

## Contents

<b>1</b>	<b>Introduzione alla probabilità</b>	<b>3</b>
1.1	Glossario . . . . .	3
1.2	Moda e Mediana . . . . .	4
1.2.1	Moda . . . . .	4
1.2.2	Mediana . . . . .	4
1.3	Media e Varianza Campionaria . . . . .	4
1.3.1	Media Campionaria . . . . .	4
1.3.2	Varianza Campionaria . . . . .	5
1.4	Disugaglianza di Chebyshev . . . . .	5
1.5	Percentile . . . . .	5
1.6	Insieme di dati Bivariati . . . . .	5
1.6.1	Coefficiente di correlazione campionario . . . . .	6
1.7	Permutazioni, Combinazioni e Disposizioni . . . . .	6
1.7.1	Permutazioni . . . . .	6
1.7.2	Combinazioni . . . . .	6
1.7.3	Disposizioni . . . . .	6
1.8	Probabilità condizionata . . . . .	7
1.8.1	Teorema di Bayes . . . . .	7
1.9	Operazioni e proprietà tra eventi . . . . .	7
<b>2</b>	<b>Variabile aleatorie</b>	<b>8</b>
2.1	Funzione di ripartizione (Tutte le variabili) . . . . .	8
2.2	Funzione di massa (Variabili discrete) . . . . .	8
2.3	Funzione della densità di probabilità (Variabili continue) . . . . .	9
<b>3</b>	<b>Funzioni a due variabili</b>	<b>11</b>
3.1	Funzione di ripartizione congiunta . . . . .	11
3.2	Funzione di massa congiunta . . . . .	11
3.3	Funzione densità congiunta . . . . .	12
3.4	Variabili aleatorie indipendenti . . . . .	13
3.4.1	X,Y indipendenti . . . . .	13
3.5	Distribuzioni condizionate . . . . .	15
3.6	funzione di massa condizionata (Discrete) . . . . .	15
3.7	funzione di densità condizionata (Continue) . . . . .	16
<b>4</b>	<b>Valore atteso</b>	<b>17</b>
4.1	Funzione di massa (Discrete) . . . . .	17
4.2	Funzione di densità (Continue) . . . . .	17
4.3	Valore atteso di una funzione . . . . .	18
4.4	Dimostrazioni . . . . .	19
4.5	Momenti N-esimi nel valore atteso . . . . .	20
4.6	Valore atteso della somma di due variabili . . . . .	20

<b>5</b>	<b>Varianza</b>	<b>21</b>
5.1	Costanti reali nella varianza . . . . .	22
<b>6</b>	<b>Deviazione Standard</b>	<b>23</b>
<b>7</b>	<b>Covarianza</b>	<b>23</b>
7.1	Proprietà della covarianza . . . . .	23
7.2	Coefficiente di correlazione lineare . . . . .	24
<b>8</b>	<b>Funzione generatrice dei momenti</b>	<b>25</b>
8.1	Disugaglianza di Markov . . . . .	25
8.2	Disugaglianza di Chebyshev . . . . .	26
<b>9</b>	<b>Legge debole dei grandi numeri</b>	<b>26</b>
<b>10</b>	<b>Modelli di variabili aleatorie</b>	<b>27</b>
10.1	Bernoulli . . . . .	27
10.2	Binomiali . . . . .	27
10.2.1	Valore atteso e varianza di Binomiali . . . . .	27
10.2.2	Funzione di massa e di ripartizione di Binomiali . . . . .	27
10.3	Poisson . . . . .	27
10.4	Ipergeometriche . . . . .	27
10.5	Uniformi . . . . .	27
10.6	Normali o Gaussiane . . . . .	27
10.7	Esponenziali . . . . .	27
10.8	Processi stocastici (Poisson) . . . . .	27
10.9	Gamma . . . . .	27
10.10	Chi-quadro . . . . .	27
10.11	Distribuzione T . . . . .	27
10.12	Distribuzione F . . . . .	27
10.13	Distribuzione logistica . . . . .	27

# 1 Introduzione alla probabilità

## 1.1 Glossario

- Sistemi non deterministici → *conoscendo i dati iniziali non possiamo determinare i dati finali*
- Incertezza degli eventi → *la varianza degli eventi che possono succedere*
- Rumore → *possiamo misurare un evento solo approssimativamente*
- Probabilità → *la materia che studia i sistemi non deterministici*
  - Frequentista → *probabilità assegnata sulla base di più esperimenti ripetuti nella stessa condizione*
  - Soggettivista → *non esiste un valore oggettivo ma ci si basa sulla fiducia e sull'incertezza che l'individuo ha riguardo l'occorrenza di un certo evento*
- Varianza → *dispersione dei dati attorno al valore centrale / media / valore*
- Confidenza → *intervallo che rappresenta una stima dei valori medi*
- Frequenza
  - Frequenza assoluta → Numero di volte che si verifica un evento
  - Frequenza relativa → Rapporto tra frequenza assoluta e il numero di prove/dati
- Dataset → numero di dati a disposizione  $D_n = \{x_1 \cdot \cdot \cdot x_n\}$
- Principio di enumerazione → Passare solo una volta da ogni elemento della raccolta
- Spazio esiti ( $s$  o  $\Omega$ ) → Tutti i possibili esiti di un evento →  $Dado = \{1 \cdot \cdot \cdot 6\}$
- Spazio eventi ( $e$ ) → Tutti i possibili risultati di un esperimento →  $Dado = \{1|2\} \leftarrow$  che esca **1** oppure **2**
- Assioma → Tre assiomi fondamentali su cui si poggia la teoria del calcolo delle probabilità
  - 1' Assioma → La probabilità di  $E$  è un numero reale **non negativo**  
 $\mathbb{P}(E) \in \mathbb{R}, \mathbb{P}(E) \geq 0 \ \forall E \subseteq \Omega \mid 0 \leq P(E) \leq 1$
  - 2' Assioma → Allo spazio degli esiti è sempre associato ad **1**  
 $\mathbb{P}(s) = 1$
  - 3' Assioma → Per ogni coppia di eventi incompatibili  $E_1, E_2 \subseteq \Omega$  la probabilità di  $E_1 \cup E_2$  è uguale alla **somma della loro probabilità**  
 $\mathbb{P}(E_1 \cup E_2) = \mathbb{P}(E_1) + \mathbb{P}(E_2)$

