CATENE DI MARKOV

Def. (Catena di Markov):
Sia $(E, \mathcal{E})$ sp. misurabile, $(\Omega, \mathcal{A}, \mathbb{P})$ sp. di prob., definiamo $\mathbb{T} = \mathbb{N} \cup \{0\}$ (oppure $\mathbb{T} = [0, +\infty)$). Sia inoltre $\mathbb{F} = (\mathcal{F}_t)_{t \in \mathbb{T}}$ una filtrazione. Diciamo che $(X_t)_{t\in\mathbb{T}}$ é una CATENA DI MARKOV (o PROCESSO DI MARKOV) per la filtrazione $\mathcal{F}$ se:

1) X_t é $\mathcal{F}_t$-misurabile $\forall t \in \mathbb{T}$ (il processo é adattato ad $\mathbb{F}$) (a valori in E)
2) Condizione (o proprietá) di Markov:
$\forall s, t \in \mathbb{T}$ $s<t$ $\forall A \in \mathcal{E}$ si ha che
$$\underset{\overset{\shortparallel}{\mathbb{E}(\mathbb{1}_{\{X_t \in A\}} \mid \mathcal{F}_s)}}{\mathbb{P}(X_t \in A \mid \mathcal{F}_s)} = \mathbb{P}(X_t \in A \mid X_s)$$

Se $\mathcal{F}_t = \sigma(\{X_s : s \leq t\})$ allora si dice che $(X_t)_{t\in\mathbb{T}}$ é DI MARKOV. Lo sp. mis. $(E, \mathcal{E})$ é detto SPAZIO DEGLI STATI DEL PROCESSO

Catene di Markov a tempo discreto ($\mathbb{T} = \mathbb{N} \cup \{0\}$) e a stati finiti (o numerabili):
Siano E insieme finito o numerabile $\mathbb{T} = \mathbb{N}_0$. Specializzando la def. precedente a questo caso si ha che $(X_n)_{n\geq 0}$ (succ. di v.a. a valori in E) é una catena di Markov se $\forall n \geq 0, \forall x_0, \dots, x_{n+1} \in E$ vale la proprietá di Markov (M):
$$\mathbb{P}(X_{n+1} = x_{n+1} \mid X_0 = x_0, \dots, X_n = x_n) = \mathbb{P}(X_{n+1} = x_{n+1} \mid X_n = x_n)$$

Osservazione:

(M) implica che $\forall k \geq 1$, $\forall x_0, \dots, x_{n+k}$ si ha:

$$\mathbb{P}(X_{n+k} = x_{n+k} \mid X_0 = x_0, \dots, X_n = x_n) = \mathbb{P}(X_{n+k} = x_{n+k} \mid X_n = x_n)$$

Infatti, per induzione su $k \geq 1$ si ha:

$k = 1 \longrightarrow$ segue direttamente da (M)

$k \longrightarrow k+1$:

$$\mathbb{P}(X_{n+k+1} = x_{n+k+1} \mid X_0 = x_0, \dots, X_n = x_n)$$

$$= \mathbb{P}\Big(\bigcup_{x_{n+1} \in E} \{X_{n+k+1} = x_{n+k+1}\} \mid X_0 = x_0, \dots, X_n = x_n\Big)$$

$$= \sum_{x_{n+1} \in E} \mathbb{P}(X_{n+k+1} = x_{n+k+1}, X_{n+1} = x_{n+1} \mid X_0 = x_0, \dots, X_n = x_n)$$

$$= \sum_{x_{n+1} \in E} \mathbb{P}(X_{n+k+1} = x_{n+k+1} \mid X_0 = x_0, \dots, X_{n+1} = x_{n+1}) \cdot \mathbb{P}(X_{n+1} = x_{n+1} \mid X_0 = x_0, \dots, X_n = x_n)$$

$(\mathbb{P}(A \cap B \mid C) = \mathbb{P}(A \mid B \cap C) \cdot \mathbb{P}(B \mid C))$

