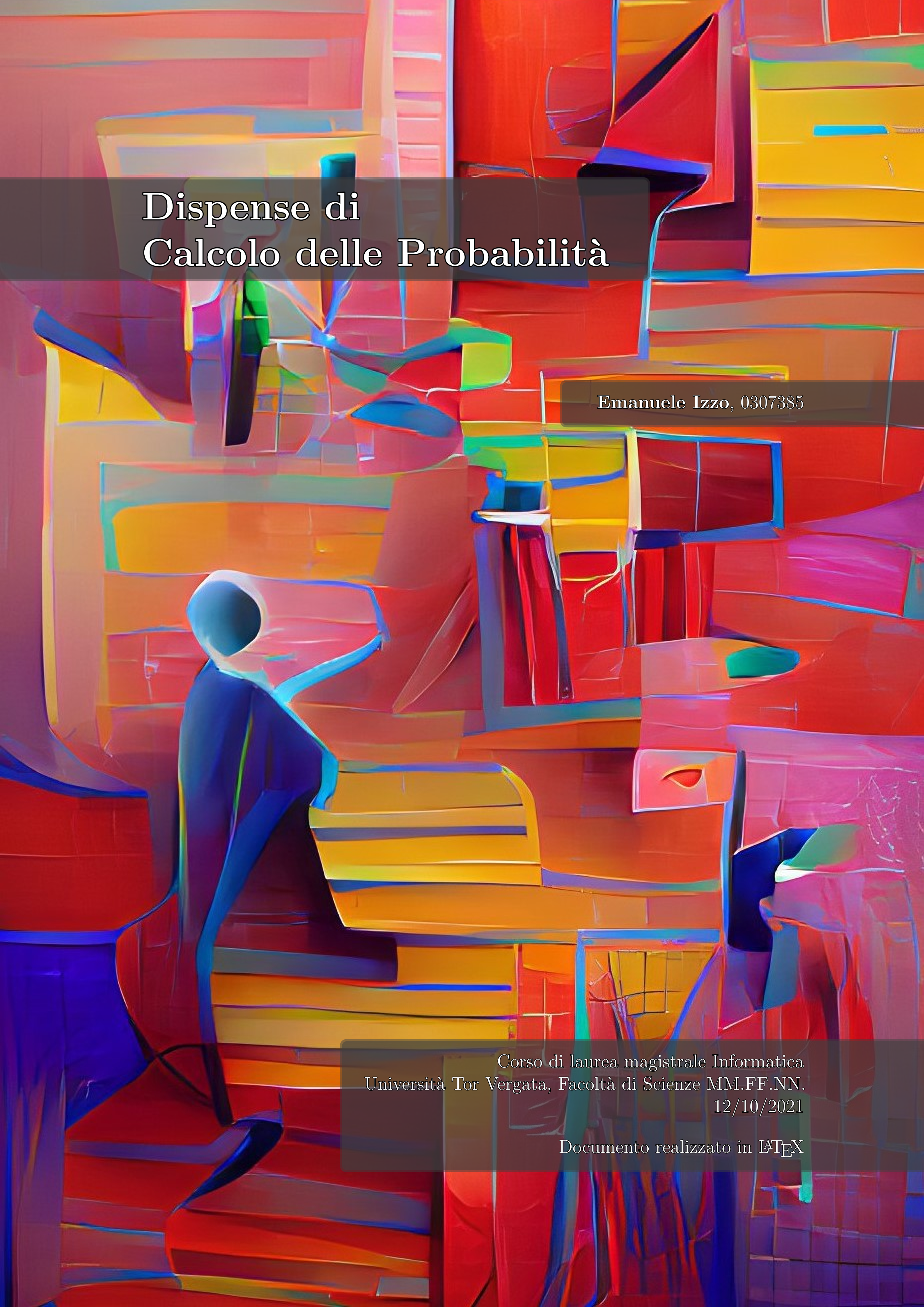

Dispense di Calcolo delle Probabilità

Emanuele Izzo, 0307385

Corso di laurea magistrale Informatica
Università Tor Vergata, Facoltà di Scienze MM.FF.NN.
12/10/2021

Documento realizzato in LaTeX

Indice

Parte I

Calcolo delle probabilità

Capitolo 1

Introduzione

1.1 Gli eventi e la probabilità

Definizione (1.1): Uno **spazio di probabilità** è una terna $(\Omega, \mathcal{F}, P)$ dove:

- Ω è un insieme non vuoto (che rappresenta l'insieme dei possibili risultati del fenomeno aleatorio in esame).
- $\mathcal{F}$ è una famiglia di sottoninsiemi di Ω detti **eventi**.
- P è una *misura di probabilità*.