## 1.2 Moda e Mediana

### 1.2.1 Moda

**Definizione:** La moda è il valore che presenta la **massima frequenza** all'interno del dataset

**Formula generica:**

$$Moda \rightarrow v_i : f_i = \max f_i \begin{cases} \text{un solo valore} & \text{Moda} \\ \text{più di un valore} & \text{Valori modali} \end{cases}$$

### 1.2.2 Mediana

**Definizione:** La mediana è il **valore centrale** all'interno del dataset (dati ordinati in ordine crescente/decesente)

**Formula generica:**

$$Mediana = \begin{cases} n \text{ pari} & \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2} \\ n \text{ dispari} & x_{[\frac{n+1}{2}]} \leftarrow \text{Intero superiore (Ceil)} \end{cases}$$

**Esempio:**

$$D_n = \{ 28, 34, 51, 19, 62, 43, 29, 38, 45, 26, 49, 33 \}$$

Per la mediana è necessario ordinare i dati in ordine crescente:

$$D_n = \{ 19, 26, 28, 29, 33, 34, 38, 43, 45, 49, 51, 62 \}$$

$$\frac{x_{\frac{12}{2}} + x_{\frac{12}{2}+1}}{2} = \frac{x_6 + x_7}{2} = \frac{34 + 38}{2} = \frac{72}{2} = 36$$

**Nota:** quando si trova ad esempio  $x_6$  bisogna andare a sostituire il valore con la posizione di x

## 1.3 Media e Varianza Campionaria

### 1.3.1 Media Campionaria

**Definizione:** La media campionaria è la **media** degli elementi di un campione.

**Formula generica:**

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$$

### 1.3.2 Varianza Campionaria

**Definizione:** La varianza campionaria è la **dispersione** degli elementi di un campione.

**Formula generica:**

$$S^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}$$

**Esempio: (Varianza e Media)**  $D_n = \{ 3, 4, 6, 7, 10 \}$

$$\text{Media del campione: } \bar{X} = \frac{(3+4+6+7+10)}{5} = \frac{30}{5} = 6$$

$$\text{Varianza campionaria: } s^2 = \frac{[(-3)^2 + (-2)^2 + 0^2 + 1^2 + 4^2]}{4} = 7.5$$

### 1.4 Disuguaglianza di Chebyshev

**Definizione:** Dice quanti dati di un campione cadono all'interno di un intervallo con centro la **media**

$$\forall k \geq 1 : k \in \mathbb{R}$$

$$(\bar{x} - k_s, \bar{x} + k_s) \longrightarrow S_k : [i : 1 \leq i \leq n, |x_i - \bar{x}| < k_s]$$

**Generalizzando:**

$$|x - \bar{x}| < 5 \longrightarrow 68\%$$

$$|x - \bar{x}| < 25 \longrightarrow 95\%$$

$$|x - \bar{x}| < 35 \longrightarrow 99.7\%$$

### 1.5 Percentile

**Definizione:** Il percentile è un indicatore che serve ad **indicare il valore minimo** sotto al quale ricade una **determinata percentuale** degli altri elementi sotto osservazione.

$$Valore \begin{cases} \geq & k \% \text{ dati} \\ \leq & 100 - k \% \text{ dati} \end{cases}$$

Prima cosa da fare è ordinare i valori in ordine crescente  
Dove il secondo quartile è sempre uguale alla **mediana**

**Esempio:**  $D_n = \{ 0, 1, 2, 3, 4, 5, 6, 7, 8 \}$

### 1.6 Insieme di dati Bivariati

**Definizione:** è lo studio della relazione di due variabili.

**Formula generica:**

$$D_n : \{(X_1, Y_1)(X_2, Y_2) \cdots (X_n, Y_n)\}$$

### 1.6.1 Coefficiente di correlazione campionario

**Definizione:** utilizzato per capire se esiste un legame **lineare** tra due serie di dati.

**Formula generica:**

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{(n-1)s_x s_y} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}}$$

## 1.7 Permutazioni, Combinazioni e Disposizioni

### 1.7.1 Permutazioni

**Definizione:** Modi possibili per sistemare **n** oggetti (**0!** = 1)

$$n! = n \cdot (n-1) \cdots (n \cdot (n-1))$$

**Esempio:** Fattoriale di 6

$$6! = 6 \cdot (6-1) \cdot (6-2) \cdots (6-5) = 720$$

### 1.7.2 Combinazioni

**Definizione:** Modi di disporre **k** elementi scelti da **n** elementi (l'ordine **non conta**)

$$\frac{n!}{k! \cdot (n-k)!} = \binom{n}{k}$$

**Esempio:** in una classe di **26** alunni si devono eleggere **2** rappresentanti

$$C_{n,k} = \frac{n!}{k! \cdot (n-k)!}$$

Sostituiamo **n** con 26 (numero di alunni) e **k** con 2 (numeri di rappresentanti)

$$C_{26,2} = \frac{26!}{2! \cdot (26-2)!} = \frac{26!}{2! \cdot 24!} = \frac{25 \cdot 26}{2} = \mathbf{325}$$

