cases} \perp\!\!\!\perp$

probabilità che quando si lascia n si va in $n+1$

$$\frac{q_{n\,n-1}}{-q_{nn}} = \frac{\mu \cdot \min\{n,k\}}{\lambda + \mu \cdot \min\{n,k\}} \Longrightarrow q_{n\,n-1} = \mu \min\{n,k\}$$

$$\longrightarrow Q = \begin{bmatrix} -\lambda & \lambda & 0 & 0 & 0 \\ \mu & -(\lambda+\mu) & \lambda & 0 & 0 \\ 0 & 2\mu & -(\lambda+2\mu) & \lambda & 0 \; \cdots \\ & \ddots & & & \end{bmatrix} \quad \lambda, \mu > 0$$

2. **Q** is irreducible (since $\mu, \lambda > 0$ (since they're $\mathcal{E}(\cdot\cdot)$ parameters))

Is it transient or recurrent?

We look at the discrete skeleton : $p_{n\,n+1} = \dfrac{\lambda}{\lambda + \mu \cdot \min\{n,k\}}$

$$p_{n\,n-1} = \frac{\mu \cdot \min\{n,k\}}{\lambda + \mu \cdot \min\{n,k\}}$$

Under what condition is transient or recurrent?

We have basically a random walk : $\overset{\mu k}{\leftarrow} \bigcirc \overset{\lambda}{\rightarrow}$

So we have:
- transient if $\lambda > \mu k$ $\longrightarrow$ the queue becomes longer and longer
- recurrent if $\lambda \leq \mu k$ $\longrightarrow$ if it's "=" it's recurrent but not fast recurrent ($\nexists$ stationary distribution); if "<" $\Rightarrow \exists$ stationary distribution

***** (Precisazione (correzione)) 13/11/2019

Relationship between invariant densities of a continuous time MC and its discrete skeleton:

- continuous time MC with transition matrix Q and invariant density $\pi Q = 0$: $(\pi Q)_j = 0 \;\; \forall j$

$$\Rightarrow \sum_{i \in E} \pi_i q_{ij} = 0 \iff \pi_j(-q_{jj}) = \sum_{\substack{i \in E \\ i \neq j}} \pi_i q_{ij}$$

If $q_{ii} \neq 0$, $q_{jj} \neq 0$:

$$\pi_j(-q_{jj}) = \sum_{i \in E} (\pi_i(-q_{ii})) \frac{q_{ij}}{-q_{ii}}$$

transient matrix of discrete skeleton $\longrightarrow \hat{p}_{ij}$

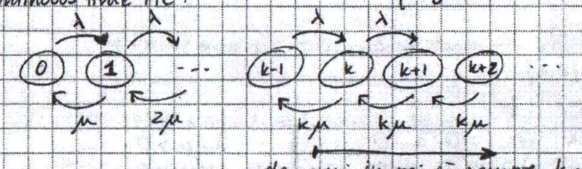
$$= \mu_j = \sum_{i \neq j} \mu_i \hat{p}_{ij}$$

$(\pi_j(-q_{jj}))_{j \in E}$ (possibly normalized) is an invariant density for the discrete skeleton

M/M/K QUEUE (pt. II)

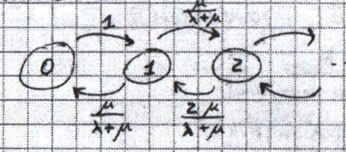
$Q = (q_{mn})_{m,n \geq 0}$: $q_{mn} = \begin{cases} \lambda & \text{if } n = m+1 \\ \mu \min\{k,m\} & n = m-1 \\ -(\lambda + \mu \cdot \min\{k,m\}) & n = m \\ 0 & \text{otherwise} \end{cases}$

Continuous time MC :



da qui in poi è sempre $k\mu$ perché ci sono **k** counters

Discrete skeleton :



$p_{mn} = \begin{cases} \dfrac{\lambda}{\lambda + \mu \cdot \min\{k,m\}} & \text{if } n = m+1 \\ \dfrac{\mu \cdot \min\{k,m\}}{\lambda + \mu \cdot \min\{k,m\}} & n = m-1 \\ 0 & \text{otherwise} \end{cases}$

So, for $m > k$:

$$\hat{p}_{mn} = \begin{cases} \frac{\lambda}{\lambda + k\mu} & n = m+1 \\ \frac{k\mu}{\lambda + k\mu} & n = m-1 \\ 0 & \text{otherwise} \end{cases}$$

this is a random walk on the natural numbers. We have a matrix that looks like:

$$\begin{pmatrix} 0 & 1 & 0 & 0 & \cdots \\ \cdots & & & & \\ \cdots & \alpha & 0 & 1-\alpha & \cdots \end{pmatrix}$$

$$\Longrightarrow \hat{P} = \begin{pmatrix} & \cdots & & \\ \cdots & \alpha & 0 & 1-\alpha & \cdots \\ & \cdots & & \end{pmatrix} : \quad \alpha > 1-\alpha \text{ iff } \alpha > \frac{1}{2} \iff \frac{k\mu}{\lambda + k\mu} > \frac{1}{2} \iff \lambda < k\mu$$

It is recurrent if $\lambda \le k\mu$.

3. Invariant densities?

$$Q = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & \cdots \\ \mu_1 & -(\mu_1+\lambda_1) & \lambda_1 & 0 & \\ & \cdots & & & \end{pmatrix}$$

, the invariant density exists iff:

$$\sum_{n \ge 1} \frac{\lambda_0 \cdots \lambda_{n-1}}{\mu_1 \cdots \mu_n} < +\infty$$

For $n > k$:

$$\frac{\lambda_0 \cdots \lambda_{n-1}}{\mu_1 \cdots \mu_n} = \frac{\lambda^n}{\underbrace{\mu \cdot (2\mu) \cdot (3\mu) \cdots}_{(k-1) \text{ factors}} \underbrace{(k\mu)(k\mu)(k\mu) \cdots}_{(n-(k-1)) \text{ factors}}} = \frac{\lambda^n}{\mu^n \, k! \, k^{n-k}}$$

$$\Longrightarrow \sum_{n \ge 1} \frac{\lambda_0 \cdots \lambda_{n-1}}{\mu_1 \cdots \mu_n} \rightsquigarrow \sum_{n=1}^{k} (\ldots) + \sum_{n > k} \left(\frac{\lambda}{\mu k}\right)^n \frac{k^k}{k!}$$

convergent $\iff \lambda < k\mu$

$\Longrightarrow$ Defining $Z_k = 1 + \sum_{n=1}^{k} \frac{\lambda^n}{n! \mu^n} + \frac{k^k}{k!} \sum_{n > k} \left(\frac{\lambda}{\mu k}\right)^n$ normalization constant

the invariant density is:

$$\pi_n = \begin{cases} \frac{1}{Z_k} \frac{\lambda^n}{\mu^n n!} & k \le n \\ \frac{1}{Z_k} \frac{\lambda^n}{\mu^n k! \, k^{n-k}} & k > n \end{cases}$$

Remark: for $n > k$, defining $\nu = \frac{\lambda}{\mu}$: $\pi_n = \left(\frac{\nu}{k}\right)^{n-k} \pi_k$

4. Average length of the queue in stationary conditions?

$$\mathbb{E}_\pi[X_t] = \sum_{n \ge 1} n \pi_n = [\ldots] = \nu + \frac{\pi_k \, \nu/k}{\left(1 - \frac{\nu}{k}\right)^2}$$

(π: because we consider stationary conditions)

assuming that $X_t \sim \pi$

5. Mean time spent in the system? (Again in stationary conditions)

$$\underbrace{W_t}_{\text{time spent in the system}} = \underbrace{S}_{\text{service time}} + \underbrace{\sum_{n \ge k} \chi_{\{X_t = n\}} \circledast}_{\text{waiting time in the queue}}$$

$\circledast$ = random variable which is the sum of $(n-k+1)$ independent random variables with $\mathcal{E}(k\mu)$

(arrivo che ci sono $n > k$ clienti: devo aspettare $n-k$ che finiscano e almeno un altro posto libero:

$$\underbrace{\circ \quad \circ \quad \circ \quad \circ}_{\text{gente dai counters } (k)} \quad \underbrace{\bullet \quad \bullet}_{n-k} \quad \overset{io}{\bullet}$$

)

$$Y_{n-k+1} \sim \Gamma(n-k+1, k\mu)$$

(since: $V_1, \ldots, V_n \sim \mathcal{E}(\theta) \longrightarrow V_1 + \ldots + V_n \sim \Gamma(n, \theta)$)

$$\Longrightarrow W_t = S + \sum_{n \ge k} Y_{n-k+1} \, \chi_{\{X_t = n\}}$$

Reminder: $V \sim \Gamma(m, \theta) \Longrightarrow \mathbb{E}[V] = \frac{m}{\theta}$

$\Longrightarrow$ the density of Y_{n-k+1} w.r.t. $\mathbb{P}(\cdot \mid X_t = n)$ is $\Gamma(n-k+1, k\mu)$

$$\Longrightarrow \mathbb{E}_\pi[W_t] = \mathbb{E}_\pi[S] + \sum_{n \ge k} \mathbb{E}_\pi\left[Y_{n-k+1} \, \chi_{\{X_t = n\}}\right]$$

$$= \frac{1}{\mu} + \sum_{n \ge k} \underbrace{\mathbb{E}_\pi\left[Y_{n-k+1} \mid \chi_{\{X_t = n\}}\right]}_{\frac{n-k+1}{k\mu}} \underbrace{\mathbb{P}_\pi(X_t = n)}_{\pi_n = \left(\frac{\nu}{k}\right)^{n-k} \pi_k \text{ for } n > k}$$

$$= \frac{1}{\mu} + \sum_{n \ge k} \frac{n-k+1}{k\mu} \left(\frac{\nu}{k}\right)^{n-k} \pi_k$$

$$= \frac{1}{\mu} + \frac{\pi_k}{k\mu} \sum_{n \ge k} (n-k+1) \left(\frac{\nu}{k}\right)^{n-k}$$

$$= \frac{1}{\mu} + \frac{\pi_k}{k\mu} \sum_{m \ge 0} (m+1) \left(\frac{\nu}{k}\right)^m$$