In genere si suppone che la famiglia $\mathcal{F}$ abbia proprietà ragionevoli, cioè che $\Omega \in \mathcal{F}$ e che presenta una *chiusura rispetto alle operazion insiemistiche*: ciò permette di identificare l'insieme $\mathcal{F}$ come una **σ-algebra**. Nel caso in cui Ω sia un insieme discreto, cioè finito o numerabile, possiamo presupporre che $\mathcal{F}$ sia l'**insieme delle parti** di Ω.

Definizione (1.2): Una **misura di probabilità** P è una funzione $P : \mathcal{F} \to [0, \infty)$ tale che $P(\Omega) = 1$ e, per ogni successione di eventi $\{E_n\}_n$ disgiunti a due a due (cioè $E_i \cap E_j = \emptyset$ per $i \neq j$), si ha $P(\cup_n E_n) = \sum_n P(E_n)$.

Tramite il teorema appena definito è possibile dimostrare le seguenti affermazioni:

- $\emptyset \in \mathcal{F}$ e $P(\emptyset) = 0$.
- Per ogni scelta di $E, F \in \mathcal{F}$ tali che $E \subset F$ si ha che $P(E) \leq P(F)$.
- Per ogni scelta di $E \in \mathcal{F}$ si ha che $P(E) \in [0, 1]$ ed inoltre si ha che $P(E) = 1 - P(E^C)$ e $P(E^C) = 1 - P(E)$.
- Per ogni famiglia finita di eventi $\{E_n\}_n$ disgiunti a due a due, si ha che $P(\cup_n E_n) = \sum_n P(E_n)$.
- Per ogni scelta di $E, F \in \mathcal{F}$ si ha che $P(E) = P(E \cap F) + P(E \cap F^C)$.

Notare che nel caso in cui $\Omega = \cup_n \{w_n\}$, per n che appartiene ad un insieme finito o numerabile, l'insieme dei valori $\{P(\{w_n\})\}_n$ consente di determinare il valore di $P(E)$ per ogni $E \in \mathcal{F}$; infatti:

$$E = \cup_{n:\omega_n \in E}\{\omega_n\} \text{ implica } P(E) = \sum_{n:\omega_n \in E} P(\{\omega_n\})$$

Questo perché $E \cup_{n:\omega_n \in E} \{\omega_n\}$ è un'unione di insieme disgiunti a due a due. Spesso nei risultati generati che vengono forniti si sottintende di avere uno spazio di probabilità $(\Omega, \mathcal{F}, P)$ assegnato, non necessariamente con Ω finito o numerabile.

1.2 Unione, intersezione e dipendenza tra eventi

Lemma (1.2a): Per ogni scelta di $E_1, E_2 \in \mathcal{F}$ si ha $P(E_1 \cup E_2) = P(E_1) + P(E_2) - P(E_1 \cap E_2)$

Lemma (1.2a): Osserviamo che $E_1 \cup E_2 = (E_1 \cap E_2^C) \cup (E_1 \cap E_2) \cup (E_1^C \cap E_2)$, la quale è l'unione di eventi disgiunti a due a due, pertanto:

$$P(E_1 \cup E_2) = P(E_1 \cap E_2^C) + P(E_1 \cap E_2) + P(E_1^C \cap E_2)$$

Allora, sommando e sottraendo $P(E_1 \cap E_2)$, si ha:

$$P(E_1 \cup E_2) = \underbrace{P(E_1 \cap E_2^C) + P(E_1 \cap E_2)}_{P(E_1)} + \underbrace{P(E_1 \cap E_2) + P(E_1^C \cap E_2)}_{P(E_2)} - P(E_1 \cap E_2)$$