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

ipotesi induttiva + (M)

$$(*) = \sum_{x_{n+1} \in E} \mathbb{P}(X_{n+k+1} = x_{n+k+1} \mid X_{n+1} = x_{n+1}, X_n = x_n) \cdot \mathbb{P}(X_{n+1} = x_{n+1} \mid X_n = x_n)$$

$$= \sum_{x_{n+1} \in E} \mathbb{P}(X_{n+k+1} = x_{n+k+1}, X_{n+1} = x_{n+1} \mid X_n = x_n)$$

$$= \mathbb{P}(X_{n+k+1} = x_{n+k+1} \mid X_n = x_n)$$

Resta da verificare l'uguaglianza (*):

$$\mathbb{P}(X_{n+k+1} = x_{n+k+1} \mid X_{n+1} = x_{n+1}) = \mathbb{P}(X_{n+k+1} = x_{n+k+1} \mid X_{n+1} = x_{n+1}, X_n = x_n)$$

$$\Leftrightarrow \mathbb{P}(X_{n+k+1} = x_{n+k+1} \mid X_{n+1} = x_{n+1}) \cdot \mathbb{P}(X_{n+1} = x_{n+1}, X_n = x_n)$$

$$= \mathbb{P}(X_{n+k+1} = x_{n+k+1}, X_{n+1} = x_{n+1}, X_n = x_n) \cdot \mathbb{P}(X_{n+1} = x_{n+1})$$

... Lo deriviamo da una proprietà generale ...

Proposizione (Proprietà di Markov generalizzata):

Supponiamo che (M) valga per $(X_n)_{n\geq 0}$ e fissiamo $n, k \geq 0$, $x \in E$. Definiamo le notazioni $X_0^n = (X_0, \dots, X_n)$, $X_n^{n+k} = (X_n, \dots, X_{n+k})$. Allora $\forall A \subseteq E^{n+1}$, $\forall B \subseteq E^{k+1}$ si ha $\mathbb{P}(X_n^{n+k} \in B | X_0^n \in A, X_n = x) = \mathbb{P}(X_n^{n+k} \in B | X_n = x)$ $= \mathbb{P}(X_0^k \in B \mid X_0 = x) =: \mathbb{P}_x(X_0^k \in B)$

Osservazione:

Scegliendo $B = \underbrace{E \times E \times \dots \times E}_{k} \times \{y\}$, $A = \{x_0, \dots, x_{n-1}, x\}$ si ha:

$$\mathbb{P}(X_{n+k} = y | X_0 = x_0, \dots, X_{n-1} = x_{n-1}, X_n = x) = \mathbb{P}(X_{n+k} = y | X_n = x)$$

Def. (CM omogenea, Probabilità e Matrice di Transizione):

Sia $(X_n)_{n\geq 0}$ una CM su uno sp. degli eventi E finito o numerabile. Se $\mathbb{P}(X_{n+1} = x \mid X_n = y)$ è ind. da n $\forall x, y \in E$ si dice che $(X_n)_{n\geq 0}$ è OMOGENEA. Le prob.

$$q(x,y) := \mathbb{P}(X_{n+1} = y \mid X_n = x)$$

sono dette PROBABILITÀ DI TRANSIZIONE e la matrice

$$Q := (q(x,y))_{x,y \in E}$$

è detta MATRICE DI TRANSIZIONE DELLA CATENA.

esempio:

Sia $(W_n)_{n\geq 1}$ una succ. di v.a. i.i.d. a valori in $\mathbb{R}$, E finito o numerabile, $\varphi : E \times \mathbb{R} \longrightarrow E$ misurabile.