Reminder:

$$\sum_{m \ge 0} (m+1) \theta^m = \sum_{m \ge 0} \frac{d}{d\theta} \left(\theta^{m+1}\right) = \frac{d}{d\theta} \sum_{m \ge 0} \theta^{m+1} = \frac{d}{d\theta} \left(\frac{\theta}{1-\theta}\right) = \frac{1}{(1-\theta)^2}$$

$$\Longrightarrow \mathbb{E}_\pi[W_t] = \frac{1}{\mu} + \frac{\pi_k}{k\mu} \cdot \frac{1}{\left(1 - \frac{\nu}{k}\right)^2}$$

Average length of the queue: $\nu + \dfrac{\pi_k \frac{\nu}{k}}{(1-\frac{\nu}{k})^2} = \lambda\left(\dfrac{1}{\mu} + \dfrac{\pi_k}{k\mu}\right), \quad \nu = \dfrac{\lambda}{\mu}$

$= \lambda \cdot \left(\text{Average waiting time}\right)$

End of cont. MC discrete state

(CURIOSITA')

Modelling epidemy:

Population of N individuals.
There is a disease:

- S = number of susceptible individuals
- I = infected individuals

Evolution:
$$\begin{cases} \dfrac{dS}{dt} = -\beta SI + \gamma I \\ \dfrac{dI}{dt} = -\gamma I + \beta SI \end{cases}$$

the sum equals 0 since we have the conservation law: $S+I=N$

γ: rate of recovery

β: rate of contact between infected and susceptible

Can we built a stochastic model? contacts and getting infected are random (variables)

19/11/2019

DETERMINISTIC SIR MODEL

Population of N individuals

- S = number of susceptible individuals
- I = number of infected individuals
- R = number of removed (recovered) individuals: they already got the disease, they recovered and they can't get the disease anymore

(Actually: S_t, I_t, R_t)

The model is:
$$\begin{cases} S'_t = \dfrac{dS_t}{dt} = -\dfrac{\beta S_t I_t}{N} + \gamma R_t \\ I'_t = \dfrac{dI_t}{dt} = \dfrac{\beta S_t I_t}{N} - \nu I_t \\ R'_t = \dfrac{dR_t}{dt} = \nu I_t - \gamma R_t \end{cases}$$

Note that $S'_t + I'_t + R'_t = 0 \quad (\Longleftrightarrow \quad S_t + R_t + I_t = const = N)$

How we construct a Markov Chain model out of this model?

We consider a simplified model: **SIS MODEL** (deterministic)

The equations of the model are the following:
$$\begin{cases} S'_t = -\dfrac{\beta I_t S_t}{N} + \nu I_t \\ I'_t = \dfrac{\beta I_t S_t}{N} - \nu I_t \end{cases} \qquad S'_t + I'_t = 0, \quad S_t + I_t = N, \quad 0 \le S_t, I_t \le N$$

$$\longrightarrow \quad S'_t = I_t\left(-\frac{\beta S_t}{N} + \nu\right): \begin{cases} \ge 0 \iff S_t \le \frac{N\nu}{\beta} \\ < 0 \iff S_t > \frac{N\nu}{\beta} \end{cases}$$

conditions for having an exploding epidemy or not

How we construct the **Markov Chain model**? (Also for SRI model)

We start with the continuous time Markov model.
(always under the condition $S_t + I_t = N$)

MC $(S_t)_{t\ge 0}$ with state space $\{0,\dots,N\}$. Q?

$q_{s,k}$ ($0<s<N$) transition $s \to k$
$$q_{s,k} = \begin{cases} \nu(N-s) & \text{if } k = s+1 \\ \dfrac{\beta s(N-s)}{N} & \text{if } k = s-1 \\ -\nu(N-s) - \dfrac{\beta s(N-s)}{N} & \text{if } k = s \\ 0 & \text{otherwise} \end{cases}$$

$q_{N,N} = 0, \dots, q_{Ns} = 0 \quad \forall s$

How can we construct a discrete time model?
We can introduce small time intervals:

$$\Delta t \text{ small}: \quad p_{s,k} = \begin{cases} q_{s,k}\,\Delta t & \text{for } k \ne s \\ 1 - q_{s,k}\,\Delta t & \text{for } k = s \end{cases}$$

We consider now a general MC (continuous or discrete) with an uncountable state of spaces (generally a subspace of $\mathbb{R}$).
But before:

CONDITIONAL EXPECTATION

$(\Omega, \mathcal{F}, P)$ probability space, $\mathcal{G} \subseteq \mathcal{F}$ sub-σ-algebra
X real random variable with $\mathbb{E}[|X|] < \infty$

Def. We call conditional expectation of X w.r.t. $\mathcal{G}$ any $\mathcal{G}$-measurable random variable Y such that:
$$\int_G X\,dP = \int_G Y\,dP \quad \forall G \in \mathcal{G} \qquad (*)$$

(X is σ-measurable $\rightsquigarrow$ Y is $\mathcal{G}$-measurable)

Remark. If $\mathcal{G} = \{\emptyset, \Omega\}$, Y $\mathcal{G}$-measurable $\xrightarrow{\text{means}}$ $\{Y \in A\} = \begin{cases} \emptyset \\ \Omega \end{cases}$

In particular: $\{Y = y\} = \begin{cases} \emptyset \\ \Omega \end{cases} \mid \forall y \implies \exists !\, y$ s.t. $Y = y$

$\implies Y$ constant.

If Y is constant $\implies$ (*) is fulfilled with $Y = \mathbb{E}[Y]$

Then $\mathcal{G} = \{\emptyset, \Omega\}$, $Y = \mathbb{E}[Y]$

Example: $X_1, \dots, X_n \sim B(1, \theta)$, $X_1, \dots, X_n$ iid: $\mathbb{P}(X_i = 1) = \theta$, $\mathbb{P}(X_i = 0) = 1 - \theta$

$S_n := X_1 + \dots + X_n$

$\mathcal{G} = \underset{\substack{\sigma\text{-alg constructed}\\ \text{by } S_n}}{\sigma(S_n)} = \sigma(\{S_n = k \mid k = 0, \dots, n\})$

$= \{\{S_n \in F\} \mid F \subseteq \{0, \dots, n\}\}$

$\mathbb{E}[X_i \mid \sigma(S_n)] = \frac{S_n}{n} \quad (\forall \theta) \quad \left(= \mathbb{E}[X_i \mid S_n = k] = \frac{k}{n}\right)$

One can check that: $\boxed{\int_G X_i \, d\mathbb{P} = \int_G \frac{S_n}{n} \, d\mathbb{P} \quad \forall G \in \sigma(S_n)}$ $\{\{S_n \in F\} \mid F \subseteq \{0..n\}\}$ — important property

(check:)

It is sufficient to check that: $\int_{\{S_n = k\}} X_i \, d\mathbb{P} = \int_{\{S_n = k\}} \frac{S_n}{n} \, d\mathbb{P} \quad \forall k$

$= \frac{k}{n} \mathbb{P}(S_n = k)$

$\rightarrow \underline{\mathbb{P}(S_n = k, X_i = 1)} = \mathbb{P}(X_i = 1, S_n - X_i = k - 1)$

$= \mathbb{P}(X_i = 1)\, \mathbb{P}(\underbrace{S_n - X_i}_{S_n - X_i \sim B(n-1, \theta)} = k - 1)$ by independence

$=$

$= \theta \binom{n-1}{k-1} \theta^{k-1} (1-\theta)^{n-k}$

On the other hand: $\underline{\frac{k}{n} \mathbb{P}(S_n = k)} = \frac{k}{n} \binom{n}{k} \theta^k (1-\theta)^{n-k}$

$= [\,\cdot\,] = \binom{n-1}{k-1} \theta^k (1-\theta)^{n-k}$ ■

$\left(\mathbb{E}[X \mid Y] = \int x\, f_{X|Y}(x|y)\, dy \right)$

Why do we care? In the MC models, by using transition matrixes we find conditional expectations.

Theorem: For all X $\mathcal{F}$-measurable (and integrable) $\mathbb{E}[|X|] < \infty$, $\mathcal{G} \subseteq \mathcal{F}$, there exists Y $\mathcal{G}$-measurable such that:

$$\int_G X \, d\mathbb{P} = \int_G Y \, d\mathbb{P} \quad \forall G \in \mathcal{G}$$

It is unique, up to sets $G \in \mathcal{G}$ of probability $\mathbb{P}(G) = 0$ and is denoted by: $\mathbb{E}[X \mid \mathcal{G}]$

proof.

Suppose $X \geq 0$ (if not $X = X_+ - X_-$)

Consider the measure $\mathbb{Q}$ on $\mathcal{G}$ defined by: $\mathbb{Q}(G) = \int_G X \, d\mathbb{P}$.

$\mathbb{Q} \ll \mathbb{P}|_{\mathcal{G}}$

($\mathbb{Q}$ is absolutely continuous wrt $\mathbb{P}$ restricted to $\mathcal{G}$)

i.e. $\mathbb{P}(G) = 0 \implies \mathbb{Q}(G) = 0$

By the Radon-Nikodyn theorem $\exists Y \geq 0$ $\mathcal{G}$-measurable s.t.

$\mathbb{Q}(G) = \int_G Y \, d\mathbb{P}$

$(= \int_G X \, d\mathbb{P})$

If X is not positive $\rightarrow$ $X = X_+ - X_-$, find Y^+ and Y^- and $\mathbb{E}[X \mid \mathcal{G}] = Y^+ - Y^-$

(but we do not insist on this part, either on the uniqueness) ■

Some properties: (X_1, X_2 random variables s.t. $\mathbb{E}[|X_1|] < \infty$, $\mathbb{E}[|X_2|] < \infty$)

1. Linearity: $\mathbb{E}[aX_1 + bX_2 \mid \mathcal{G}] = a\,\mathbb{E}[X_1 \mid \mathcal{G}] + b\,\mathbb{E}[X_2 \mid \mathcal{G}]$
2. Positivity: $X_1 \geq 0 \implies \mathbb{E}[X_1 \mid \mathcal{G}] \geq 0$
3. Normalization: $X_1 = $ constant, $X_1 = 1 \implies \mathbb{E}[X_1 \mid \mathcal{G}] = 1$

proof. (1.)