Lemma (1.2b) (Union Bound): Per ogni famiglia finita o successione $\{E_n\}_n$ di elementi di $\mathcal{F}$ si ha $P(\cup_n E_n) \le \sum_n P(E_n)$

Lemma (1.2b): Consideriamo i seguenti eventi $\{\hat{E}_n\}_n$:

$$\hat{E}_1 = E_1$$
$$\hat{E}_n = E_n \cap (E_1 \cup E_2 \cup ... \cup E_{n-1})^C \text{ per } n \ge 2$$

Allora si ha $\cup_n E_n = \cup_n \hat{E}_n$ e gli eventi $\{\hat{E}_n\}_n$ sono disgiunti a due a due. Quindi, essendo anche per costruzione $\hat{E}_n \subset E_n$ per ogni scelta di n, si ha:

$$P(\cup_n E_n) = P(\cup_n \hat{E}_n) = \sum_n P(\hat{E}_n) \le \sum_n P(E_n)$$

Lemma (1.2c) (Principio di inclusione-esclusione): Per ogni famiglia finita o successione $\{E_n\}_n$ di elementi di $\mathcal{F}$ si ha che:

$$P(E_1 \cup E_2 \cup ... \cup E_n) = \sum_i P(E_i) - \sum_{i<j} P(E_i \cap E_j) + \\ + \sum_{i<j<k} P(E_i \cap E_j \cap E_k) + ... + (-1)^n P)E_1 \cap E_2 \cap ... \cap E_n)$$

Procediamo per induzione su n:

- Per $n = 2$ il risultato è vero per il lemma (1.2a).
- Per $n \geq 3$ applichiamo il passo induttivo pnendo che la formula sia vera per $n - 1$. Con riferimento al caso $n = 2$ per $E_1 \cup E_2 \cup ... \cup E_{n-1}$ e E_n si ha:

$$\begin{aligned} &P(E_1 \cup E_2 \cup ... \cup E_{n-1} \cup E_n) = \\ &\quad = P(E_1 \cup E_2 \cup ... \cup E_{n-1}) + P(E_n) - P((E_1 \cup E_2 \cup ... \cup E_{n-1} \cap E_n)) = \\ &\quad = P(E_1 \cup E_2 \cup ... \cup E_{n-1}) + P(E_n) - P(\cup_{i=1}^{n-1}(E_i \cap E_n)) \end{aligned}$$

A questo punto si usa la formulta per il caso $n - 1$ per il primo e il terzo termine. Quindi si ha:

$$\begin{aligned} &P(E_1 \cup E_2 \cup ... \cup E_{n-1} \cup E_n) = \\ &\quad = \sum_i P(E_i) - \sum_{i<j} P(E_i \cap E_j) + \sum_{i<j<k} P(E_i \cap E_j \cap E_k) + ... \\ &\quad\quad ... + (-1)^{n-1-1} P(\cap_{i=1}^{n-1} E_i) + P(E_n) + \\ &\quad\quad - \{\sum_i P(E_i \cap E_n) - \sum_{i<j} P(E_i \cap E_n \cap E_j \cap E_n) + ... \\ &\quad\quad ... + (-1)^{n-1-1} P(\cap_{i=1}^{n}(E_i \cap E_n))\} = \\ &\quad = \sum_i P(E_i) - \sum_{i<j} P(E_i \cap E_j) + \sum_{i<j<k} P(E_i \cap E_j \cap E_k) + ... \\ &\quad\quad ... + (-1)^{n-1-1} P(\cap_{i=1}^{n-1} E_i) + P(E_n) + \\ &\quad\quad - \{\sum_i P(E_i \cap E_n) - \sum_{i<j} P(E_i \cap E_j \cap E_n) + ... \\ &\quad\quad ... + (-1)^{n-1-1} P(\cap_{i=1}^{n} E_i)\} \end{aligned}$$

A questo punto non resta che combinare i seguenti termini:

$$\begin{aligned}
&\sum_i P(E_i) \text{ e } P(E_n) \\
&\sum_{i<j} P(E_i \cap E_j) \text{ e } \sum_i P(E_i \cap E_n) \\
&\sum_{i<j<k} P(E_i \cap E_j \cap E_k) \text{ e } \sum_{i<j} P(E_i \cap E_j \cap E_n) \\
&... \\
&-(-1)^{n-1-1} P(E_1 \cap E_2 \cap ... \cap E_{n-1} \cap E_n) = \\
&\quad = (-1)^{n-1} P(E_1 \cap E_2 \cap ... \cap E_{n-1} \cap E_n)
\end{aligned}$$

