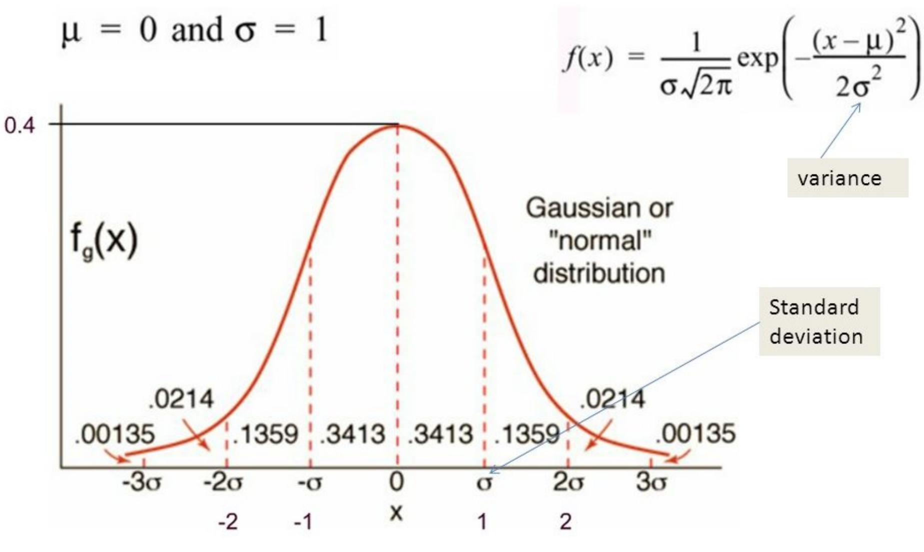
Dispense di
Probabilità e statistica

Emanuele Izzo, 0253052



Corso di laurea triennale Informatica
Università Tor Vergata, Facoltà di Scienze MM.FF.NN.
04/09/2020

Documento realizzato in LaTeX

Indice

Parte I

Probabilità nel discreto e nel continuo

Capitolo 1

Introduzione

1.1 Fenomeno casuale

Un fenomeno casuale è un fenomeno il cui esito non è certo; esempio:

- Il numero che esce lanciando un dado (caso discreto);
- Il tempo di funzionamento di una lampadina (caso continuo).

1.2 Insieme dei "risultati possibili" o "eventi elementari"

$\Omega := \{x_1, x_2, ..., x_n\}$ è l'insieme di tutti i risultati possibili del fenomeno casuale, e gli eventi saranno individuati da sottoinsieme di Ω; esempio:

- $\Omega := \{1, 2, 3, 4, 5, 6\}, E =: \{2, 4, 6\}$
  [oppure $E := \{\text{esce un numero pari}\}$] (caso discreto);
- $\Omega := [0, \infty), E := [5, \infty)$ (caso continuo).

$\mathcal{A}$ è una famiglia di sottoinsiemi di Ω che rappresentano gli eventi.

1.3 σ-algebra

Una famiglia di sottoinsiemi di Ω è detta σ-algebra se valgono le seguenti proprietà:

1. $\Omega \in \mathcal{A}$;
2. $A \in \mathcal{A} \Rightarrow A^{\mathrm{C}} \in \mathcal{A}$;
3. $\{A_n\}_{n\geq 1} \in \mathcal{A} \Rightarrow \displaystyle\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$.

Corollario

a. Dalle proprietà [1] e [2] abbiamo che $\emptyset \in \mathcal{A}$;

b. $\forall k \geq 1, A_1, A_2, ..., A_k \in \mathcal{A} \Rightarrow A_1 \cup A_2 \cup ... \cup A_k \in \mathcal{A}$.

c. Dalle proprietà [1] e [3] abbiamo che $\{A_n\}_{n\geq 1} \Rightarrow \bigcap_{n=1}^{\infty} A_n \in \mathcal{A}$;

d. $\forall k \geq 1, A_1, A_2, ..., A_k \in \mathcal{A} \Rightarrow A_1 \cap A_2 \cap ... \cap A_k \in \mathcal{A}$.

1.4 Spazio di probabilità

Sia Ω un insieme, sia $\mathcal{A}$ una σ-algebra, allora $(\Omega, \mathcal{A}, \mathrm{P})$ è uno spazio di probabilità se $\mathrm{P} : \mathcal{A} \to [0, \infty)$, dove P è detta misura di probabilità, soddisfa le seguenti proprietà:

1. $\mathrm{P}(\Omega) = 1$;

2. $\forall \{A_n\}_{n\geq 1} \subset \mathcal{A} | \forall n \neq m : A_n \cap A_m \neq \emptyset \Rightarrow \mathrm{P}(\sum_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mathrm{P}(A_n)$.