è possibile anche semplificare i fattoriali come in questo caso

### 1.7.3 Disposizioni

**Definizione:** Modi di disporre **k** elementi scelti da **n** elementi (l'ordine **conta**)

$$\frac{n!}{(n-k)!}$$

**Esempio:** Quante parole si possono ottenere usando 4 **diverse** lettere da *youmath*  
 In questo caso dobbiamo contare le **disposizioni** senza ripetizione di **classe 4 di 7**

$$D_{7,4} = \frac{7!}{(7-4)!} = \frac{7!}{3!} = \frac{5040}{6} = 840$$

## 1.8 Probabilità condizionata

**Definizione:** è la probabilità che succeda un evento **E** dato un evento **F**

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

**Esempio:** 3 scatole con contenuto nascosto dove in una è presente il premio

$$P(Vincita) = \frac{1}{3}$$

$$P(Vincita|1' \text{ pacco contiene un gatto}) = \frac{1}{2}$$

$$P(Vincita|1' \text{ pacco NON contiene un gatto}) = 0$$

### 1.8.1 Teorema di Bayes

**Formula generica:**

$$P(F_j|E) = \frac{P(F_j \cap E)}{P(E)} = \frac{P(E|F_j) \cdot P(F_j)}{\sum_{i=1}^p P(E|F_i) \cdot P(F_i)}$$

Probabilità di  $F_j$  sapendo che si sia verificato l'evento **E**

## 1.9 Operazioni e proprietà tra eventi

**Definizione:** Prendiamo come esempio **E** ed **F** come eventi

- $E \cup F \leftarrow$  Unione
- $E \cap F \leftarrow$  Intersezione
- $E \subset F \mid E \subseteq F \leftarrow$  Contenuto
- $E \supset F \mid E \supseteq F \leftarrow$  Contiene
- $E^c \leftarrow$  Complemento

**Le seguenti operazioni possono essere combinate tra di loro:** formando così le proprietà che seguono:

- $E \cup (F \cap G) = (E \cup F) \cap G \rightarrow$  Associativa unione
- $(E \cup F) \cap G = (E \cap G) \cup (F \cap G) \rightarrow$  Distributiva intersezione
- $E \cap (F \cup G) = (E \cap F) \cup (E \cap G) \rightarrow$  Associativa intersezione
- $(E \cap F) \cup G = (E \cup G) \cap (F \cup G) \rightarrow$  Distributiva unione
- $(E \cup F)^c = \frac{E^c \cap F^c}{(E \cap F)^c} = E^c \cap F^c$

## 2 Variabile aleatorie

**Definizione:** La variabile aleatoria è una variabile che può assumere **valori diversi** in dipendenza da *qualche esperimento casuale*.

$$X \begin{cases} \text{Discrete} & \text{Solo valori finiti} \\ \text{Continue} & \text{Possono assumere range illimitati} \end{cases}$$

### 2.1 Funzione di ripartizione (Tutte le variabili)

**Definizione:** La Probabilità che la variabile aleatoria **X** assuma un valore minore o uguale ad x

**Formula generica:**  $F(x) = P(X \leq x)$

- F = funzione di ripartizione
- X = variabile aleatoria
- x = variabile normale

**Esempio :**

$$P(a < X \leq b)$$

$$P(X \leq b) = P(X \leq a) + P(a < X \leq b)$$

$$P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a)$$

### 2.2 Funzione di massa (Variabili discrete)

**Formula generica:**  $p(a) = P(X = a)$

se si ha la funzione di ripartizione è possibile ottenere la funzione di massa perche:

$$X \leq a = \cup X_i$$



**Formula generica:**

$$F(x) = P(X \leq a) = \sum_{x \leq a} p(x_i)$$

TODO- GRAFICO

**Esempio:** variabile aleatoria X che può assumere valori **1, 2 o 3**

Dato che  $p(1) + p(2) + p(3) = 1$

**Se:**

$$p(1) = \frac{1}{2}$$

$$p(2) = \frac{1}{3}$$

**Allora:**

$$p(3) = \frac{1}{6}$$

La funzione di ripartizione F di X è data da:

$$F(a) = \begin{cases} 0 & a < 1 \\ \frac{1}{2} & 1 \leq a < 2 \\ \frac{5}{6} & 2 \leq a < 3 \\ 1 & 3 \leq a \end{cases}$$

## 2.3 Funzione della densità di probabilità (Variabili continue)

**Formula generica:**

$$P(X \in B) = \int_B f(x) dx$$

$$P(X \in (-\infty, +\infty)) = \int_{-\infty}^{+\infty} f(x) dx = 1$$

integrando  $-\infty$  a  $+\infty$  la probabilità che avvenga x è per forza 1 perché andiamo ad includere tutti i valori di  $\mathbb{R}$

Se abbiamo che  $\mathbf{B} = [\mathbf{a}, \mathbf{b}] \rightarrow P(a \leq X \leq b) = \int_a^b f(x) dx$

Se abbiamo che  $\mathbf{B} = [\mathbf{a}] \rightarrow P(X = a) = \int_a^a f(x) dx = 0$

Relazione che lega la funzione di ripartizione **F** alla densità **f**:

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(x) dx$$

Derivando entrambi i membri otteniamo che:

$$\frac{d}{da} F(a) = f(a)$$

**Esempio:** Sia assegnata una variabile aleatoria  $X$  con densità data da:

$$f(x) = \begin{cases} C(4x - 2x^2) & 0 < x < 2 \\ 0 & \text{altrimenti} \end{cases}$$