Definiamo:

$$X_0 \text{ v.a. ind. da } W_n, \quad X_{n+1} = \varphi(X_n, W_{n+1}) \quad \forall n \geq 0$$

Allora $(X_n)_{n\geq 0}$ è una CM omogenea:

$$\mathbb{P}(X_{n+1}=y \mid X_0=x_0,\dots,X_{n-1}=x_{n-1}, X_n=x)$$

$$= \mathbb{P}(\varphi(x,W_{n+1})=y \mid X_0=x_0,\dots,X_{n-1}=x_{n-1},X_n=x)$$

W_{n+1} ind. da $X_0,\dots,X_n$ essendo queste ultime funzioni di $X_0, W_1,\dots,W_n$

$$= \mathbb{P}(\varphi(x,W_{n+1})=y)$$

e allo stesso punto saremmo arrivati partendo da

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

<u>Osservazione:</u>

Quando (M) NON dipende da n (CM omogenea) si ha:

$$\mathbb{P}(X_{n+1}=y \mid X_n=x) = \mathbb{P}(X_1=y \mid X_0=x)$$

Notare che $q(x,y)\geq 0 \wedge \sum_{y\in E} q(x,y) = 1$. Una matrice con tali proprietà è detta MATRICE STOCASTICA.

Sia $(X_n)_{n\geq 0}$ una CM omogenea, siano inoltre $x_0,\dots,x_n$ elementi di E. Calcoliamo:

$$\mathbb{P}(X_n=x_n,\dots,X_0=x_0) = \mathbb{P}(X_n=x_n \mid X_{n-1}=x_{n-1},\dots,X_0=x_0)\cdot \mathbb{P}(X_{n-1}=x_{n-1},\dots,X_0=x_0)$$

$$= \mathbb{P}(X_n=x_n \mid X_{n-1}=x_{n-1})\cdot\mathbb{P}(X_{n-1}=x_{n-1},\dots,X_0=x_0)$$

$$= q(x_{n-1},x_n)\cdot\mathbb{P}(X_{n-1}=x_{n-1},\dots,X_0=x_0) = (\text{induzione})$$

$$= q(x_{n-1},x_n)\cdot\ldots\cdot q(x_0,x_1)\cdot\mathbb{P}(X_0=x_0) = \underbrace{\mathbb{P}(X_0=x_0)}_{\mu_0(x_0)}\prod_{i=0}^{n} q(x_{i-1},x_i)$$

distribuzione iniziale

$$(\Rightarrow \mathbb{P}(X_n=x_n,\dots,X_1=x_1 \mid X_0=x_0)$$

$$= q(x_n,x_m)\cdot\ldots\cdot q(x_0,y_0)$$

Come conseguenza, $\mu_n(x) =: \mathbb{P}(X_n=x) = \sum_{x_0,\dots,x_{n-1}\in E}\mathbb{P}(X_n=x,\dots,X_0=x_0)$

$$= \sum_{x_0,\dots,x_{n-1}\in E}\mu_0(x_0)\, q(x_0,x_0)\cdot\ldots\cdot q(x_{n-1},x) = \sum_{x_0,\dots,x_{n-1}\in E} Q^n(x_0,y)\,\mu_0(x_0)$$

dove $Q^n(x,y)$ é l'entrata x,y della matrice Q^n; $\mu_0 = (\mu_0(x))_{x\in E}$ é interpretato come un vettore riga. In sintesi:

$$\mu_n = \mu_0 Q^n$$

Inoltre, dalla proprietà di Markov generalizzata si ha:

$$\mathbb{P}(X_0^n \in A, X_n^{n+k} \in B \mid X_n = x) = \mathbb{P}(X_0^n \in A \mid X_n = x) \cdot \mathbb{P}(X_n^{n+k} \in B \mid X_n = x)$$

(Passato e futuro sono indipendenti condizionatamente al presente)

Dim.