Definizione (1.3) (Indipendenza tra eventi): Due eventi $E, F \in \mathcal{F}$ sono indpendenti se $P(E \cap F) = P(E)P(F)$. In generale, data un'arbitraria famiglia di eventi $\{E_n\}_n$, abbiamo eventi indipendenti se, presa una qualsiasi sottofamiglia finita con almeno due eventi $E_{i_1}, E_{i_2}, ..., E_{i_k}$ con $k \geq 2$, si ha che:

$$P(E_{i_1} \cap E_{i_2} \cap ... \cap E_{i_k}) = P(E_{i_1})P(E_{i_2})...P(E_{i_k})$$

Definizione (1.4) (Probabilità condizionata): Dati due eventi $E, F \in \mathcal{F}$ tali che $P(F) \neq 0$, la **probabilità condizionata** di E sapendo che si è verificato F è definita come segue:

$$P(E|F) = \frac{P(E \cap F)}{P(F)}$$

Spesso viene usata la terminologia semplificata *"probabilità condizionata di E dato F"*. In generale è come se si passasse da Ω a F e quindi si normalizza dividendo per $P(F)$ in modo che $P(\Omega|F) = 1$. Se $P(F) \neq 0$, allora E ed F sono **indipendenti** se e solo se $P(E|F) = P(E)$. È possibile usare la formula della probabilità condizionata per calcolare la probabilità dell'intersezione di due eventi:

$$P(E \cap F) = P(E|F)P(F)$$

Questa uguaglianza vale anche se $P(F) = 0$ perché. anche se non possiamo definire $P(E|F)$, possiamo dire che:

$$P(E \cap F) = P(E|F)\underbrace{P(F)}_{=0} = 0$$

Teorema (1.1) (Formula delle probabilità totali): Sia $\{E_n\}_n$ una famiglia finita o una successione di eventi disgiunti a due a due, e tali che $\cup_n E_n = \Omega$ (ossia $\{E_n\}_n$ è una partizione finita o numerabile). Allora, per ogni $B \in \mathcal{F}$, si ha che:

$$P(B) = \sum_n P(B|E_j)P(E_j)$$

Teorema (1.1): Osservando che gli eventi $\{B \cap E_n\}_n$ sono disgiunti a due a due, si ha:

$$\begin{aligned} P(B) = P(B \cap \Omega) &= P(B \cap (\cup_n E_n)) = \\ &= P(\cup_n (B \cap E_n)) = \\ &= \sum_n P(B \cap E_n) = \\ &= \sum_n P(B|E_n)P(E_n) \end{aligned}$$

Teorema (1.2) (Formula di Bayes): Nelle ipotesi del Teorema (1.1), e supponendo che $P(B) \neq 0$, si ha che:

$$P(E_j|B) = \frac{P(B|E_j)P(E_j)}{\sum_n P(B|E_n)P(E_n)} \text{ per ogni indice } j$$

Teorema (1.2): Si tratta di una verifica diretta, usando la definizione della probabilità condizionata e la formula delle probabilità totali:

$$P(E_j|B) = \frac{P(B \cap E_j)}{P(B)} = \frac{P(B|E_j)P(E_j)}{\sum_n P(B|E_n)P(E_n)} \text{ per ogni indice } J$$

Capitolo 2

Variabili aleatorie discrete e media

2.1 Le variabili aleatorie discrete

Definizione (2.1): Una **variabile aleatoria** è un'applicazione $X : \Omega \to \mathbb{R}$ con opportune proprietà (le controimmagini di intervalli finiti appartengono alla σ-algebra). UNa varaibile aleatoria è **discreta** se l'insieme dei valori assunti dalla variabile aleatoria è finito o numerabile.