**(a)** quanto vale  $C$ ? **(b)** quanto vale  $P(X > 1)$ ?

**(a)** siccome  $f$  è una densità allora:

$$\begin{aligned} 1 &= C \int_0^2 (4x - 2x^2) dx \\ &= C \left[ 2x^2 - \frac{2x^3}{3} \right] \Big|_{x=0}^{x=2} = C \cdot \frac{8}{3} \\ &= C = \frac{3}{8} \end{aligned}$$

**(b)** conoscendo ora la densità  $f$  possiamo trovare la  $P(X > 1)$ :

$$P(X > 1) = \int_1^\infty f(x) dx = \frac{3}{8} \int_1^2 (4x - 2x^2) dx = \frac{1}{2}$$

### 3 Funzioni a due variabili

Questo tipo di funzioni ci sono utili quando l'utilizzo di una sola variabile è impossibile poichè *l'oggetto in questione è basato sulla relazione di due variabili aleatorie*

#### 3.1 Funzione di ripartizione congiunta

**Definizione:** Funzione di ripartizione a due variabili aleatorie  $X$  e  $Y$

**Formula generica:**

$$F(x, y) = P(X \leq x, Y \leq y)$$

Se vogliamo trovare solamente la funzione di ripartizione di una singola variabile aleatoria:

$$\begin{aligned} F_X(x) &= P(X \leq x) \\ &= P(X \leq x, Y \leq \infty) \\ &= F(x, \infty) \end{aligned}$$

Applicabile anche alla  $F_Y(y)$

$$F_Y(y) = F(\infty, y)$$

#### 3.2 Funzione di massa congiunta

**Definizione:** Probabilità che accadano due eventi ( $\mathbf{X}$  e  $\mathbf{Y}$ ) nello stesso istante.

**Formula generica:**  $p(x_i, y_j) = P(X = x_i, Y = y_j)$

Se vogliamo trovare solamente la funzione di massa di una singola variabile aleatoria:

$$\begin{aligned} p_X(x_i) &:= P(X = x_i) \\ &= P(\cup_j \{X = x_i, Y = y_j\}) \\ &= \sum_j P(X = x_i, Y = y_j) \\ &= \sum_j p(x_i, y_j) \end{aligned}$$

Applicabile anche alla  $p_Y$

$$p_Y(y_j) = \sum_i p(x_i, y_j)$$

$$\sum_x \sum_y p(x, y) = 1$$

### 3.3 Funzione densità congiunta

**Definizione:** Due variabili aleatorie  $X$  e  $Y$  sono *congiuntamente continue* se esiste una funzione non negativa  $f(x,y)$  definita per tutti gli  $x$  e gli  $y$

**Formula generica:**

$$P((X,Y) \in C) = \int \int_{(x,y) \in C} f(x,y) dx dy$$

se  $A$  e  $B$  sono sottoinsiemi qualsiasi di  $\mathbb{R}$  e  $C := A \times B$

$$C := (x,y) \in \mathbb{R}^2 : x \in A, y \in B$$

Possiamo riscrivere la funzione di ripartizione congiunta di  $X$  e  $Y$  come segue:

$$\begin{aligned} F(a,b) &= P(X \leq a, Y \leq b) \\ &= P(X \in a, Y \in b) \\ &= \int_B \int_A f(x,y) dx dy \\ &= \int_{-\infty}^a \int_{-\infty}^b f(x,y) dx dy \end{aligned}$$

**Esempio:** Siano  $X$  e  $Y$  due variabili aleatorie congiuntamente continue con densità di probabilità data da:

$$f(x,y) = \begin{cases} 2e^{-x}e^{-2y} & x > 0, y > 0 \\ 0 & \text{altrimenti} \end{cases}$$

Si calcolino **(a)**  $P(X > 1, Y < 1)$

$$\begin{aligned} P(X > 1, Y < 1) &= \int_0^1 \int_1^\infty 2e^{-x}e^{-2y} dx dy \\ &= \int_0^1 2e^{-2y} \left( \int_1^\infty e^{-x} dx \right) dy \\ &= \int_0^1 2e^{-2y} \{-e^{-x}\}|_{x=1}^\infty dy \\ &= e^{-1} \int_0^1 2e^{-2y} dy \\ &= e^{-1}(1 - e^{-2}) \end{aligned}$$

In questo caso si è integrato prima in una variabile e poi nell'altra

## 3.4 Variabili aleatorie indipendenti

### 3.4.1 X, Y indipendenti

**Definizione:** Un evento su una variabile non influenza l'altra.

**Formula generica:** Se soddisfano questa richiesta le variabili si dicono *indipendenti*

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$$

Usando gli assiomi della probabilità è possibile dimostrare che la definizione di sopra è equivalente a:

$$P(X \leq a, Y \leq b) = P(X \leq a)P(Y \leq b)$$

$$\forall a, b \in \mathbb{R}$$

Ovvero che la funzione di ripartizione congiunta sia il prodotto delle marginali:

$$F(a, b) = F_X(a)F_Y(b)$$

**Funzione di massa:**

$$p(a, b) = p_X(a)p_Y(b)$$

$$P(X = a, Y = b) = P(X = a)P(Y = b)$$

**Dimostrazione:**

$$\begin{aligned} P(X \in A, Y \in B) &= \sum_{x \in A} \sum_{y \in B} p(x, y) \\ &= \sum_{x \in A} \sum_{y \in B} p_X(x)p_Y(y) \\ &= \sum_{x \in A} p_X(x) \sum_{y \in B} p_Y(y) \\ &= P(X \in A)P(Y \in B) \end{aligned}$$