$\int_G \mathbb{E}[aX_1 + bX_2 \mid \mathcal{G}] \, d\mathbb{P} \overset{\text{by def.}}{=} \int_G (aX_1 + bX_2) \, d\mathbb{P}$

$(G \in \mathcal{G})$

$= a \int_G X_1 \, d\mathbb{P} + b \int_G X_2 \, d\mathbb{P}$

$= a \int_G \mathbb{E}[X_1 \mid \mathcal{G}] \, d\mathbb{P} + b \int_G \mathbb{E}[X_2 \mid \mathcal{G}] \, d\mathbb{P}$

$= \int_G \left(a\,\mathbb{E}[X_1 \mid \mathcal{G}] + b\,\mathbb{E}[X_2 \mid \mathcal{G}] \right) d\mathbb{P}$ ■

Projective Property (most important)

If $\mathcal{H} \subseteq \mathcal{G}$ then:

$$\mathbb{E}\big[\mathbb{E}[X \mid \mathcal{G}] \mid \mathcal{H}\big] = \mathbb{E}[X \mid \mathcal{H}]$$

("proiezioni" perché se proiettiamo X su $\mathcal{G}$ e poi proiettiamo quello che abbiamo ottenuto su $\mathcal{H}$ (sottoinsieme di $\mathcal{G}$) è uguale a proiettare subito su $\mathcal{H}$)

proof.

$H \in \mathcal{H}$: $\int_H \mathbb{E}\big[\mathbb{E}[X \mid \mathcal{G}] \mid \mathcal{H}\big] \, d\mathbb{P} = \int_H \mathbb{E}[X \mid \mathcal{G}] \, d\mathbb{P}$ — because $H \in \mathcal{H}$

$= \int_H X \, d\mathbb{P} = (\ast)$ — because $H \in \mathcal{H} \subseteq \mathcal{G} \rightarrow H \in \mathcal{G}$

$\rightarrow$ by def of $\mathbb{E}[\,\cdot \mid \mathcal{H}]$:

$(\ast) = \int_H \mathbb{E}[X \mid \mathcal{H}] \, d\mathbb{P}$ ■

If X, Y have joint density $f_{X,Y}$ and conditional density $f_{X|Y}$

$\longrightarrow$ $\mathbb{E}[X \mid \sigma(Y)]$ is the r.v. $\omega \mapsto \int x \, f_{X|Y}(x \mid \underbrace{Y(\omega)}_{y}) \, dx$

(notation : $\mathbb{E}[X \mid \underbrace{Y(\omega)}_{Y=y}]$)

20/11/2019

$(X_n)_{n \geq 0}$ discrete time MC , P transition matrix

$f : \underset{\text{set of states}}{E} \longrightarrow \mathbb{R}$ s.t. $\mathbb{E}[|f(X_n)|] < +\infty \quad \forall n$

! $\boxed{\mathbb{E}[f(X_{n+1}) \mid \sigma(X_n, \dots, X_0)] = Pf(X_n)}$ ← this sort of came from the Markov property (for the proof we use the M prop)

where $\boxed{(Pf)(j) = \sum_{k \in E} p_{jk} \, f(k)} \quad \forall j \in E$

Remark. $f = \chi_A \quad A \subseteq E$: $\mathbb{E}[\chi_A(X_{n+1}) \mid \dots] = (P\chi_A)(X_n)$

$= \mathbb{P}(X_{n+1} \in A \mid \sigma(X_n, \dots, X_0))$

$= \sum_{k \in E} p_{X_n k} \, \chi_A(k)$

$= \sum_{k \in A} \mathbb{P}(X_{n+1} = k \mid X_n = \cdot)$

$= \mathbb{P}(X_{n+1} \in A \mid X_n = \cdot)$

the probability given all the past or given just the last one is the same (Markov property)

What happen when we have more than $n+1$?

$\mathbb{E}[f(X_{n+2}) \mid \sigma(X_n, \dots, X_0)] = \mathbb{E}\Big[\mathbb{E}[f(X_{n+2}) \mid \underbrace{\sigma(X_{n+1}, \dots, X_0)}_{\text{bigger } \sigma\text{-alg of}}] \mid \sigma(X_n, \dots, X_0)\Big]$

projective property

$= \mathbb{E}[Pf(X_{n+1}) \mid \sigma(X_n, \dots, X_0)]$

$= (PPf)(X_n) = P^2 f(X_n)$

$\rightarrow \boxed{\mathbb{E}[f(X_{n+m}) \mid \sigma(X_n, \dots, X_0)] = P^m f(X_n)}$

$(X_t)_{t \geq 0}$ continuous time MC, $(P_t)_{t \geq 0}$ transition semigroup

$f : E \rightarrow \mathbb{R}$ s.t. $\mathbb{E}[|f(X_t)|] < +\infty \quad \forall t$

! $\boxed{\mathbb{E}[f(X_t) \mid \sigma(X_r \mid r \leq s)] = (P_{t-s} f)(X_s)} \quad s < t$

proof. (Formula in the continuous case)

We must check that $\forall G \in \sigma(X_r \mid r \leq s)$: $\int_G f(X_t) \, d\mathbb{P} = \int_G (P_{t-s} f)(X_s) \, d\mathbb{P}$

It is sufficient to check it for G in a set of generators of $\sigma(X_r \mid r \leq s)$ such as: $\{X_s = j, X_{s_n} = j_n, \dots, X_0 = j_0\}$ $\quad 0 < s_1 < \dots < s_n < s$

For these events one can do an explicit computation:

$\int_{\{X_s = j, X_{s_n} = j_n, \dots, X_0 = j_0\}} f(X_t) \, d\mathbb{P} = \sum_{k \in E} f(k) \, \mathbb{P}(X_t = k, X_s = j, X_{s_n} = j_n, \dots)$

$= \sum_{k \in E} f(k) \, \mathbb{P}(X_t = k \mid X_s = j \, \cancel{\dots}) \, \mathbb{P}(X_s = j, \dots)$

by Markov property

$= \sum_{k \in E} f(k) \, p_{jk}(t-s) \int_{\{X_s = j, X_{s_n} = j_n, \dots\}} d\mathbb{P}$

$= \int_{\{X_s = j, X_{s_n} = j_n, \dots\}} (P_{t-s} f)(j) \, d\mathbb{P}$

↑ $j = X_s$ because

■

Another **property** (useful in computation):

X random variable $\mathcal{F}$-measurable , $\mathbb{E}[|X|^p] < +\infty$

Y random variable $\mathcal{G}$-measurable , $\mathbb{E}[|Y|^q] < +\infty$ $\quad \Big\} \ \frac{1}{p} + \frac{1}{q} = 1$

Then: $\boxed{\mathbb{E}[XY \mid \mathcal{G}] = \mathbb{E}[X \mid \mathcal{G}] \cdot Y}$ **(E1)** $\quad (\mathcal{G} \subseteq \mathcal{F})$

proof. (come in ARF (in generale nella measure theory) : indicatrici $\longrightarrow$ simple functions $\Longrightarrow$ generic functions)

Suppose $X, Y \geq 0$ (if not $X = X^+ - X^-$, $Y = Y^+ - Y^-$, ...)

(1) (E1) is true when $Y = \chi_G \quad G \in \mathcal{G}$: $(G' \in \mathcal{G}$: $)$

$\int_{G'} \mathbb{E}[X \chi_G \mid \mathcal{G}] \, d\mathbb{P} = \int_{G'} \chi_G \, X \, d\mathbb{P}$

$= \int_{\underbrace{G' \cap G}_{\in \mathcal{G}}} X \, d\mathbb{P}$

$= \int_{G' \cap G} \mathbb{E}[X \mid \mathcal{G}] \, d\mathbb{P}$

$= \int_G \underbrace{\chi_{G'} \, \mathbb{E}[X \mid \mathcal{G}]}_{\mathcal{G}\text{-measurable}} \, d\mathbb{P}$

$\rightarrow \mathbb{E}[\chi_G X \mid \mathcal{G}] = \chi_G \, \mathbb{E}[X \mid \mathcal{G}]$

(2) (E1) is true for $Y = \sum_j a_j \chi_{G_j}$ simple $\quad (G_j \in \mathcal{G})$

(3) Approx Y by $Y_n \uparrow Y$ $\quad Y_n$ simple functions ■

Remark: X $\mathcal{F}$-measurable, $\mathcal{G} \subseteq \mathcal{F}$. Suppose that X and $\mathcal{G}$ are independent (a random variable and a σ-algebra), i.e.
$\mathbb{P}(\{X\in A\} \cap G) = \mathbb{P}(X\in A)\cdot \mathbb{P}(G)$
(in another way: $\underbrace{\sigma(X)}_{\sigma \text{ generated by } X}$ and $\mathcal{G}$ are independent).

$$\Longrightarrow \mathbb{E}[X \mid \mathcal{G}] = \mathbb{E}[X]$$

because: $\int_G X \, d\mathbb{P} = \mathbb{E}[X]\,\mathbb{P}(G)$ by independence
$= \int_G \mathbb{E}[X]\, d\mathbb{P}$

Another **property**: X with $\mathbb{E}[|X|] < \infty \Longrightarrow \mathbb{E}[|\mathbb{E}[X|\mathcal{G}]|] < +\infty$

Also, $\forall p \in [1,+\infty)$: $\mathbb{E}[|X|^p] < +\infty \Longrightarrow \mathbb{E}[|\mathbb{E}[X|\mathcal{G}]|^p] \leq \mathbb{E}[|X|^p]$ ⊛

By Jensen: (Jensen inequality) $\varphi: \mathbb{R}\to\mathbb{R}$ convex, $\mathbb{E}[X] < +\infty$, $\mathbb{E}[|\varphi(X)|] < +\infty$
$\longrightarrow \varphi(\mathbb{E}[X|\mathcal{G}]) \leq \mathbb{E}[\varphi(X)|\mathcal{G}]$