Corollario

a. $\mathrm{P}(\emptyset) = 0$, infatti $\forall n \geq 1, A_n = \emptyset$: questo si può dimostrate tramite la proprietà [2], infatti $\mathrm{P}(\sum_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mathrm{P}(A_n) \Rightarrow \mathrm{P}(\sum_{n=1}^{\infty} \emptyset) = \sum_{n=1}^{\infty} \mathrm{P}(\emptyset)$, e se fosse $\mathrm{P}(\emptyset) > 0$, il secondo membro sarebbe infinito (impossibile), quindi $\mathrm{P}(\emptyset) = 0$;

b. $\forall k \geq 1, B_1, B_2, ..., B_k \in \mathcal{A} | \forall n \neq m, A_n \cap A_m = \emptyset \Rightarrow$
$\Rightarrow \mathrm{P}(\sum_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} \mathrm{P}(B_n)$: infatti basta considerare la proprietà [2] con $B_n = A_n$ se $1 \leq n \leq k$ e $B_n = \emptyset$ se $n > k$;

c. Prendiamo $k = 2$ ed $E, F \in \mathcal{A}$, e poniamo $B_1 = E \cap F, B_2 = E \cap F^{\mathrm{C}}$: se calcoliamo $\mathrm{P}(B_1 \cup B_2) = \mathrm{P}(B_1) + \mathrm{P}(B_2) \Rightarrow$
$\Rightarrow \mathrm{P}(E) = \mathrm{P}(E \cap F) + \mathrm{P}(E \cap F^{\mathrm{C}})$, e se poniamo $E = \Omega$ abbiamo $\mathrm{P}(\Omega) = \mathrm{P}(\Omega \cap F) + \mathrm{P}(\Omega \cap F^{\mathrm{C}}) \Rightarrow 1 = \mathrm{P}(F) + \mathrm{P}(F^{\mathrm{C}})$, da cui ricaviamo che $\mathrm{P}(F) = 1 - \mathrm{P}(F^{\mathrm{C}})$ e $\mathrm{P}(F^{\mathrm{C}}) = 1 - \mathrm{P}(F)$;

d. $E, F \in \mathcal{A} | E \subset F \Rightarrow \mathrm{P}(F) = \mathrm{P}(E) + \mathrm{P}(F/E)$, e tramite la il punto [c], ponendo $k = 2, B_1 = E, B_2 = F/E$, abbiamo che $\mathrm{P}(E) \leq \mathrm{P}(F)$;

e. Tramite i punti [1], [a], [c] e [d], possiamo dire che $\mathrm{P}(:)\mathcal{A} \to [0, 1]$ anziché $\mathrm{P}(:)\mathcal{A} \to [0, \infty)$;

f. Siano $A, B \in \mathcal{A}$, abbiamo che $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)$;

g. Generalizzando il punto [f], siano $A, B, C \in \mathcal{A}, \mathrm{P}(=)\mathrm{P}(A) + \mathrm{P}(B) +$
$+ \mathrm{P}(C) - \mathrm{P}(A \cap B) - \mathrm{P}(A \cap C) - \mathrm{P}(B \cap C) + \mathrm{P}(A \cap B \cap C)$.

1.5 Spazio di probabilità uniforme discreto

Sia Ω un insieme finito, sia $\mathcal{A} = \mathcal{P}(\Omega)$ l'insieme delle parti (tutti i possibili sottoinsiemi di Ω), si deve avere $\forall \omega \in \Omega, \mathrm{P}(\{\omega\}) = p$ costante. Segue necessariamente che $p = \frac{1}{\#\Omega}$ $(\#\Omega = |\Omega|)$, infatti
$1 = \mathrm{P}(\Omega) = \mathrm{P}(\bigcup_{\omega\in\Omega}\{\omega\}) = \sum_{\omega\in\Omega}\mathrm{P}(\{\omega\}) = \#\Omega \cdot p \Rightarrow p = \frac{1}{\#\Omega}$. Consideriamo un generico $A \in \Omega, A = \{a_1, a_2, ..., a_k\}$ con $\#A = k$, abbiamo che
$\mathrm{P}(A) = \mathrm{P}(\bigcup_{n=1}^{k}\{a_n\}) = \sum_{n=1}^{k}\mathrm{P}(\{a_n\}) = \sum_{n=1}^{k}\frac{1}{\#\Omega} = \frac{k}{\#\Omega} = \frac{\#A}{\#\Omega}$. Osserviamo che $\mathrm{P}(A) = 0$ se $A = \emptyset$ e $\mathrm{P}(A) = 1$ se $A = \Omega$.