Mostriamo che $\forall A \in E^{n+1}$, $B \in E^{k+1}$ vale la seguente:

$$\mathbb{P}(X_0^n \in A, \underbrace{X_n = x, X_n^{n+k} \in B) = \mathbb{P}(X_0^n \in A, X_n = x)}_{(*)}\mathbb{P}(X_0^k \in B)$$

Osserviamo che da $(*)$ seguono tutte le affermazioni:

1) $A = E^{n+1}$, dividendo $(*)$ per $\mathbb{P}(X_n = x)$ si ha
$$\mathbb{P}(X_n^{n+k} \in B \mid X_n = x) = \mathbb{P}_x(X_0^k \in B)$$
2) Dividendo $(*)$ per $\mathbb{P}(X_0^n \in A, X_n = x)$ si ha:
$$\mathbb{P}(X_n^{n+k} \in B \mid X_0^n \in A, X_n = x) = \mathbb{P}_x(X_0^k \in B)$$ e
quindi la prima affermazione é dimostrata.
3) l'indipendenza condizionale tra passato e futuro si ottiene dividendo $(*)$ per $\mathbb{P}(X_n = x)$.

Dimostriamo ora $(*)$. Usando l'additività della probabilità, basta mostrare $(*)$ per $A = \{(x_0, \dots, x_n)\}$,

$B = \{(y_0, \dots, y_k)\}$ dato che ogni altro A o B è unione disgiunta di insiemi di tale forma. Quindi dobbiamo dimostrare:

$$\mathbb{P}(X_0^n = x_0^n, X_n = x, X_n^{n+k} = y_0^k) = \mathbb{P}(X_0^n = x_0^n, X_n = x) \cdot \mathbb{P}_x(X_0^k = y_0^k)$$

Se $x_n \neq x$, entrambi i membri sono 0, altrimenti entrambi sono:

$$\underbrace{\mu_0(x_0)\, q(x_0, x_1) \cdots q(x_{n-1}, x)}_{\mathbb{P}(X_0^n = x_0^n, X_n = x)}\ \underbrace{q(x, y_1)\, q(y_1, y_2) \cdots q(y_{k-1}, y_k)}_{\mathbb{P}_x(X_0^k = y_0^k)}$$

$\square$

<u>esempio:</u>

$$\mathbb{P}(X_{n+k} = y \mid X_n = n, X_n = n, X_0^1 \in A)$$
$$= \mathbb{P}(X_{n+k} = y \mid X_n = k) = \mathbb{P}_x(X_k = y) = Q^k(x, y)$$

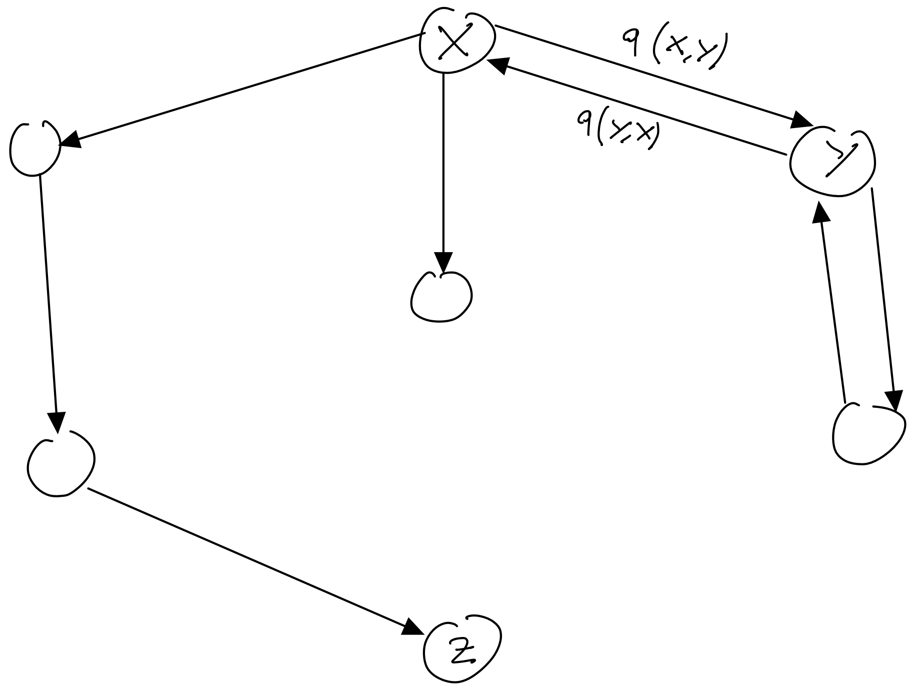
cioè $Q^k(x,y)$ è la probabilità di passare da x a y in k passi.