We use the Jensen inequality for ⊛ with $\varphi(x) = |x|^p$

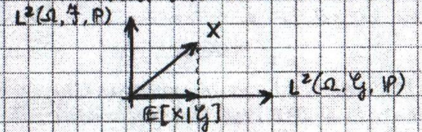
Theorem: If $\mathbb{E}[|X|^2] < +\infty$, $\mathcal{G}\subseteq\mathcal{F}$ then $\mathbb{E}[|\mathbb{E}[X|\mathcal{G}]|^2] < +\infty$ and

$$\min_{\substack{Y\ \mathcal{G}\text{-measurable} \\ \text{s.t. } \mathbb{E}[|Y|^2]<\infty}} \mathbb{E}[|X-Y|^2] = \mathbb{E}[|X - \mathbb{E}[X|\mathcal{G}]|^2]$$

(X – $\mathcal{F}$-measurable: the minimum square error we do by approximating X $\mathcal{F}$-measurable with Y $\mathcal{G}$-measurable is given by ↑)

i.e. the conditional expectation is the MSE (min square error) approx of X with a $\mathcal{G}$-measurable random variable

We know that: $\mathcal{G}\subseteq\mathcal{F} \Longrightarrow L^2(\Omega,\mathcal{G},\mathbb{P}) \subseteq L^2(\Omega,\mathcal{F},\mathbb{P})$
this theorem can be viewed as an orthogonal projection:



proof.
Let $V := \mathbb{E}[X|\mathcal{G}]$:

$\mathbb{E}[|X-Y|^2] = \mathbb{E}[|(X-V)+(V-Y)|^2] = \mathbb{E}[$
$= \mathbb{E}[|X-V|^2] + \mathbb{E}[|V-Y|^2] + 2\mathbb{E}[(X-V)(V-Y)]$

$\mathbb{E}[(X-V)(V-Y)] = \mathbb{E}[\mathbb{E}[(X-V)\underbrace{(V-Y)}_{\mathcal{G}\text{-measurable}}|\mathcal{G}]]$
$= \mathbb{E}[(V-Y)\,\mathbb{E}[X-V|\mathcal{G}]]$

but $\mathbb{E}[X-V|\mathcal{G}] = 0$ because $\mathbb{E}[X|\mathcal{G}] - V = 0$
$\Longrightarrow$ the double product $2\mathbb{E}[(X-V)(V-Y)] = 0$
$\Longrightarrow \mathbb{E}[|X-Y|^2] = \mathbb{E}[|X-V|^2] + \underbrace{\mathbb{E}[|V-Y|^2]}_{\substack{\text{minimized by } V=Y \\ (Y\ \mathcal{G}\text{-measurable})}}$ ■

Lemma: (usefull for computation)
X random variable
$\mathcal{G}$ independent of X
X has density f
$Y_1,\dots,Y_n$ $\mathcal{G}$-measurable
$\Longrightarrow \mathbb{E}[h(X,Y_1,\dots,Y_n)|\mathcal{G}] = \int_{\mathbb{R}} h(x,Y_1,\dots,Y_n)\, f(x)\, dx$

26/11/2019

MARKOV PROCESSES (arbitrary set of values, discrete or continuous time)

$(\Omega,\mathcal{F},\mathbb{P})$ probability space
$(E,\mathcal{E})$ measurable space (typical choices $E\subseteq\mathbb{N}^d$, $E\subseteq\mathbb{Z}^d$ with $\mathcal{E} = \mathcal{P}(E)$, $E\subseteq\mathbb{R}^d$ Borel $\mathcal{E} = \mathcal{B}(E)$)
$t \in$ set of times (typical choices: $\subseteq\mathbb{N}$ or interval $\subseteq\mathbb{R}$)

Def. $(X_t)_t$ collection of random variables $X_t:\Omega\to E$ is a Markov process if $\forall m$ $\forall t_1 < t_2 < \dots < t_m$, $\forall E_1,\dots,E_{m+1}\in\mathcal{E}$:
$\mathbb{P}(X_{t_{m+1}}\in E_{m+1} \mid X_{t_m}\in E_m,\dots,X_{t_1}\in E_1) = \mathbb{P}(X_{t_{m+1}}\in E_{m+1}\mid X_{t_m}\in E_m)$
(Note: (notation) if E is discrete $\longrightarrow$ it's a Markov chain)

(Def.)
The Markov process is time homogeneous if:
$\mathbb{P}(X_{t+s}\in E_2 \mid X_s\in E_1)$ does not depend on s $\forall t>0$

TRANSITION KERNELS (heuristic definition):
$$P_t(x,A) = \mathbb{P}(X_t\in A\mid X_0 = x)$$

(Formal definition): A transition kernel is a collection of functions $(P_t)_t$:
$P_t: E\times\mathcal{E}\to[0,1]$ s.t.:

1. $P_t(x,\cdot)$, i.e. $A\mapsto P_t(x,A)$ is a probability measure on $\mathcal{E}$ $\forall x\in E$
2. $P_t(\cdot,A)$, i.e. $x\mapsto P_t(x,A)$ is $\mathcal{E}$-measurable for all $A\in\mathcal{E}$
3. $\underbrace{\mathbb{P}(X_{t+s}\in A\mid X_s = x)}$ $= \int_A P_t(x,dy)$
$= \mathbb{E}[\chi_A(X_{t+s})\mid X_s = x]$

Comment: if X_{t+s}, X_s have joint distribution: f_{X_{t+s},X_s} and conditional distribution: $f_{X_{t+s}|X_s}$ then:
$\underbrace{\mathbb{P}(X_{t+s}\in A\mid X_s = x)}_{\substack{P_t(x,A) \\ \text{probability starting from } x \\ \text{to be in } A \text{ after a time } t}} = \int_A f_{X_{t+s}|X_t}(y|x)\,dy$
$\longleftarrow$ $f_{X_{t+s}|X_s}(\cdot|x)$ is the density of $A\mapsto P_t(x|A)$ w.r.t. the Lebesgue measure

Remark: $(X_t)_{t\geq 0}$ continuous time MC with E countable:

$P_t(x,A)$? : $\mathbb{P}(X_t \in A \mid X_0 = x) = \sum_{j\in A} \mathbb{P}(X_t = j \mid X_0 = x) = \sum_{j\in A} p_{ij}(t)$

$$\longrightarrow \boxed{p_{ij}(t) = P_t(i, \{j\})}$$

$(X_t)_{t\geq 0}$ Markov process, $\mathbb{E}[f(X_{t+s}) \mid \sigma(X_s)] = ?$ P_t ?

If $f = \chi_A$, $A \in \mathcal{E}$: $\mathbb{E}[f(X_{t+s}) \mid \sigma(X_s)] = \mathbb{P}(X_{t+s} \in A \mid \sigma(X_s))$

$$= P_t(X_s, A)$$

$$= \int_A \chi_A(y)\, P_t(X_s, dy)$$

Summing up: $\mathbb{E}[\chi_A(X_{t+s}) \mid \sigma(X_s)] = \int_E \chi_A(y)\, P_t(X_s, dy)$

For $f : E \to \mathbb{R}$ measurable s.t. $\mathbb{E}[|f(X_{t+s})|] < \infty$ then:

$$\boxed{\mathbb{E}[f(X_{t+s}) \mid \sigma(X_s)] = \underbrace{\int_E f(y)\, P_t(X_s, dy)}_{\omega \longmapsto \int_E f(y)\, P_t(X_s(\omega), dy)}}$$

CHAPMAN-KOLMOGOROV EQUATION

In the discrete E case was:

$$\mathbb{P}(X_t = j \mid X_r = k) = \sum_h \mathbb{P}(X_t = j \mid X_s = h)\, \mathbb{P}(X_s = h \mid X_r = k)$$

$$p_{kj}(t-r) = \sum_h p_{kh}(s-r)\, p_{hj}(t-s)$$

In the continuous E case is:

$$\mathbb{P}(X_t \in A \mid X_r = x) = \text{“}\int_E \underbrace{\mathbb{P}(X_t \in A \mid X_s = y)}_{P_{t-s}(y,A)}\, \underbrace{\mathbb{P}(X_s = y \mid X_r = x)\, dy}_{P_{s-r}(x,dy)}\text{”}$$

$$= \int_E P_{t-s}(y,A)\, P_{s-r}(x,dy)$$

Discrete time case: $P_{n+m}(x,A) = \int_E P_n(y,A)\, P_m(x,dy)$

Example 1: AUTO-REGRESSIVE GAUSSIAN PROCESS

$b \in \mathbb{R}$, $\sigma > 0$, $(Z_n)_{n\geq 1}$ iid $N(0,1)$ r.v.

- initial condition X_0
- $X_{n+1} = b X_n + \sigma Z_{n+1}$

(X_n function of $Z_1, \dots, Z_n$ is independent of Z_{n+1})

Markov property $\mathbb{P}(X_3 \in E_3 \mid X_2 \in E_2, X_1 \in E_1) \overset{*}{=}$

$$= \mathbb{P}(bX_2 + \sigma Z_3 \in E_3 \mid X_2 \in E_2, X_1 \in E_1)$$

(Remember that $Z_3 \perp\!\!\!\perp X_1, X_2$)

$$= \mathbb{E}[\chi_{E_3 - bX_2}(Z_3) \mid X_2 \in E_2, X_1 \in E_1]$$

Nel caso in cui non fossimo convinti della formula:

$$* = \frac{\mathbb{P}(X_3 \in E_3, X_2 \in E_2, X_1 \in E_1)}{\mathbb{P}(X_2 \in E_2, X_1 \in E_1)}; \quad X_2 = bX_1 + \sigma Z_2, \quad X_3 = bX_2 + \sigma Z_3 = b^2 X_1 + b\sigma Z_2 + \sigma Z_3$$

$$\int_E dx_1 \int_E dz_2 \int_E dz_3\, \chi_{E_3}(b^2 x_1 + \sigma z_2 + \sigma z_3)\, \chi_{E_2}(bx_1 + \sigma z_2) \cdot \chi_{E_1}(x_1)\, f_{X_1}(x_1)\, \frac{1}{\sqrt{2\pi}} e^{-\frac{z_2^2}{2}}\, \frac{1}{\sqrt{2\pi}} e^{-\frac{z_3^2}{2}}$$

In realtà l'ha fatto solo per farci vedere quanto sia sbatti il calcolo quindi dobbiamo fidarci della formula

More intuitive: (more heuristic) Consider first the case $E_2 = I_2 = (x_2 - \varepsilon, x_2 + \varepsilon)$ if we prove the formula on a small interval then we can sum the intervals obtaing that we want

$$\mathbb{P}(X_3 \in E_3 \mid X_2 \in E_2, X_1 \in E_1)$$

$$= \mathbb{P}(\underbrace{bX_2}_{\approx bx} + \sigma Z_3 \in E_3 \mid X_2 \approx x, X_1 \in E_1)$$

$$= \mathbb{P}\left(Z_3 \in \frac{E_3 - bx}{\sigma} \mid X_2 \approx x\right) \longleftarrow \perp\!\!\!\perp X_1$$

$$\approx \mathbb{P}(bX_2 + \sigma Z_3 \in E_3 \mid X_2 \approx x) = \mathbb{P}(X_3 \in E_3 \mid X_2 \in E_2)$$

What are the transition kernels?