**Funzione di densità:**

$$f(x, y) = f_X(x)f_Y(y)$$

$$\forall x, y \in \mathbb{R}$$

**Esempio con variabili indipendenti continue e con stessa funzione di densità:**

$$f_X(t) = f_Y(t) = \begin{cases} e^{-t} & t > 0 \\ 0 & \text{altrimenti} \end{cases}$$

Quale è la densità di probabilità della variabile aleatoria data dal rapporto  $X/Y$

$$\begin{aligned} F_{X|Y}(a) &= P(X|Y \leq a) \\ &= \int_{(x,y)} \int_{x \leq ay} f(x,y) dx dy \\ &= \int_{(x,y)} \int_{x \leq ay} f(x)f(y) dx dy \\ &= \int_0^\infty \int_0^{ay} e^{-x} f(x)f(y) dx dy \\ &= \int_0^\infty e^{-y} \left( \int_0^{ay} e^{-x} dy \right) dy \\ &= \int_0^\infty e^{-y} (1 - e^{-ay}) dy \\ &= \left[ -e^{-y} + \frac{e^{-(a+1)y}}{(a+1)} \right] \Big|_0^\infty \\ &= 1 - \frac{1}{a+1} \end{aligned}$$

La funzione di densità si ricava infine **derivando** la funzione di ripartizione

$$f_{X|Y}(a) = \frac{d}{da} \left( 1 - \frac{1}{a+1} \right) = \frac{1}{(a+1)^2} a > 0$$

### 3.5 Distribuzioni condizionate

**Definizione:** La distribuzione condizionata di Y dato X è la probabilità di X quando è conosciuto il valore assunto da X.

A ogni distribuzione condizionata è associato un valore atteso condizionato e una varianza condizionata

**Formula generica:**  $P(E|F) = \frac{P(E \cap F)}{P(F)}$

### 3.6 funzione di massa condizionata (Discrete)

**Formula generica:**

$$\begin{aligned} p_{X|Y}(X|Y) &= P(X = x, Y = y) \\ &= \frac{P(X = x, Y = y)}{P(Y = y)} \\ &= \frac{p(X, Y)}{p_Y(x, y) > 0} \end{aligned}$$

$$\forall x, \forall y \text{ con } p_Y(y) > 0$$

Se y non è un valore possibile di Y, ovvero se  $P(Y = y) = 0$ , la quantità  $p_{X|Y}(x|y)$  non è definita

**Esempio:** Siano X e Y due variabili aleatorie discrete con funzione di massa congiunta p dato che:

$$p(0, 0) = 0.4 \quad p(0, 1) = 0.2 \quad p(1, 0) = 0.1 \quad p(1, 1) = 0.3$$

Calcolare la massa di X condizionata da  $Y = 1$

$$P(Y = 1) = \sum_x p(x, 1) = p(0, 1) + p(1, 1) = 0.5$$

**Quindi:**

$$P(X = 0|Y = 1) = \frac{p(0, 1)}{P(Y = 1)} = \frac{2}{5}$$

$$P(X = 1|Y = 1) = \frac{p(1, 1)}{P(Y = 1)} = \frac{3}{5}$$

Se X e Y sono variabili congiuntamente continue, non è possibile utilizzare la definizione di distribuzione condizionata valida per quelle discrete, infatti sappiamo che  $P(Y = y) = 0$  per tutti i valori di y

### 3.7 funzione di densità condizionata (Continue)

**Formula generica:**

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)}$$

Se  $X$  e  $Y$  sono congiuntamente continue e  $A$  è un sottoinsieme di numeri reali per ogni  $y$  si può definire:

$$P(X \in A|Y = y) := \int_A f_{X|Y}(x|y) dx$$

Notiamo che  $X$  e  $Y$  sono indipendenti allora:

$$f_{X|Y}(x, y) = f_X(x) \qquad P(X \in A|Y = y) = P(X \in A)$$

**Esempio:** è data la seguente densità congiunta di  $X$  e  $Y$

$$f(x, y) = \begin{cases} \frac{12}{5}x(2 - x - y) & 0 < x < 1, 0 < y < 1 \\ 0 & \text{altrimenti} \end{cases}$$

Si calcoli la densità condizionata di  $X$  rispetto a  $Y = y$  per  $0 < y < 1$ .  
Se questi due numeri sono compresi tra 0 e 1 abbiamo che:

$$\begin{aligned} f_{X|Y}(x, y) &:= \frac{f(x, y)}{f_Y(y)} \\ &= \frac{f(x, y)}{\int_{-\infty}^{\infty} f(x', y) dx'} \\ &= \frac{x(2 - x - y)}{\int_0^1 x'(2 - x' - y) dx'} \\ &= \frac{x(2 - x - y)}{\frac{2}{3} - \frac{y}{2}} \\ &= \frac{6x(2 - x - y)}{4 - 3y} \end{aligned}$$



## 4 Valore atteso

**Definizione:** Rappresenta la media pesata dei valori di una variabile aleatoria

### 4.1 Funzione di massa (Discrete)

$$\mathbb{E}[X] := \sum_i x_i P(X = x_i)$$

Si può dire quindi che il valore atteso è anche detto *media* di X oppure *aspettazione*

**Esempio semplice:** Se X è una variabile aleatoria con funzione di massa

$$p(0) = \frac{1}{2} = p(1)$$

Allora:

$$\mathbb{E}[X] = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{0+1}{2} = \frac{1}{2}$$

**Esempio dado fair 6 facce**  $P(x_i = i) = \frac{1}{6}$

$$\mathbb{E}[X] = \sum_{i=1}^6 i \frac{1}{6} = \frac{1}{6} \sum_{i=1}^6 i = \frac{21}{6} = \frac{7}{2} = 3.5$$