Per come è definita una variabile aleatoria discreta, per ogni $a \in \mathbb{R}$ l'evento:

$$\{X = a\} := \{\omega \in \Omega : X(\omega) = a\}$$

È costituito da un insieme finito o numerabile di punti, e quindi si ha:

$$P(X = a) = \sum_{\omega \in \Omega : X(\omega) = a} P(\{\omega\})$$

Definizione (2.2): Due variabili aleatorie discrete X, Y (definite su uno stesso spazio di probabilità) sono **indipendenti** se:

$$P(\{X = x\} \cap \{Y = y\}) = P(X = x)P(Y = y) \text{ per ogni } x, y \in \mathbb{R}$$

La definizione si estende al caso più di due variabili aleatorie: precisamente, date $\{X_i\}_{i \in I}$, queste sono indipendenti se, per ogni sottoinsieme finito $\{i_1, i_2, ..., i_k\} \subset I$ con $k \geq 2$, si ha che:

$$\begin{aligned} P(\{X_{i_1} = x_{i_1}\} \cap \{X_{i_2} = x_{i_2}\} \cap ... \cap \{X_{i_n} = x_{i_n}\}) = \\ = P(X_{i_1} = x_{i_1})P(X_{i_2} = x_{i_2})...P(X_{i_n} = x_{i_n}) \end{aligned}$$

Nel caso in cui si ha $I = \{1, 2, ..., n$, basta chiedere:

$$P(\{X_1 = x_1\} \cap \{X_2 = x_2\} \cap ... \cap \{X_n = x_n\}) = P(X_1 = x_1)P(X_2 = x_2)...P(X_n = x_n)$$

2.2 La media

Una variabile aleatoria discreta X ha **valore atteso** o **media** finita se:

$$\sum_{x \in X} |x| P(X = x) < \infty$$

In tal caso il valore atteso è uguale a:

$$\mathbb{E}[X] = \sum_{x \in X} x P(X = x)$$

Possiamo notare che se X assume un numero finito di valori, la condizione per l'esistenza del valore atteso è sicuramente verificata.

Teorema (2.1) (linearità): Siano $X_1, X_2, ..., X_n$ variabili aleatorie discrete definite su uno stesso spazio di probabilità, e con valore atteso finito. Allore, per ogni $c_1, c_2, ..., c_n \in \mathbb{R}$, la variabile aleatoria $\sum_{i=1}^{n} c_i X_i$ ha speranza matematica finita e si ha:

$$\mathbb{E}[\sum_{i=1}^{n} c_i X_i] = sum_{i=1}^{n} c_i \mathbb{E}[X_i]$$

Teorema (2.1): La condizione di valore atteso finito è soddisfatta poiché

$$\sum_{x_1, x_2, ..., x_n} |c_1 x_1 + c_2 x_2 + ... + c_n x_n| P(X_1 = x_1, X_2 = x_2, ..., X_n = x_n) \leq$$

$$\leq \sum_{i=1}^{n} |c_i| \underbrace{\sum_{x_i} |x_i| P(X_i = x_i)}_{<\infty} < \infty$$

Inoltre si ha:

$$\begin{aligned}
\mathbb{E}[\sum_{i=1}^{n} c_i X_i] &= \sum_{x_1, x_2, ..., x_n} (c_1 x_1 + c_2 x_2 + ... + c_n x_n) P(X_1 = x_1, X_2 = x_2, ..., X_n = x_n) = \\
&= c_1 \sum_{x_1, x_2, ..., x_n} x_1 P(X_1 = x_1, X_2 = x_2, ..., X_n = x_n) + \\
&\quad + c_2 \sum_{x_1, x_2, ..., x_n} x_2 P(X_1 = x_1, X_2 = x_2, ..., X_n = x_n) + ... + \\
&\quad + c_n \sum_{x_1, x_2, ..., x_n} x_n P(X_1 = x_1, X_2 = x_2, ..., X_n = x_n)
\end{aligned}$$

In generale, per ogni $i \in \{1, 2, ..., n\}$, nella i-esima sommatoria si ha:

$$\sum_{x_1, x_2, ..., x_{i-1}, x_{i+1}, ..., x_n} P(X_1 = x_1, X_2 = x_2, ..., X_n = x_n) = P(X_i = x_i)$$