$P(x,A) = \mathbb{P}(X_{n+1} \in A \mid X_n = x)$

or (more correctly) $P(X_n, A) = \mathbb{E}[\chi_A(X_{n+1}) \mid \sigma(X_n)]$, $X_{n+1} = bX_n + \sigma Z_{n+1}$

$$= \int_E \chi_A(bX_n + \sigma y)\, \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}\, dy$$

change of variables: $bX_n + \sigma y = w$, $y = \frac{w - bX_n}{\sigma}$

$$= \int_E \chi_A(w)\, \frac{1}{\sqrt{2\pi}} e^{-\frac{(w-bX_n)^2}{2\sigma^2}}\, \frac{dw}{\sigma}$$

$$= \int_A \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(w-bX_n)^2}{2\sigma^2}}\, dw$$

$$= P_1(X_n, A)$$

$X_{n+1} = bX_n + \sigma Z_{n+1}$, fixed X_n we have: $X_{n+1} \sim N(bX_n, \sigma^2)$

or:

$$\mathbb{E}[X_{n+1} \mid \sigma(X_n)] = \mathbb{E}[bX_n \mid \sigma(X_n)] + \mathbb{E}[\sigma Z_{n+1} \mid \sigma(X_n)] + \underbrace{\mathbb{E}[\sigma Z_{n+1}]}_{=0}$$

they're ⫫

$$= b X_n$$

27/11/2019

Example 2: (a M.P. with $t \in [0,\infty)$ and values $\mathbb{R}$)

BROWNIAN MOTION

$(B_t)_{t \geq 0}$ r.v. with $B_0 = 0$ and:

1. $\forall n \;\; \forall 0 < t_1 < \dots < t_n$: $B_{t_1}, B_{t_2 - t_1}, \dots, B_{t_n} - B_{t_{n-1}}$ are ⫫ (we have independent increments)
2. $B_t - B_s \sim N(0, \sigma^2(t-s))$

If we have $t_3 > t_2 > t_1$:

$$\mathbb{P}(B_{t_3} \in E_3 \mid \underbrace{B_{t_2} \in E_2}_{\text{we think about } E_2 \text{ like } (x-\epsilon, x+\epsilon)}, B_{t_1} \in E_1) = \mathbb{P}(\underbrace{B_{t_3} - B_{t_2}}_{\text{independent of } B_{t_1}} \in E_3 - x \mid B_{t_2} \in (x-\epsilon, x+\epsilon), \cancel{B_{t_1} \in E_1})$$

Transition kernel: $P_t(x, A) = \mathbb{P}(B_{s+t} \in A \mid B_s = x)$

probability of starting from x and reach A

$$\mathbb{E}[X_n(B_{t+s}) \mid \sigma(B_s)]$$

$$= \mathbb{P}(\underbrace{B_{t+s} - B_s}_{\sim N(0,\sigma^2 t)} + x \in A \mid B_s = x)$$

$$= \frac{1}{\sqrt{2\pi\sigma^2 t}} \int_A e^{-\frac{(y-x)^2}{2\sigma^2 t}} \, dy$$

if we fix the x we have a measure with the density:

$$\frac{1}{\sqrt{2\pi\sigma^2 t}} e^{-\frac{(y-x)^2}{2\sigma^2 t}} \, dy$$

Often $P_t(x, A) = \int_A k_t(x,y) \, dy$ (often we identify the ~~kernel~~ density with the kernel)

$A \in \mathcal{B}(\mathbb{R})$

BROWNIAN MOTION AS SCALING LIMIT OF SYMMETRIC RW

$(X_n)_{n \geq 0}$ MC with states $\mathbb{Z}$ and transition matrix:

$X_0 = 0$

$$\begin{bmatrix} \ddots & & & & \\ \cdots 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \cdots \\ & & & & \ddots \end{bmatrix}$$

$(X_n)_n$ symmetric random walk

$$(n-1) \xleftarrow{\frac{1}{2}} (n) \xrightarrow{\frac{1}{2}} (n+1)$$

~~We know from the central limit theorem~~

~~$\forall t$: $\frac{X_1 + \dots + X_{[nt]}}{\sqrt{n}} \xrightarrow{\text{in law}} B_t \sim N(0,t)$~~

~~convergence to a random variable B_t~~

~~$[nt] \sim nt$ integer part of nt~~

~~We fix $s < t$ and we consider the increment:~~

~~$\frac{X_{[ns]+1} + \dots + X_{[nt]}}{\sqrt{n}} \longrightarrow B_t - B_s$~~

~~$X_1 + \dots + X_{[ns]}$ ⫫~~

$(X_n)_{n \geq 0}$ symmetric random walk with $X_0 = 0$

$X_n = Y_1 + \dots + Y_n$ with $Y_j \overset{iid}{\sim}$: $\mathbb{P}(Y_j = 1) = \mathbb{P}(Y_j = -1) = \frac{1}{2}$

$\mathbb{E}[Y_j] = 0, \quad V[Y_j] = 0$

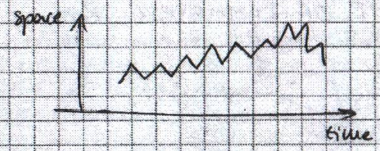
$$\frac{X_n}{\sqrt{n}} = \frac{Y_1 + \dots + Y_n}{\sqrt{n}} \xrightarrow{\text{in law}} N(0,1)$$

Central limit theorem: $\forall t \quad \frac{X_{[nt]}}{\sqrt{n}} = \frac{Y_1 + \dots + Y_{[nt]}}{\sqrt{n}} \longrightarrow N(0,t)$ since $[nt] \sim nt$

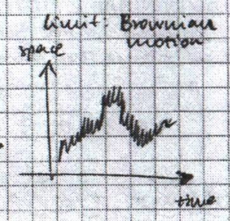
lo e la fra
supponiamo che:
X_n = posizione al tempo n
Y_n = se vado avanti di 1 o indietro di 1 al tempo n

Considering $t, s > 0$: $\frac{X_{[n(t+s)]} - X_{[nt]}}{\sqrt{n}} = \frac{Y_{[nt]+1} + \dots + Y_{[n(t+s)]}}{\sqrt{n}} \xrightarrow{n \to \infty} N(0,s)$

$$\Rightarrow \begin{cases} \frac{X_{[n(t+s)]} - X_{[nt]}}{\sqrt{n}} \longrightarrow B_{t+s} - B_t \sim N(0,s) \\ \frac{X_{[nt]}}{\sqrt{n}} \longrightarrow B_t \sim N(0,t) \end{cases}$$



we rescale the time and the space, and in the limit we got the Brownian motion



fine dell'esempio

OBSERVE THAT : (A formula holds:)

$$\mathbb{E}\left[f(X_{t+s})\, g(X_s) \right] = \int_E g(x)\, \mu_s(dx) \underbrace{\int_E f(y)\, P_t(x, dy)}_{} \qquad (*)$$

formula which relate the kernel with the expectation

case: $f = \chi_A$

$P_t(x, A) = \mathbb{P}(X_{t+s} \in A \mid X_s = x)$

$X_s \sim \mu_s$

If (X_{t+s}, X_s) has joint density ϕ :

$$\mathbb{E}\left[f(X_{t+s})\, g(X_s) \right] = \int_E \int_E f(y)\, g(x)\, \underbrace{\phi(x,y)}_{\text{can be written as}}\, dx\, dy$$

$\phi_{X_{t+s}|X_s}(y|x)\, \phi_{X_s}(x)$ by def. of conditional density

$$= \int_E g(x)\, \phi_{X_s}(x)\, dx \int_E f(y) \boxed{\phi_{X_{t+s}|X_s}(y|x)\, dy}$$

$ \quad = P_t(x, dy)$

comparing this integral with $(*)$ we got

STATIONARY DISTRIBUTIONS

$(X_t)_{t \geq 0}$ MP and μ a measure on $\mathcal{E}$, μ is stationary (or invariant) if when $X_0 \sim \mu$ then $X_t \sim \mu$ $\forall t$

($\equiv$ if all the random variable has the same distribution: $X_0 \sim$ distr 1 $\implies X_t \sim$ distr 1 $\forall t$)

If we look at the formula $\uparrow$ $(*)$

$f = \chi_A, \quad g = 1$:

$$\mathbb{P}(X_{t+s} \in A) = \mu_{t+s}(A) = \int_E \mu_s(dx)\, P_t(x, A)$$

A μ is stationary if and only if $\boxed{\mu(A) = \int_E \mu(dx)\, P_t(x, A) \quad \forall t}$!

Proposition: A measure μ is invariant if and only if: $\mu(A) = \int_E \mu(dx)\, P_t(x, A) \quad \forall t \quad \forall A$

Brownian motion: has an invariant measure?