Oppure:

$$\mathbb{E}[X] := 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = \frac{7}{2} = 3.5$$

dove il risultato è la media dei valori che X può assumere

**Se N è molto grande** allora  $N_i \approx N_p(x_i)$

$$\sum_i^n x_i p(x_i) \approx \sum_i^n x_i \frac{n_i}{n}$$

### 4.2 Funzione di densità (Continue)

**Formula generica:**

$$\mathbb{E}[X] := \int_{-\infty}^{+\infty} x f(x) dx$$

**Esempio:** Siamo in attesa di una comunicazione che deve arrivare dopo le ore 17. a partire dalle 17 è una variabile aleatoria con funzione di densità data da:

$$f(x) = \begin{cases} \frac{1}{1.5} & \text{se } 0 < x < 1.5 \\ 0 & \text{altrimenti} \end{cases}$$

Il valore atteso del tempo che trascorre tra le 17 e il momento di arrivo della comunicazione è quindi:

$$\mathbb{E}[X] = \int_0^{1.5} \frac{x}{1.5} dx = 0.75$$

### 4.3 Valore atteso di una funzione

**Definizione:** è possibile calcolare il valore atteso di una funzione  $g(X)$  notando che essa stessa è una variabile aleatoria quindi si applicano le stesse proprietà, come segue:

Variabile discreta:

$$\mathbb{E}[g(X)] = \sum_i g(x_i)p(x_i)$$

Variabile Continua:

$$\mathbb{E}[g(x)] = \int_{-\infty}^{+\infty} g(x)f(x) dx$$

**Esempio (discrete):** quanto vale il valore atteso del quadrato di una variabile  $X$  con le seguenti funzioni di massa?

$$p(0) = 0.2$$

$$p(1) = 0.5$$

$$p(2) = 0.3$$

Se poniamo  $Y := X^2$  questa diventa una variabile che può assumere i valori  $0^2, 1^2, 2^2$

$$p_Y(0) := P(Y = 0^2) = 0.2$$

$$p_Y(1) := P(Y = 1^2) = 0.5$$

$$p_Y(4) := P(Y = 2^2) = 0.3$$

Quindi:

$$\mathbb{E}[X^2] = \mathbb{E}[Y] = 0 \cdot 0.2 + 1 \cdot 0.5 + 4 \cdot 0.3 = 1.7$$

Oppure (utilizzando la proposizione delle variabili discrete)

$$\mathbb{E}[X^2] = 0^2 \cdot 0.2 + 1^2 \cdot 0.5 + 2^2 \cdot 0.3 = 1.7$$

**Esempio (continue):** Il tempo – in ore – necessario per localizzare un guasto nell'impianto elettrico di una fabbrica è una variabile aleatoria  $X$  con funzione di densità

$$f(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{altrimenti} \end{cases}$$

Se il danno economico provocato da una interruzione di  $x$  ore è  $x^3$ , qual è il valore atteso di questo costo?

**Applicando la proposizione della variabile continua** possiamo ottenere quanto segue:

$$\mathbb{E}[X^3] = \int_0^1 x^3 dx = \frac{1}{4}$$

#### 4.4 Dimostrazioni

Sia per discreto che per continuo si applicano le seguenti proprietà:

$$\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$$

Se proviamo a ponere  $a = 0$  scopriamo che:

$$\mathbb{E}[b] = b$$

Se proviamo a ponere  $b = 0$  scopriamo che:

$$\mathbb{E}[aX] = a\mathbb{E}[X]$$

Ovvero, il valore atteso di un fattore costante moltiplicato per una variabile aleatoria, è pari alla costante per il valore atteso della variabile aleatoria.

**Per caso discreto:**

$$\begin{aligned} \mathbb{E}[aX + b] &= \sum_x (ax + b)p(x) \\ &= a \sum_x xp(x) + b \sum_x p(x) \\ &= a\mathbb{E}[X] + b \end{aligned}$$

**Per caso continuo:**

$$\begin{aligned} \mathbb{E}[aX + b] &= \int_{-\infty}^{+\infty} (ax + b)f(x) dx \\ &= a \int_{-\infty}^{+\infty} xf(x) dx + b \int_{-\infty}^{+\infty} f(x) dx \\ &= a\mathbb{E}[X] + b \end{aligned}$$

## 4.5 Momenti N-esimi nel valore atteso

**Definizione:** se  $n = 1, 2 \dots n$ , la quantità  $\mathbb{E}[X^n]$  se esiste viene detta *momento n-esimo* della variabile aleatoria  $X$ .

è possibile applicare le formule di prima, come segue:

$$\mathbb{E}[X^n] = \begin{cases} \sum_x x^n p(x) & \text{se } X \text{ è discreta} \\ \int_{-\infty}^{+\infty} x^n f(x) dx & \text{se } X \text{ è continua} \end{cases}$$

## 4.6 Valore atteso della somma di due variabili

**Definizione:** è possibile applicare le formule viste sopra anche quando abbiamo due variabili aleatorie

se in questo caso  $\mathbb{E}[g(X, Y)]$  esiste allora:

$$\mathbb{E}[g(X, Y)] = \begin{cases} \sum_x \sum_y g(x, y) p(x, y) & \text{Se discreto} \\ \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x, y) f(x, y) dx dy & \text{Se continuo} \end{cases}$$

se  $g(X, Y)$  come  $g = X + Y$  allora

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$

**Dimostrazione caso discreto:**

$$\begin{aligned} \mathbb{E}[X + Y] &= \sum_x \sum_y (x + y) p(x, y) \\ &= \sum_x x \cdot [\sum_j p(x_i, y_j)] + \sum_x y \cdot [\sum_i p(x_i, y_j)] \\ &= \sum_x x p_X(x) + \sum_y y p_Y(y) \\ &= \mathbb{E}[X] + \mathbb{E}[Y] \end{aligned}$$