Sia X una v.a. a valori in E, $(W_n)_{n \geq 1}$ v.a. i.i.d., $\varphi : E \times \mathbb{R} \longrightarrow E$, definiamo $X_{n+1} := \varphi(X_n, W_{n+1})$. Allora $(X_n)_{n \geq 0}$ è una CM omogenea con matrice di transizione $q(x,y) = \mathbb{P}(\varphi(x, W_1) = y)$. Tale risultato vale anche nella "direzione opposta":

Se Q è matrice stocastica, allora $\exists \varphi$ come sopra t.c. la CM $X_{n+1} = \varphi(X_n, U_{n+1})$ ha matrice di transizione Q e (U_n) sono v.a. uniformi su $(0,1)$ indipendenti.

Notiamo che una CM a stati finiti si può rappresentare

Con un grafo pesato diretto:



(Se un arco é assente, allora $q(x,y)=0$)

<u>Def.</u> (Stato Accessibile):

$y \in E$ É ACCESSIBILE da $x \in E$ per una CM di matrice di transizione Q) se $\exists\, m \geq 0$ t.c. $Q^m(x,y)>0$ ovvero $\mathbb{P}(X_m=y \mid X_0=x)>0$. In tal caso si scrive $x \rightsquigarrow y$.

Notiamo che $\rightsquigarrow$ é transitiva:

$$x \rightsquigarrow y \wedge y \rightsquigarrow z \Rightarrow x \rightsquigarrow z$$

Infatti $x \rightsquigarrow y \Rightarrow \exists\, m$ t.c. $Q^m(x,y)>0$, $y \rightsquigarrow z \Rightarrow \exists\, n$ t.c. $Q^n(y,z)>0 \Rightarrow Q^{n+m}(x,z)>0$ dato che:

$$Q^{n+m}(x,z) = Q^m Q^n(x,z) = \sum_w Q^m(x,w)\,Q^n(w,z)$$
$$\geq Q^m(x,y)\,Q^n(y,z)>0$$

<u>Osservazione:</u>

Sia y accessibile da x, allora $\exists\, m$ t.c. $Q^m(x,y)>0$.

$$\Rightarrow Q^m(x,y) = \sum_{x_1,\dots,x_{m-1}} q(x,x_1)\cdot\ldots\cdot q(x_{m-1},y)$$

$\Rightarrow Q^n(x,y)>0 \Leftrightarrow \exists\, x_1,\dots,x_{n-1}$ t.c. $q(x,x_1),\dots,q(x_{n-1},y)>0$

<u>Def.</u> (Stati comunicanti)

$x,y\in E$ COMUNICANO (e scriviamo $x \leftrightarrow y$) se é si ha $x \rightsquigarrow y \wedge y \rightsquigarrow x$. La relazione $\leftrightarrow$ é riflessiva, simmetrica e transitiva, quindi é rel. di equivalenza.

<u>Def.</u> (CM irriducibile):

Una CM si dice IRRIDUCIBILE se $\forall x,y\in E$ si ha $x\leftrightarrow y$

<u>esempi</u>:

1) Urna di Ehrnfest:

Consideriamo 2 urne (A,B) e N palline numerate da 1 ad N. Suddividiamo le palline tra A e B. Scegliamo casualmente (con prob. uniforme) un elemento $i\in\{1,\dots,N\}$ e spostiamo l'i-sima pallina nell'altra urna. Denotiamo con X_n il numero di palline in A dopo n estrazioni. Notiamo che se $X_n=i$, allora $X_{n+1}=i-1 \vee i+1$