Before of that: the ~~m~~ symmetric random walk has an invariant measure? (we did it at the beginning) No

$\implies$ Since the Brownian is the limit of a process which doesn't have an invariant measure, ~~[illegible]~~ the Brownian motion doesn't have an invariant measure

$$\underset{\substack{\text{transition}\\ \text{kernel of}\\ \text{Brownian measure}}}{P_t(x, dy)} = \frac{1}{\sqrt{2\pi t}}\, e^{-\frac{(y-x)^2}{2t}}\, dy$$

(continuation of Brownian)

If μ is a probability measure on $\mathcal{B}(\mathbb{R})$ s.t.:

$$\mu(A) = \int_{\mathbb{R}} \mu(dx) \int_A \frac{1}{\sqrt{2\pi t}}\, e^{-\frac{(y-x)^2}{2t}}\, dy \qquad \forall t\ \forall A$$

$\forall t$: $t \to \infty$: $\frac{1}{\sqrt{2\pi t}} \to 0 \quad \to \mu(A) \xrightarrow{t \to \infty} 0$

$\implies \nexists$ a μ defined as $(*)$ that is a stationary measure

If $\exists \mu$ s.t. $B_t \sim \mu$ $\forall t$, then

$$B_{t+s} = B_s + \underbrace{(B_{t+s} - B_s)}_{}$$

$$\mu \overset{?}{=} \mu + N(0,t)$$

this is abnormal, because: (in realtà funziona solo con le gaussiane)

$$\sim \mu \sim \underbrace{\mu * N(0, \sigma^2)}_{\substack{\text{convolution of}\\ \mu \text{ and } N(0,\sigma^2)}} \implies \mu \sim N(\vartheta, \sigma^2)$$

// tutto per dire: the Brownian motion doesn't have an invariant distribution since it's the limit of a process which doesn't have an invariant distribution

---

Example: Autoregressive Gaussian process:

$$X_{n+1} = b X_n + \sigma Z_{n+1} \qquad (Z_n)_{n \geq 1},\ N(0,1) \text{ indep.}$$

Do we have an invariant distribution here?

We try with Gaussians: $N(0, \theta)$

Let $\phi_{X_{n+1}}(t) = \mathbb{E}\left[e^{it X_{n+1}} \right] \underset{\uparrow}{=} e^{-\frac{\theta t^2}{2}}$

Gaussian characteristic function

$$\phi_{bX_n + \sigma Z_{n+1}}(t) = \phi_{bX_n}(t)\, \phi_{\sigma Z_{n+1}}(t)$$

characteristic functions property

$$= \phi_{X_n}(bt)\, \phi_{Z_{n+1}}(\sigma t)$$

$$= e^{-\frac{\theta b^2 t^2}{2}}\, e^{-\frac{\sigma^2 t^2}{2}}$$

$$\phi_{X_{n+1}}(t) = \phi_{bX_n + \sigma Z_{n+1}}(t) \quad \forall t$$

$$-\frac{\theta t^2}{2} = -\frac{\theta b^2 t^2}{2} - \frac{\sigma^2 t^2}{2} \longleftrightarrow \theta = \frac{\sigma^2}{1-b^2}$$

$\implies$ We have an invariant density

if $X_n \sim N\left(0, \frac{\sigma^2}{1-b^2}\right) \implies X_{n+1} \sim N\left(0, \frac{\sigma^2}{1-b^2}\right)$

(ATTENZIONE: if $|b| < 1$)

$N\left(0, \frac{\sigma^2}{1-b^2}\right)$ it's the invariant density

IRREDUCIBILITY

(MC irreducible $\longrightarrow$ the invariant measure is unique)

The irreducibility has something to do with $P_t(x, A)$: if E is discrete, $x = i$, $A = \{j\}$ $\longrightarrow$ $P_t(x, A) = p_{ij}(t)$

3/12/2019

$(X_t)_{t\geq 0}$ Markov processes with values in $(E, \mathcal{E})$
$(P_t)_{t\geq 0}$ transition kernels

IRREDUCIBLE with respect to a reference measure φ on E if $\forall x \in E$, $\forall A \in \mathcal{E}$ with $\varphi(A) > 0$ $\exists\, t > 0$ s.t. $P_t(x, A) > 0$

Remark: if E is discrete and φ is the counting measure ($\varphi(A) = |A|$ cardinality of a subset) then:
(φ-irreducible $\longleftrightarrow$ irreducible in the sense of previous)

Remark: if a MP is irreducible w.r.t. several measures, one can always find a "maximal reference measure" φ^* s.t. "given any other reasonable measure φ then $\varphi \ll \varphi^*$" (vedi Zorn's lemma & Co.)

Recurrence (we consider only the discrete time and $E \subseteq \mathbb{R}^d$, $\mathcal{E} = \mathcal{B}(E)$ for simplicity)

$$A \subset E, \quad \underbrace{N_A}_{\text{number of visits in the set } A} = \sum_{n\geq 0} \chi_A(X_n) = \sum_{n\geq 0} \chi_{(X_n \in A)}$$

Def. For a MP $(X_n)_{n\geq 0}$ a set A is called Harris recurrent if $\forall x \in A$ $\underbrace{\mathbb{P}_x(N_A = +\infty) = 1}_{\text{starting from } x \text{ the prob. of visiting } A \ \infty \text{ times} = 1}$

Def. For a MP $(X_n)_{n\geq 0}$, $(X_n)_{n\geq 0}$ is Harris recurrent if $\exists \varphi$ reference measure s.t. all $A \subset E$ with $\varphi(A) > 0$ is Harris recurrent

recurrent :
- set : if we return ∞ times in the set
- $(X_n)_{n\geq 0}$ (MP) : if $\forall$ set is recurrent

Remark: it could be shown that $\mathbb{P}_x(N_A = +\infty) = 1 \Leftrightarrow \mathbb{E}_x[N_A] = +\infty$
this holds because of the (strong) Markov property

Example: $X_{n+1} = b X_n + \sigma Z_{n+1}$ with $|b| < 1$, $(Z_n)_{n\geq 1} \overset{iid}{\sim} N(0,1)$

We proved that $N\left(0, \frac{\sigma^2}{1-b^2}\right)$ is an invariant measure

Is the process irreducible, transient, recurrent?

invariant irreducible $\longrightarrow$ recurrent

(Example) ① Irreducibility?

$$P_1(x, A) = \mathbb{P}(X_{n+1} \in A \mid X_n = x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_A e^{-\frac{(y-bx)^2}{2\sigma^2}}\, dy \qquad \Big|\ X_{n+1} - bx \sim N(0, \sigma^2)$$

$E = \mathbb{R}$, φ = lebesgue measure

$\forall x \quad P_1(x, E) > 0 \iff \varphi(A) > 0$

$\Longrightarrow$ is φ-irreducible

② Harris recurrence? We expect it to be recurrent because it has an invariant measure

$\forall A$ s.t. $\varphi(A) > 0$ (suppose $\varphi(A^c) > 0$) $\quad Y_n = \begin{cases} 1 & \text{if } X_n \in A \\ 0 & \text{if } X_n \notin A \end{cases} = \chi_A(X_n)$

$(Y_n)_{n\geq 0}$ is a $\{0,1\}$-valued MC with transition matrix

$$\underbrace{\begin{array}{c|cc} & 0 & 1 \\ \hline 0 & \alpha & 1-\alpha \\ 1 & 1-\beta & \beta \end{array}}_{\text{finite number of states and irreducible}}, \quad \alpha = \mathbb{P}(Y_{n+1} = 0 \mid Y_n = 0) = \mathbb{P}(X_{n+1} \in A^c \mid X_n \in A^c)$$

$\alpha > 0$

$\beta > 0$

$\Longrightarrow$ we visit A and A^c ∞ time $\Longrightarrow$ it's ~~recurrent~~ Harris recurrent

Exercise: (see lecture notes) 1. $X_n \xrightarrow{\text{in law}} N\left(0, \frac{\sigma^2}{1-b^2}\right)$

2. $X_n \overset{L^2}{\not\longrightarrow} N\left(0, \frac{\sigma^2}{1-b^2}\right)$

Remark : LAW OF LARGE NUMBERS
$(X_n)_{n\geq 0}$ Harris recurrent Markov process w.r.t. φ
For any $\mathcal{E}$-measurable function $f: E \to \mathbb{R}$ s.t.

$$\int_E |f(x)|\, \varphi(dx) < +\infty \quad : \quad \text{INVARIANT MEASURE}$$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} f(X_k) = \int_E f(x)\, \varphi(dx)$$

Discrete space analogy: $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n} f(X_k) = \sum_{j\in E} f(j)\, \pi_j$

with $(\pi_j)_j$ invariant distribution

end of the real value MC #

MARTINGALES

$(\Omega, \mathcal{F}, \mathbb{P})$ probability space, $(M_t)_{t\geq0}$ real random variables — $(\mathcal{F}_t)_{t\geq0}$ increasing* family of sub σ-algebra of $\mathcal{F}$

Def. $(M_t)_{t\geq0}$ is called a martingale if:

1. $\mathbb{E}[|M_t|] < +\infty \quad \forall t$
2. M_t is $\mathcal{F}_t$-measurables $\forall t$
3. $\forall s<t$: $\mathbb{E}[M_t | \mathcal{F}_s] = M_s$:= MARTINGALE PROPERTY

* increasing: $s<t \Rightarrow \mathcal{F}_s \subseteq \mathcal{F}_t$ ⇒ := FILTRATION

Remark. If 1.+2. and $\forall s<t$ $\mathbb{E}[M_t|\mathcal{F}_s] \geq M_s$ ⤳ SUBMARTINGALE
If 1.+2. and $\forall s<t$ $\mathbb{E}[M_t|\mathcal{F}_s] \leq M_s$ ⤳ SUPERMARTINGALE

Example: (random walk) $(X_n)_{n\geq0}$: $\begin{cases} X_0 = x \in \mathbb{Z} \\ X_{n+1} = X_n + Y_n \end{cases}$ with Y_n s.t. $\begin{cases} \mathbb{P}(Y_n = 1) = p \\ \mathbb{P}(Y_n = -1) = 1-p = q \end{cases}$

Exercise: $M_n = X_n - n(p-q)$, $(M_n)_{n\geq0}$ is a martingale w.r.t. $(\mathcal{F}_n)_{n\geq0}$ s.t. $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$

We check 1., 2., 3.:

1. $|M_n| \leq |X_n| + n|p-q| \leq |x_0| + n + n|p-q|$
   $\Longrightarrow \mathbb{E}[|M_n|] < +\infty$
2. M_n is $\mathcal{F}_n$-measurable by definition
3. $\mathbb{E}[M_n | \mathcal{F}_m] \overset{?}{}$ $m<n$:

- First $m = n-1$:

$$\mathbb{E}[M_n|\mathcal{F}_{n-1}] = \mathbb{E}\Big[\underbrace{(X_n - X_{n-1})}_{Y_n} + \underbrace{X_{n-1} - n(p-q)}_{\substack{\mathcal{F}_{n-1}\text{- measurable} \\ \text{so goes out } \mathbb{E}[\cdot|\mathcal{F}_{n-1}]}} \Big| \mathcal{F}_{n-1}\Big]$$

$$= \mathbb{E}[Y_n|\mathcal{F}_{n-1}] + X_{n-1} - n(p-q)$$

$Y_n \perp X_{n-1}, X_{n-2}, \dots$ generate $\mathcal{F}_{n-1}$

$$= \mathbb{E}[Y_n] + X_{n-1} - n(p-q)$$
$$= (p-q) + X_{n-1} - n(p-q)$$
$$= X_{n-1} - (n-1)(p-q) = M_{n-1}$$

- Second: (generic $m<n$):

$$\mathbb{E}[M_n|\mathcal{F}_m] = \mathbb{E}\big[\mathbb{E}[M_n|\mathcal{F}_{n-1}]\,\big|\,\mathcal{F}_n\big] = \mathbb{E}[M_{n-1}|\mathcal{F}_m] = \dots = M_m$$

More generally, $(X_n)_{n\geq0}$ MC discrete time & discrete state

$$\mathbb{E}[f(X_n)|\mathcal{F}_m] = (P^{n-m}f)(X_m) \quad \forall m<n, \quad \mathcal{F}_m = \sigma(X_0, \dots, X_m)$$

Proposition: $M_n := f(X_n) - \sum_{k=1}^{n-1} (Pf(X_k) - f(X_k))$ define a Martingale w.r.t. $(\mathcal{F}_n)_{n\geq0}$, $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$

(when we consider $f(X_n)$, $\forall f$, the only thing we need is the integrability of f)

proof.

$$\mathbb{E}[M_{n+1}|\mathcal{F}_n] \overset{?}{=} M_n \iff \mathbb{E}[M_{n+1} - M_n|\mathcal{F}_n] = 0$$

$$\mathbb{E}\Big[f(X_{n+1}) - \textstyle\sum_{k=0}^{n}(\dots) - f(X_n) + \sum_{k=0}^{n}(\dots)\Big|\mathcal{F}_n\Big] =$$

$$= \mathbb{E}\Big[f(X_{n+1}) \underbrace{- \cancel{f(X_n)} - (Pf(X_n) - \cancel{f(X_n)})}_{\mathcal{F}_n\text{- measurable}} \Big| \mathcal{F}_n\Big]$$

$$= \mathbb{E}[f(X_{n+1})|\mathcal{F}_n] - Pf(X_n) = 0 \quad \blacksquare$$

04/12/2019

MARTINGALE STOPPING THEOREM

$(\Omega, \mathcal{F}, \mathbb{P})$ probability space, $(\mathcal{F}_t)_t$ filtration

Def. $T: \Omega \to [0, +\infty]$ is a stopping time for the filtration $(\mathcal{F}_t)_t$ if $\{T \leq t\} \in \mathcal{F}_t \quad \forall t$.

Equivalently one can say:
- $\{T > t\} \in \mathcal{F}_t \quad \forall t$
- for discrete times: $\{T = n\} \in \mathcal{F}_n$
  $\{T \leq n\} \setminus \{T \leq n-1\}$

Theorem (Martingale stopping theorem) $(M_t)_{t\geq0}$ martingale, $t \in$ discrete set, martingale w.r.t. a filtration $(\mathcal{F}_t)_t$. T is a stopping time of the same filtration. The stopped process $(M_{T\wedge t})_t$ is also a martingale and:

$$\mathbb{E}[M_{T\wedge t}] = \mathbb{E}[M_0] = \mathbb{E}[M_t] \quad \forall t$$

immediate consequence of martingale's propriety:

$$\mathbb{E}[M_{T\wedge t}] = \mathbb{E}[\mathbb{E}[M_{T\wedge t}|\mathcal{F}_0]] = \mathbb{E}[M_0]$$

(Martingales have constant expectation because of their def.)

proof. In the notes. ■

Example: $(X_n)_{n\geq0}$ symmetric random walk on integers.
$X_0 = 0$, $-a < 0 < b$ $a, b \in \mathbb{N}$
$T = \min\{n \geq 1 \mid X_n = -a \text{ or } X_n = b\}$
exit time from $(-a, b)$

Compute:
1. probability of exit from $-a$ (i.e. $\mathbb{P}(X_T = -a)$) or b
2. mean exit time

(1) Consider the martingale $(X_n)_{n\geq0}$ w.r.t. $(\mathcal{F}_n)_n$: $\mathcal{F}_n = \sigma(X_0, \dots, X_n)$ (when it's not written anything ⟹ it's the natural filtration)
The stopping theorem says:

$$\mathbb{E}[X_{T\wedge n}] = \mathbb{E}[X_0] = 0 \quad \forall n$$

$|X_{T\wedge n}| \leq \max\{a,b\} \quad \forall n$

When $n \to \infty$: $\begin{cases} T\wedge n \to T \\ X_{T\wedge n} \to -a\,\chi_{\{X_T=a\}} + b\,\chi_{\{X_T=b\}} \end{cases}$

we're assuming that T is almost surely finite (*)

this because we can't wondering forever in the interval

(*) Suppose $\mathbb{P}(T<\infty)=1$.
Using the dominated convergence:
$\mathbb{E}[-a\,\mathbb{1}_{\{X_T=a\}} + b\,\mathbb{1}_{\{X_T=b\}}] = 0$

$= -a\,\mathbb{P}(X_T=-a) + b\,\mathbb{P}(X_T=b) = 0$

Note that: $\mathbb{P}(X_T=-a) + \mathbb{P}(X_T=b) = 1$
because the exit can be either from the side $(-a)$ or (b), not both

$\longrightarrow$ Solve: $\mathbb{P}(X_T=-a) = \frac{b}{a+b}$, $\mathbb{P}(X_T=b) = \frac{a}{a+b}$

(2) Consider the martingale $M_n = X_n^2 - n$.
We check that is a martingale:

$X_n = \sum_{k=1}^n Y_k$, Y_k: $\mathbb{P}(Y_k=1)=\mathbb{P}(Y_k=-1)=\frac{1}{2}$, Y_k iid

$\mathbb{E}[X_{n+1}^2 - (n+1) \mid \mathcal{F}_n] = \mathbb{E}[(X_n+Y_{n+1})^2 \mid \mathcal{F}_n] - (n+1)$

$= \mathbb{E}[X_n^2 + 2X_nY_{n+1} + \underbrace{Y_{n+1}^2}_{=1} \mid \mathcal{F}_n] - (n+\cancel{1})$

since X_n is $\mathcal{F}_n$-measurable

$= X_n^2 + 2X_n\,\mathbb{E}[Y_{n+1} \mid \mathcal{F}_n] - n$

are independent

$= X_n^2 + 2X_n\,\mathbb{E}[Y_{n+1}] - n = X_n^2 - n$

$\mathbb{E}[Y_{n+1}]=0$

Alternative check:

$f(X_n) - \sum_{k=0}^{n-1}(Pf(X_k) - f(X_k))$

$f(i) = i^2$, $P = \begin{bmatrix} 0 & 0 & 0 & \cdots \\ 1/2 & 0 & 1/2 & \cdots \\ 0 & 1/2 & 0 & 1/2 & \cdots \\ & & \cdots \end{bmatrix}$

$Pf(i) = \frac{(i-1)^2}{2} + \frac{(i+1)^2}{2} = i^2+1 \implies Pf(i) - f(i) = 1$

$\longrightarrow \sum_{k=0}^{n-1}(Pf(X_k) - f(X_k)) = \sum_{k=0}^{n-1} 1 = n$

We want now to compute the mean.
We apply the stopping theorem:

$\mathbb{E}[X_{T\wedge n}^2 - T\wedge n] = \mathbb{E}[M_0] = 0$

$\Longrightarrow \mathbb{E}[T\wedge n] = \mathbb{E}[X_{T\wedge n}^2] \leq \max\{a^2,b^2\}$

We use Fatou's Lemma:
$\mathbb{E}[T] \leq \liminf_n \mathbb{E}[T\wedge n] < +\infty$

$\longrightarrow \mathbb{P}(T<+\infty)=1$

We use monotone convergence: $\mathbb{E}[T] = \lim_{n\to\infty}\mathbb{E}[T\wedge n]$

By dominated convergence: $\mathbb{E}[X_T^2] = \lim_{n\to\infty}\mathbb{E}[X_{T\wedge n}^2]$

$\Longrightarrow \mathbb{E}[T] = \mathbb{E}[X_T^2] = a^2\,\mathbb{P}(X_T=-a) + b^2\,\mathbb{P}(X_T=b)$

$= \frac{a^2b}{a+b} + \frac{b^2a}{a+b} = ab$

(this is something we already knew from the Gambler's problem)
(with just a translation)

In the stopping theorem we always end up with something like:

$\mathbb{E}[M_{T\wedge n}] = \mathbb{E}[M_0] \quad \forall n$

If we are not able to remove the truncation it's a problem.

Warning: it is not always true that $\mathbb{E}[M_T] = \lim_{n\to\infty}\mathbb{E}[M_{T\wedge n}]$

Counter example: $(X_n)_{n\geq0}$ symmetric random walk on $\mathbb{Z}$
$X_0 = 0$ and $T = \min\{n\geq1 \mid X_n = 1\}$
We know that $(X_n)_{n\geq0}$ is a martingale
$\xRightarrow{\text{stopping theorem}} \mathbb{E}[X_{T\wedge n}] = \mathbb{E}[X_0] = 0$
However: $T\wedge n \xrightarrow[(a.s.)]{\text{almost surely}} T \quad \mathbb{P}(T<+\infty)=1$
$\Longrightarrow \lim_{n\to\infty} X_{T\wedge n} = X_T = 1$
$\Longrightarrow \mathbb{E}[X_T] = 1 \neq \lim_{n\to\infty}\mathbb{E}[X_{T\wedge n}] = 0$

NOTE that $\mathbb{E}[T]=\infty$ because it's slowly recurrent (even if $\mathbb{P}(T<\infty)=1$)

We can't control the negative part of T, we ~~have no~~ can't apply the dominated convergence theorem $\Longrightarrow$ we cannot exchange the limit and the expectation

10/12/2019

$(M_t)_{t\in D}$ Martingale with respect to a filtration $(\mathcal{F}_t)_{t\in D}$

Theorem: T stopping time of $(\mathcal{F}_t)_{t\in D}$, $(\{T\leq t\}\in\mathcal{F}_t \;\; \forall t\in D)$
$(M_{T\wedge t})_{t\in D}$ is a martingale. In particular:
$\mathbb{E}[M_{T\wedge t}] = \mathbb{E}[M_0] \quad (=\mathbb{E}[M_t]) \quad \forall t\in D$

proof.