1.6 Probabilità condizionata (condizionale)

Sia $(\Omega, \mathcal{A}, \mathrm{P})$ uno spazio di probabilità, si suppone che un certo evento $B \in \mathcal{A}$ si è verificato, dobbiamo costruire un nuovo spazio di probabilità $(\Omega, \mathcal{A}, \mathrm{P}_B)$, dove $\forall A \in \mathcal{A}, \mathrm{P}_B(A) = \mathrm{P}(A|B) = c_B \cdot \mathrm{P}(A \cap B)$. Il valore c_B deve essere tale che $\mathrm{P}(\Omega|B) = 1$, da cui abbiamo che
$1 = c_B \cdot \mathrm{P}(\Omega \cap B) = c_B \cdot \mathrm{P}(B) \Rightarrow c_B = \frac{1}{\mathrm{P}(B)}$ dato che $\Omega \cap B = B$. Quindi abbiamo che $\mathrm{P}(A|B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)}$ solo se $\mathrm{P}(B) \neq 0$, ossia se $B \neq \emptyset$.

Corollario

a. $\mathrm{P}(B|B) = \frac{\mathrm{P}(B \cap B)}{\mathrm{P}(B)} = 1$;

b. Se $A \cap B = \emptyset$, allora $\mathrm{P}(A|B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} = 0$;

c. Se $\mathrm{P}(B) = 1$, allora $\forall A \in \mathcal{A}, \mathrm{P}(A|B) = \mathrm{P}(A)$, infatti
$\mathrm{P}(A|B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} = \mathrm{P}(A \cap B) = \mathrm{P}(A) - \mathrm{P}(A \cap B^{\mathrm{C}}) = \mathrm{P}(A)$,
poiché $\mathrm{P}(A \cap B^{\mathrm{C}}) \leq \mathrm{P}(B^{\mathrm{C}})$ perché $A \cap B^{\mathrm{C}} \subset B^{\mathrm{C}}$ e sarà uguale a 0 dato che $\mathrm{P}(B^{\mathrm{C}}) = 1 - \mathrm{P}(B) = 1 - 1 = 0$.

Parte II

Le catene di Markov

Capitolo 2

Introduzione

Le catene di Markov sono un esempio di processo stocastico $\{X_t : t \in \mathrm{T}\}$, dove $\mathrm{T} \subset \mathbb{R}^+$ è l'insieme dei tempi, definite come:

- $\mathrm{T} \subset \mathbb{N}$;
- Le variabili aleatorie $\{X_t : t \in \mathrm{T}\}$ sono a valori compresi in un insieme numerabile E, detto spazio degli stati;
- $\forall n \geq 1, \forall i, j, i_{n-1}, ..., i_1 \in E$
  $P(x_{n+1} = j | X_n = i, X_{n-1} = i_{n-1}, ..., X_i = i_1) = P(x_{n+1} = j | X_n = i)$
  dove $P(A|B, C) = P(A|B \cap C)$, e nei casi le probabilità condizionate sono ben definite.

2.1 Catena omogenea

Si dice che la catena è omogenea se i valori $p_{ij}(n)$ non dipendono da n. In tal caso si scriverà p_{ij} anziché $p_{ij}(n)$.

Proposizione

$\forall n \geq 1, \forall i \subset E, \sum_{j \in \mathrm{E}} p_{ij} = 1$ dove $\sum_{j \in \mathrm{E}} p_{ij}$ è la somma magli elementi della riga i della matrice $\mathrm{P}(n) = (p_{ij}(n))_{i,j \in \mathrm{E}}$ per ogni n fissato. Sapendo che $\bigcup_{j \in \mathrm{E}} \{X_{n+1} = j\} = \Omega$, abbiamo che
$\sum_{j \in \mathrm{E}} p_{ij} = \sum_{j \in \mathrm{E}} P(X_{n+1} = j | X_n = i) =$
$= P(\bigcup_{j \in \mathrm{E}} \{X_{n+1} = j\} | X_n = i) = P(\Omega | X_n = i) = 1.$

Osservazione

La matrice $\mathrm{P}(n) = (p_{ij}(n))_{i,j \in \mathrm{E}}$ ha autovalore 1 con autovettore composto da tutti 1, infatti