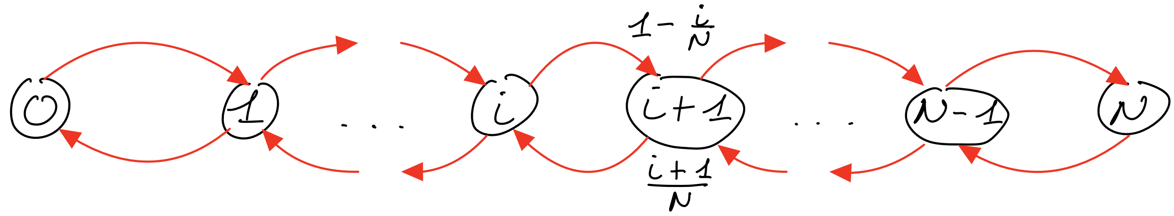
$\Rightarrow \mathbb{P}(X_{n+1}=i-1 \mid X_n=i) = \frac{i}{N}$

$\Rightarrow \mathbb{P}(X_{n+1}=i+1 \mid X_n=i) = 1-\frac{i}{N}$

$\Rightarrow$ si ha:

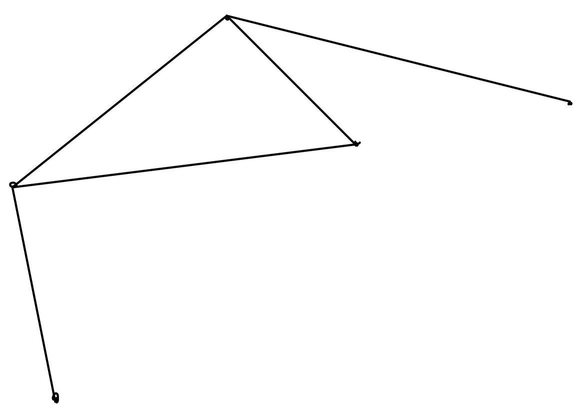
$$Q=\begin{pmatrix} 0 & 1 & 0 & \cdots & \cdots & 0\\ \frac{1}{N} & 0 & \frac{N-1}{N} & \cdots & \cdots & 0\\ & \frac{2}{N} & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ 0 & \cdots & \cdots & \cdots & 1 & 0 \end{pmatrix}$$

$\Rightarrow$ disegnamo il grafo:



$\Rightarrow$ la CM è irriducibile

2) Sia (E, V) un grafo NON diretto, ovvero E insieme finito, $V = E \times E$ simmetrico $\left((x,y) \in V \Rightarrow (y,x) \in V \right)$



Definiamo:

$$q(x,y) = \begin{cases} \frac{1}{\deg(x)} & \text{se } (x,y) \in V \\ 0 & \text{altrimenti} \end{cases}$$

dove $\deg(x) := |\{ y : (x,y) \in V \}|$

(conveniamo che $(x,y) \notin V$ $\forall x \in E$ e se $\deg(x) = 0$ allora $q(x,y) = 1$).

$\Rightarrow$ Una CM con tale Q è detta PASSEGGIATA ALEATORIA sul grafo (E, V)

N.B.
Tale catena è irriducibile se e solo se il grafo è connesso.

Tempi di Ingresso:

Sia (X_n) una CM omogenea con matrice di transizione

Q, e sia $y \in E$. Definiamo:

$$T_y := \min\{n \geq 1 : X_n = y\} \quad (\min \phi = +\infty)$$

Notare che $\{T_y > m\} = \{X_1 \neq y, \dots, X_m \neq y\} \in \sigma(X_1, \dots, X_m)$, quindi T_y é un tempo d'arresto per la filtrazione generata dal processo (X_n). Notare che NON necessariamente $T_y < +\infty$ q.c. Ad esempio se $E = \{1, 2\}$, $Q = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, si ha $\mathbb{P}_1(T_2 = +\infty) = 1$

<u>Proposizione</u>

Supponiamo che la CM sia irriducibile. Allora:

$$\mathbb{P}_x(T_y < +\infty) = 1 \wedge \mathbb{E}_x(T_y) < +\infty \quad \forall x, y \in E$$