**Dimostrazione caso continuo:**

$$\begin{aligned} \mathbb{E}[X + Y] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x + y) f(x, y) dx dy \\ &= \int_{-\infty}^{\infty} x [\int_{-\infty}^{\infty} f(x, y) dy] dx + \int_{-\infty}^{\infty} y [\int_{-\infty}^{\infty} f(x, y) dx] dy \\ &= \int_{-\infty}^{\infty} x f_X(x) dx + \int_{-\infty}^{\infty} y f_Y(y) dy \\ &= \mathbb{E}[X] + \mathbb{E}[Y] \end{aligned}$$

è possibile applicare la ricorsione per il numero di variabili aleatori

$$\begin{aligned}\mathbb{E}[X + Y + Z] &= \mathbb{E}[(X + Y) + Z] \\ &= \mathbb{E}[X + Y] + \mathbb{E}[Z] \\ &= \mathbb{E}[X] + \mathbb{E}[Y] + \mathbb{E}[Z]\end{aligned}$$

In generale per ogni  $n$

$$\mathbb{E}[X_1 + X_2 + \dots + X_n] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + \dots + \mathbb{E}[X_n]$$

**Esempio: 2 dadi a 6 facce**

$$\begin{aligned}\mathbb{E}[X + Y] &= \mathbb{E}[X] + \mathbb{E}[Y] \\ &= \sum_{i=1}^6 x_i p(x_i) + \sum_{i=1}^6 y_i p(y_i) \\ &= \sum_{i=1}^6 x_i \frac{1}{6} + \sum_{i=1}^6 y_i \frac{1}{6} \\ &= \frac{7}{2} + \frac{7}{2} = 7\end{aligned}$$

Dove **7** è il valore atteso della somma dei due dadi.

Se vogliamo predire il valore di  $X$  possiamo scegliere un numero che sarà uguale ad  $X$ . L'errore che commetteremo sarà di  $(X - c)^2$

Se  $c = \mathbb{E}[X]$  l'errore sarà minimizzato  $\mu := \mathbb{E}[X]$

$$\mathbb{E}[(X - c)^2] \geq \mathbb{E}[(X - \mu)^2]$$

## 5 Varianza

**Definizione:** Indica di quanto i dati si discostano dalla media al quadrato

$$\mu = \mathbb{E}[X] \leftarrow \text{Primo momento} \qquad \mathbb{E}[X^2] \leftarrow \text{Momento secondo}$$

**Formula generica:**

$$\text{Var}(X) := \mathbb{E}[(X - \mu)^2]$$

**Generalizzazione:**

$$\begin{aligned}\text{Var}(X) &= \mathbb{E}[(X - \mu)^2] \\ &= \mathbb{E}[X^2 - 2\mu \cdot X + \mu^2] \\ &= \mathbb{E}[X^2] - 2\mu \cdot \mathbb{E}[X] + \mu^2 \\ &= \mathbb{E}[X^2] - \mu^2 \longrightarrow \mathbb{E}[X^2] - \mathbb{E}[X]^2\end{aligned}$$

### Esempio: Varianza di un dado

$$\begin{aligned}\mathbb{E}[X^2] &= \sum_1^6 i^2 P(X=i) \\ &= 1^2 \cdot \frac{1}{6} + 2^2 \cdot \frac{1}{6} + 3^2 \cdot \frac{1}{6} + 4^2 \cdot \frac{1}{6} + 5^2 \cdot \frac{1}{6} + 6^2 \cdot \frac{1}{6} \\ &= \frac{91}{6}\end{aligned}$$

Sapendo che  $\mathbb{E}[X] = \frac{7}{2}$

$$Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}$$

## 5.1 Costanti reali nella varianza

Una utile identità che riguarda la varianza è la seguente (per ogni coppia di costanti reali  $a$  e  $b$ )

$$Var(aX + b) = a^2 \cdot Var(X)$$

Per dimostrare ciò ricordiamoci sempre di  $\mu := \mathbb{E}[X]$

### Dimostrazione:

$$\begin{aligned}Var(aX + b) &:= \mathbb{E}[(aX + b - \mathbb{E}[aX + b])^2] \\ &= \mathbb{E}[(aX + b - a\mu - b)^2] \\ &= \mathbb{E}[a^2(X - \mu)^2] \\ &= a^2 \mathbb{E}[(X - \mu)^2] \\ &= a^2 Var(X)\end{aligned}$$

Se sostituiamo i valori di **a** e **b** troviamo che:

SE  $a = 0 \rightarrow Var(b) = 0 \rightarrow$  le costanti hanno varianza **nulla**

SE  $a = 1 \rightarrow Var(X+b) = Var(X) \rightarrow$  sommando una const. non cambia la varianza

SE  $b = 0 \rightarrow Var(aX) = a^2 \cdot Var(X)$

## 6 Deviazione Standard

**Definizione:** Indica di quanto dei dati si **discostano dalla media** (non al quadrato)

**Formula generica:**

$$S = \sqrt{\text{Var}(X)}$$

$$\text{Var}(X + X) = \text{Var}(2 \cdot X) = 4 \cdot \text{Var}(X)$$

**Se X è indipendente allora:**

$$\text{Var}(X) + \text{Var}(X) = \text{Var}(X + X)$$

## 7 Covarianza

**Definizione:** Misura la **variazione** tra due variabili aleatorie associate tra di loro

**Formula generica:**

$$\text{Cov}(X, Y) := \mathbb{E}[(X - \mu_x)(Y - \mu_y)] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