1. $\mathbb{E}[|M_{T\wedge t}|] < \infty$

$T = \sum_{r\in D} r\,\mathbb{1}_{\{T=r\}} \longrightarrow M_{T\wedge t} = \sum_{r\in} M_{r\wedge t}\,\mathbb{1}_{\{T=r\}}$

$= \underbrace{\sum_{r\in D\cap[0,t)} M_r}_{\text{finite}} + M_t\,\mathbb{1}_{\{T\geq t\}}$

$$|M_{T\wedge t}| \le \underbrace{\sum_{r\in D\cap[0,t]} |M_r| + |M_t|\, \mathbb{1}_{\{T\ge t\}}}_{\text{Finite sum}}$$

$\Rightarrow \mathbb{E}[|M_{T\wedge t}|] < \infty$

2. $M_{T\wedge t}$ is $\mathcal{F}_t$-measurable

$$M_{T\wedge t} = \sum_{r\in D\cap[0,t)} \underbrace{M_r \mathbb{1}_{\{T=r\}}}_{\mathcal{F}_r\text{-measurable}} + \underbrace{M_t \mathbb{1}_{\{T\ge t\}}}_{\mathcal{F}_t\text{-measurable}} \Rightarrow M_{T\wedge t} \text{ is } \mathcal{F}_t\text{-meas.}$$

$\in \mathcal{F}_r \subseteq \mathcal{F}_t$

3. $\mathbb{E}[M_{T\wedge t} | \mathcal{F}_s] = M_{T\wedge s} \quad \forall s<t$

namely $\forall A \in \mathcal{F}_s$: $\int_A M_{T\wedge t}\, d\mathbb{P} = \int_A M_{T\wedge s}\, d\mathbb{P}$

Since $T = \sum_{r\in D} r\, \mathbb{1}_{\{T=r\}}$: For each r the set $\in \mathcal{F}_r$

$$\int_A M_{T\wedge s}\, d\mathbb{P} = \int_{A\cap\{T\le s\}} M_T\, d\mathbb{P} + \int_{\underbrace{A}_{\in\mathcal{F}_s}\cap\underbrace{\{T>s\}}_{\in\mathcal{F}_s}} M_s\, d\mathbb{P}$$

$\in \mathcal{F}_s$ $\Rightarrow M_s$ può essere M_t $\forall t\ge s$ siccome $\mathcal{F}_s \subset \mathcal{F}_t$

$$= \int_{A\cap\{T\le s\}} M_T\, d\mathbb{P} + \int_{A\cap\{T>s\}} M_t\, d\mathbb{P}$$

$$= \int_{A\cap\{T\le s\}} M_T\, d\mathbb{P} + \sum_{r\in D\cap(s,t)} \int_{A\cap\{T=r\}} M_t\, d\mathbb{P} + \int_{A\cap\{T\ge t\}} M_t\, d\mathbb{P}$$

$A \in \mathcal{F}_s$ $(s<r)$, $\{T=r\} \in \mathcal{F}_r$, $\in \mathcal{F}_r$

$$= \int_{A\cap\{T\le s\}} M_T\, d\mathbb{P} + \sum_{r\in D\cap(s,t]} \int_{A\cap\{T=r\}} M_r\, d\mathbb{P} + \int_{A\cap\{T>t\}} M_{T\wedge t}\, d\mathbb{P}$$

M_T (because $\{T=r\}$)

$$= \int_{A\cap\{T\le s\}} M_{T\wedge t}\, d\mathbb{P} + \int_{A\cap\{s<T\le t\}} M_{T\wedge t}\, d\mathbb{P} + \int_{A\cap\{T>t\}} M_{T\wedge t}\, d\mathbb{P}$$

$$= \int_A M_{T\wedge t}\, d\mathbb{P}$$

In particular $\mathbb{E}[M_{T\wedge t}] = \mathbb{E}[\mathbb{E}[M_{T\wedge t}|\mathcal{F}_0]] = \mathbb{E}[M_0] =$

$\mathbb{E}[M_t] = \mathbb{E}[\mathbb{E}[M_t|\mathcal{F}_0]] = \mathbb{E}[M_0]$ ■

(1) $(X_n)_{n\ge0}$ discrete time, discrete space MC, $f: E\to\mathbb{R}$ s.t.

$\mathbb{E}[|f(X_n)|] < +\infty \quad \forall n$, then:

$f(X_n) - \sum_{k=0}^{n-1} ((Pf)(X_k) - f(X_k))$ is a martingale w.r.t. the natural filtration of the process

(2) $(X_t)_{t\ge0}$ continuous time, discrete space MC s.t.

$\mathbb{E}[|f(X_t)|] < \infty$ and $\mathbb{E}[|(Qf)(X_t)|] < \infty$ ($\equiv$ regular MC)

then: $f(X_t) - \int_0^t (Qf)(X_s)\, ds$ is a martingale

proof.

$\mathbb{E}[f(X_t)|\mathcal{F}_s] = (P_{t-s}f)(X_s) \quad s<t$

$\mathcal{F}_s$ generated by events $\{X_s=i, X_{s_1}=i_1, \dots\}$ $s>s_1>\dots>0$

Show that: $\mathbb{E}[f(X_t) - \int_0^t \underbrace{(\cdot)}_{(Qf)(X_r)}\, dr \,|\, \mathcal{F}_s] = f(X_s) - \int_0^s (Qf)(X_r)\, dr$

$$\int_0^t (Qf)(X_r)\, dr = \underbrace{\int_0^s (Qf)(X_r)\, dr}_{\mathcal{F}_s\text{-meas.}} + \int_s^t (Qf)(X_r)\, dr$$

$$\Rightarrow \mathbb{E}\left[\int_0^s (Qf)(X_r)\, dr \,\Big|\, \mathcal{F}_s\right] = \int_0^s (Qf)(X_r)\, dr$$

$\Rightarrow$ We have to prove that $\mathbb{E}[f(X_t)|\mathcal{F}_s] = f(X_s)$

$$\mathbb{E}[f(X_t) - f(X_s)|\mathcal{F}_s] = \mathbb{E}\left[\int_s^t (Qf)(X_r)\, dr \,\Big|\, \mathcal{F}_s\right]$$

FUBINI'S

$$= \int_s^t \mathbb{E}[(Qf)(X_r)|\mathcal{F}_s]\, dr$$

$$= \int_s^t (P_{r-s}Qf)(X_s)\, dr$$

F K E | Kolmogorov forward equation $\left(\frac{d}{dt}P_t = P_t Q\right)$

$$= \int_s^t \frac{d}{dr}(P_{r-s}f)(X_s)\, dr$$

$$= (P_{t-s}f)(X_s) - f(X_s)$$

$\mathbb{E}[f(X_t) - f(X_s)|\mathcal{F}_s] = (P_{t-s})(X_s) - f(X_s)$ ■

Application: $M|M|1$ QUEUE : $Q = \begin{bmatrix} -\lambda & \lambda & 0 & 0 & \dots \\ \mu & -(\lambda+\mu) & \lambda & 0 & \dots \\ 0 & \mu & -(\lambda+\mu) & \lambda & 0 \dots \\ & & \vdots & & \end{bmatrix}$

$0<\lambda<\mu$

$T = \min\{t>0 \,|\, X_t = 0\}$ = first time for empty queue

(oppure lo si può pensare come birth/death process dove T è il first time of extinction)

We know that $\exists$ invariant density :

$\pi_n = \left(1-\frac{\lambda}{\mu}\right)\left(\frac{\lambda}{\mu}\right)^n \Rightarrow (X_n)_{n\ge0}$ is fast recurrent

: $\mathbb{P}_n(T < \infty) = 1$.

Compute $\mathbb{E}_n[T]$.

Consider $f: \mathbb{N} \to \mathbb{N}$: $f(n) = n$ (identity function)

$$(Qf)(n) = \begin{bmatrix} \lambda(f(1)-f(0)) \\ \vdots \\ \mu(f(n-1)-f(n)) + \lambda(f(n+1)-f(n)) \\ \vdots \end{bmatrix} = \begin{bmatrix} \lambda \\ -\mu+\lambda \\ \vdots \\ -\mu+\lambda \\ \vdots \end{bmatrix}$$

$$\Rightarrow X_t - \int_0^t \left(\lambda \mathbb{1}_{\{X_s=0\}} - (\mu-\lambda)\mathbb{1}_{\{X_s>0\}}\right) ds$$

$X_n = n$, $\quad n = \mathbb{E}\left[X_{T\wedge t} - \int_0^{T\wedge t} \left(\lambda \mathbb{1}_{\{X_s=0\}} - (\mu-\lambda)\mathbb{1}_{\{X_s>0\}}\right) ds\right]$

before T, $X_s > 0 \implies \lambda \mathbb{1}_{\{X_s=0\}} - (\mu-\lambda)\mathbb{1}_{\{X_s=0\}} = -(\mu-\lambda)$

then: $\quad n = \mathbb{E}\left[X_{T\wedge t} + \int_0^{T\wedge t} (\mu-\lambda)\, ds\right]$

$\Rightarrow n = \mathbb{E}\left[X_{T\wedge t} + (\mu-\lambda)(T\wedge t)\right]$

a good choice of f is a function which is constant at least in the interval in which we're integrating

$t \to \infty$: $X_{T\wedge t} \xrightarrow{\text{as } t\to\infty} X_T = 0$

(limit almost sure, and also in L^1)

$$n = \mathbb{E}\left[0 + (\mu-\lambda)T\right] \implies \mathbb{E}[T] = \frac{n}{\mu-\lambda}$$