$\mathrm{E} = \{1, 2, ..., r\}$

$$\begin{bmatrix} p_{11}(n) & p_{12}(n) & \dots & p_{1r}(n) \\ p_{21}(n) & p_{22}(n) & \dots & p_{2r}(n) \\ \vdots & \vdots & \ddots & \vdots \\ p_{r1}(n) & p_{r2}(n) & \dots & p_{rr}(n) \end{bmatrix} \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

Esempio con E finito: rovina del giocatore

Abbiamo una moneta X definita in questo modo:

$$\begin{cases} P(X = \mathrm{T}) = p \text{ probabilità che esca testa} \\ P(X = \mathrm{C}) = 1 - p \text{ probabilità che esca croce} \end{cases}$$

E abbiamo due giocatori A e B, con capitale rispettivamente $a \geq 1$ e $b \geq 2$. Poniamo di porre i seguenti casi:

$$\begin{cases} \text{Esce testa} \to A \text{ prende un'unità da } B \\ \text{Esce croce} \to B \text{ prende un'unità da } A \end{cases}$$

Il gioco viene descritto bene da $X_n = \{\text{capitale di } A \text{ dopo } n \text{ lanci}\}$.

Matrice di transizione

$\mathrm{E} = \{0, 1, 2, ..., a+b-2, a+b-1, a+b\}$

$$\mathrm{P}(n) = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 1-p & 0 & p & \dots & 0 & 0 & 0 \\ 1 & 1-p & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 1 & 0 & 0 & \dots & 0 & p & 0 \\ 1 & 0 & 0 & \dots & 1-p & 0 & p \\ 1 & 0 & 0 & \dots & 0 & 0 & 1 \end{pmatrix}$$

2.2 Leggi congiunte (caso omogeneo)

Vogliamo calcolare $P(X_{n_1} = i_1, X_{n_2} = i_2, ..., X_{n_k} = i_k)$, dove $P(A, B) = P(A \cup B)$, con $i_1, i_2, ..., i_k \in \mathrm{E}$ e $1 \leq n_1 \leq n_2 \leq ... \leq n_k$ interi. Per rispondere a questa domanda bisogna fare riferimento alla matrice di transizione e m passi $P^{(m)} = (p_{ij}^{(m)}) = P(X_m = j | X_0 = i)$. C'è un legame tra $\mathrm{P} = (p_{ij})$ e $\mathrm{P}^{(m)} = (p_{ij}^{(m)})$ (sapendo che per $m = 1$ devono coincidere)?

$$p_{ij}^{(m)} = P(X_m = j | X_0 = i) = \sum_{h \in \mathrm{E}} P(\{X_m = j\} \cup \{X_{m-1} = h\} | X_0 = i) =$$

$$= \sum_{h \in \mathrm{E}} P(X_m = j | \{X_{m-1} = k\} \cup \{X_0 = i\}) P(X_{m-1} = h | X_0 = i) =$$

$$= \sum_{h \in \mathrm{E}} \underbrace{P(X_m = j | X_{m-1} = h)}_{= p_{hj}} \underbrace{P(X_{m-1} = h | X_0 = 1)}_{p_{ih}^{(m-1)}} = \mathrm{P} * \mathrm{P}^{(m-1)}$$

In conclusione, per arbitrarietà di m, $\mathrm{P}^{(m)} = \mathrm{P}^{(m-1)}\mathrm{P} \Rightarrow \mathrm{P} = \underbrace{\mathrm{PP}...\mathrm{P}}_{m \text{ volte}}$. Pertanto, la formula per le leggi congiunte è

$\forall k \geq 1, \forall i_1, i_2, ..., 1_k \in \mathrm{E}, \forall n_1 \leq n_2 \leq ... \leq n_k textinteri,$

$$P(X_{n_1} + i_1, X_{n_2} + i_2, ..., X_{n_k} + i_k) = \sum_{j \in \mathrm{E}} P(X_0 = j) p_{ji_1}^{(n_1)} p_{i_1 i_2}^{(n_2 - n_1)} ... p_{i_{k-1} i_k}^{(n_k - n_{k-1})}$$

Osservazione

Punto di vista righe per colonne
$v^{(0)} = (P(X_0 = i))_{i \in \mathrm{E}}$ vettore riga 0
$\forall i \in \mathrm{E}, P(X_n = j) = \sum_{j \in \mathrm{E}} P(X_0 = j) p_{ij}^{(n)}$
$v^{(n)} = (P(X_n = i))_{i \in \mathrm{E}}$ vettore riga n
$v^{(n)} = v^{(0)} \mathrm{P}^{(n)}$
Ovviamente i vettori sono con coordinate non negative a somma 1 (stesse proprietà delle densità discrete).