Quindi, ritornando alla quantità da calcolare, si ha:

$$\mathbb{E}[\sum_{i=1}^{n} c_i X_i] = c_1 \sum_{x_1} x_1 P(X_1 = x_1) + c_2 \sum_{x_2} x_2 P(X_2 = x_2) + ...+$$
$$+ c_n \sum_{x_n} x_n P(X_n = x_n) = \sum_{i=1}^{n} c_i \mathbb{E}[X_i]$$

La formula appena dimostrata non solo è valida anche nel caso di variabili aleatorie non indipendenti, ma è estendibile anche ad una famiglia di variabili aleatorie numeriche $\{X_i\}_{i\geq 1}$, supponendo che le costanti $\{c_i\}_{i\geq 1}$ scelte arbitrariamente siano tali che $\sum_{i\geq 1} |c_i| \mathbb{E}[X_i] < \infty$.

Teorema (2.2) (Disuguaglianza di Jensen): Sia X una variabile aleatoria (non necessariamente discreta). Sia $f : \mathbb{R} \to \mathbb{R}$ una funzione convessa. Allora $\mathbb{E}[\mho(\mathbb{X})] \geq f(\mathbb{E}[X])$.

Teorema (2.2): Vogliamo considerare il caso restrittivo in cui f ammetta la possibilità di avere una formula di Taylor del secondo ordine, cioè:

$$f(x) = f(y) + f'(y)(x-y) + \frac{f''(c)}{2}(x-y)^2$$

Con c punto medio opportune che appartiene all'intervallo $[x, y]$. Si vuole far riferimento a tale formula per $y = \mathbb{E}[X]$. Inoltre si sfrutta anche il fatto che, per convessità di f, si ha $f''(c) \geq 0$. Quindi:

$$f(x) = f(\mathbb{E}[X]) + f'(\mathbb{E}[X])(x - \mathbb{E}[X]) + \frac{f''(c)}{2}(x - \mathbb{E}[X])^2$$

Da cui, essendo $\frac{f''(c)}{2}(x - \mathbb{E}[X])^2 \geq 0$, si ha:

$$f(x) \geq f(\mathbb{E}[X]) + f'(\mathbb{E}[X])(x - \mathbb{E}[X])$$

Notare che nel teorema sopra si fa riferimento ad una funzione convessa. Si dice che una funzione è *convessa* se per ogni $x_1, x_2 \in \mathbb{R}$ e per ogni $\lambda \in [0, 1]$ si ha che $f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2)$, ossia che la corda che congiunge i due punti distinti $(x_1, f(x_1)), (x_2, f(x_2))$ è sopra al grafico della funzione f tra x_1 e x_2. Un esempio di possibile funzione f convessa è $f(x) = x^k$ con k intero positivo pari (per ogni $x \in \mathbb{R}$).

Parte II

Distribuzioni notevoli

Capitolo 3

Distribuzione di Bernouli

La **Distribuzione di Bernouli** è una distribuzione di probabilità su due soli valori: 0 ed 1, detti anche **fallimento** e **successo**. Più informalmente, può essere pensta come il modello per il set dei possibili risultati di un singolo esperimento che cerca di rispondere ad una domanda generale (ossia ad una domanda che che accetta solo due risposte: una affermativa e una contraria).

3.1 Definizione

Una variabile aleatoria X ha distribuzione di Bernouli se, preso $p \in [0, 1]$, si ha che:

$$P(X = k) = \begin{cases} p & \text{se } k = 1 \\ q = 1 - p & \text{se } k = 0 \\ 0 & \text{altrimenti} \end{cases}$$

In questo casi si scriverà che $X \sim 1_X(p)$.

3.2 Media

La media di una variabile aleatoria X con distribuzione di Bernouli è pari a:

$$\mathbb{E}[X] = 1 \cdot P(X = 1) + 0 \cdot P(X = 0) = 1 \cdot p = p$$

3.3 Varianza

La varianza di una variabile aleatoria X con distribuzione di Bernouli è pari a:

$$\begin{aligned} Var[X] = \mathbb{E}[X^2] - \mathbb{E}^2[X] = \\ &= 1^2 \cdot P(X = 1) + 0^2 \cdot P(X = 0) - p^2 = \\ &= 1 \cdot p - p^2 = p - p^2 = \\ &= p(1 - p) = pq \end{aligned}$$

Capitolo 4

Distribuzione binomiale