**Dove:**

$$\mu_x = \mathbb{E}[X]$$

$$\mu_y = \mathbb{E}[Y]$$

*La covarianza può essere negativa, positiva o nulla*

**Positivo**  $\rightarrow$  Le due variabili crescono o decrescono insieme

**Negativo**  $\rightarrow$  Quando una variabile cresce l'altra decresce

**Nulla**  $\rightarrow$  Le due variabili sono indipendenti

### 7.1 Proprietà della covarianza

$$\text{Cov}(X, Y) = \text{Cov}(Y, X) \leftarrow \text{Commutativo}$$

$$\text{Cov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

$$\text{Cov}(X, X) = \text{Var}(X)$$

$$\text{Cov}(X + Y, Z) = \text{Cov}(X, Z) + \text{Cov}(Y, Z)$$

$$\text{Cov}(X + Y, Z + W) = \text{Cov}(X, Z) + \text{Cov}(X, W) + \text{Cov}(Y, Z) + \text{Cov}(Y, W)$$

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)$$

- Se  $X_1 \dots X_n$  e  $Y_1 \dots Y_n$  sono variabili aleatorie qualsiasi allora:

$$\text{Cov}\left(\sum_{i=1}^n x_i, \sum_{j=1}^m y_j\right) = \sum_{i=1}^n \sum_{j=1}^m \text{Cov}(X_i, Y_j)$$

Se  $X$  e  $Y$  sono **indipendenti**:

$$\mathbb{E}[XY] = \mathbb{E}[X]\mathbb{E}[Y]$$

$$\text{Cov}(X, Y) = 0 \longrightarrow \text{se sono indipendenti}$$

## 7.2 Coefficiente di correlazione lineare

**Definizione:** numero puro che tiene conto della deviazione standard di  $X$  e  $Y$

**Formula generica:**

$$\text{Corr}(X, Y) := \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \cdot \text{Var}(Y)}}$$

La correlazione può assumere valori compresi tra **-1** e **1**

**-1**  $\longrightarrow$  Le due variabili sono inversamente proporzionali

**0**  $\longrightarrow$  Le due variabili sono indipendenti

**1**  $\longrightarrow$  Le due variabili sono crescono o decrescono con lo stesso rapporto



## 8 Funzione generatrice dei momenti

**Definizione:** Funzione che ci permette di calcolare i momenti della distribuzione.

**Formula generica:**

$$\phi(t) = \mathbb{E}[e^{tX}]$$

Dove  $X$  è una variabile aleatoria e  $t$  è un parametro reale

$$\phi(t) = \mathbb{E}[e^{tX}] \begin{cases} \sum_x e^{tx} p(x) & \text{se } X \text{ discreta} \\ \int_{-\infty}^{+\infty} e^{tx} f(x) dx & \text{se } X \text{ continua} \end{cases}$$

Derivando la funzione si ottengono i momenti:

$$\phi'(t) = \frac{d}{dt} \mathbb{E}[e^{tX}] = \mathbb{E}\left[\frac{d}{dt} e^{tX}\right] = \mathbb{E}[X e^{tX}] \longrightarrow \phi'(0) = \mathbb{E}[X]$$

**Generalizzando:**

$$\phi^n(0) = \mathbb{E}[X^n]$$

Media:

$$\mu_x = \phi'(0)$$

Varianza:

$$\sigma_x^2 = \phi''(0) = \{\phi'(0)\}'$$

### 8.1 Disuguaglianza di Markov

**Definizione:** Ci permette di sapere la probabilità che una variabile assuma valori molto grandi

**Definizione:** Serve per calcolare che una variabile aleatoria assuma un minimo di "a"  $\longrightarrow a \in \mathbb{R}$

**Solo per variabili positive:**  $X \in (0, +\infty)$

**Formula generica:**

$$P(X \geq a) \leq \frac{\mathbb{E}[X]}{a}$$

**Generalizzazione:**

$$\begin{aligned} \mathbb{E}[X] &= \int_0^{+\infty} x f(x) dx \\ &= \int_0^a x f(x) dx + \int_a^{+\infty} x f(x) dx \geq \int_a^{+\infty} x f(x) dx \geq \int_a^{+\infty} a f(x) dx \\ &= a \int_a^{+\infty} f(x) dx \\ &= a P(X \geq a) \end{aligned}$$

## 8.2 Disuguaglianza di Chebyshev

**Definizione:** Ci permette di sapere la probabilità che una variabile si discosti dalla media per più di un certo numero di deviazioni standard.

$$X \text{ var aleatoria } \begin{cases} \mu & \text{Media} \\ \sigma^2 & \text{Varianza} \end{cases}$$

Per ogni  $r > 0 \rightarrow$  valore che indica il discostamento dalla media

**Formula generica:**

$$P(|X - \mu| \geq r) \leq \frac{\sigma^2}{r^2}$$

Serve per ottenere le stime di probabilità di eventi rari di variabili  
cui conosciamo solo la **media** e la **varianza**

## 9 Legge debole dei grandi numeri

**Definizione:** Dice che la probabilità che la differenza tra la media campionaria e il valore atteso superi una determinata soglia diventa sempre più piccola all'aumentare del numero di osservazioni

## 10 Modelli di variabili aleatorie

10.1 Bernoulli

10.2 Binomiali

10.2.1 Valore atteso e varianza di Binomiali

10.2.2 Funzione di massa e di ripartizione di Binomiali

10.3 Poisson

10.4 Ipergeometriche

10.5 Uniformi

10.6 Normali o Gaussiane

10.7 Esponenziali

10.8 Processi stocastici (Poisson)

10.9 Gamma

10.10 Chi-quadro

10.11 Distribuzione T

10.12 Distribuzione F

10.13 Distribuzione logistica