<u>Dim.:</u>

Dato che la catena é irriducibile, $\forall x, y \in E \; \exists m_{x,y} \geq 1$ t.c. $Q^{m_{x,y}}(x, y) > 0$. Sia $m = \max\limits_{x,y}\{m_{x,y}\}$, e sia $\delta = \min\limits_{x,y \in E}\{Q^{m_{x,y}}(x,y)\} > 0$. Calcoliamo:

$$\mathbb{P}_x(T_y > m) = 1 - \mathbb{P}_x(T_y \leq m) = 1 - \mathbb{P}_x(\exists\, 1 \leq i \leq m : X_i = y)$$
$$\leq 1 - \mathbb{P}_x(X_{m_{x,y}} = y) = 1 - Q^{m_{x,y}}(x, y) \leq 1 - \delta$$

Mostriamo per induzione su $K \geq 1$ che $\mathbb{P}_x(T_y > Km) \leq (1-\delta)^K$

$K = 1$: ✓

$K - 1 \longrightarrow K$:

Sia $A_i := \{X_{(i-1)m+1} \neq y, \dots, X_{im} \neq y\}$

$$\Rightarrow \{T_y > Km\} = \bigcap_{i=1}^{K} A_i \Rightarrow \mathbb{P}(A_i \mid X_{(i-1)m} = x') \underset{\uparrow}{=} \mathbb{P}_{x'}(A_1)$$

omogeneitá temporale

$$= \mathbb{P}_{x'}(T_y > m) \leq 1 - \delta$$

Allora:

$$\mathbb{P}_x(T_y > km) = \mathbb{P}_x(A_1 \cap \ldots \cap A_k)$$
$$= \sum_{\substack{x' \in E \\ x' \neq y}} \mathbb{P}_x(A_1 \cap \ldots \cap A_k \cap \{X_{(k-1)m} = x'\})$$
$$= \sum_{\substack{x' \in E \\ x' \neq y}} \mathbb{P}_x(A_k | \{X_{(k-1)m} = x'\} \cap A_1 \cap \ldots \cap A_{k-1}) \cdot \mathbb{P}(\{X_{(k-1)m} = x'\} \cap A_1 \cap \ldots \cap A_{k-1})$$
$$= \sum_{x' \in E} \mathbb{P}_x(A_k | X_{(k-1)m} = x') \cdot \mathbb{P}(\{X_{(k-1)m} = x'\} \cap A_1 \cap \ldots \cap A_{k-1})$$

Proprietà di Markov generalizzata

$$\leq (1-\delta)\, \mathbb{P}(A_1 \cap \ldots \cap A_{k-1}) = (1-\delta)\, \mathbb{P}(T_y > (k-1)m)$$
$$\leq (1-\delta)(1-\delta)^{k-1} = (1-\delta)^k$$

ipotesi induttiva

$\Rightarrow$ si ha quindi $\mathbb{P}_x(T_y > km) \leq (1-\delta)^k$

$$\Rightarrow \mathbb{P}_x(T_y = +\infty) = \mathbb{P}_x\Big(\bigcap_{k \geq 1} \{T_y > km\}\Big) = \lim_{k \to +\infty} \mathbb{P}(T_y > km)$$
$$\leq \lim_{k \to +\infty} (1-\delta)^k = 0 \quad \text{dato che } \delta > 0 \Rightarrow 1 - \delta < 1$$

Ricordiamo che se a X è a valori in $\mathbb{N}$, allora:
$$\mathbb{E}(X) = \sum_{n=0}^{+\infty} \mathbb{P}(X > n)$$

Quindi:
$$\mathbb{E}_x(T_y) = \sum_{n=0}^{+\infty} \mathbb{P}_x(T_y > n) = \sum_{k=0}^{+\infty} \sum_{n=km}^{(k+1)m-1} \mathbb{P}_x(T_y > n)$$
$$\leq \sum_{k=0}^{+\infty} \sum_{n=km}^{(k+1)m-1} \mathbb{P}_x(T_y > km) \leq m \sum_{k=0}^{+\infty} (1-\delta)^k < +\infty$$

$\square$