# STATISTICA

# Corso A

# Autore

Giuseppe Acocella 2024/25

Ultima Compilazione - March 16, 2025

# Contents

1	Sta	tistica Descrittiva	3
	1.1	Frequenze e Campioni	3
	1.2	Caratteri e Rappresentazioni Grafiche	3
		1.2.1 Classi di grafici	5
	1.3	Indici	6
		1.3.1 Indici di Centralità - (Media, Mediana, Moda)	6
		1.3.2 Indici di Dispersione - (Varianza, Deviazione Standard)	7
	1.4	Funzione di Ripartizione Empirica (FDR/ECDF)	7
	1.5	Quantili, Percentili, BoxPlot, Outlier	7
2	Dat	i Bivariati	8
	2.1	Scatterplot e Regressione	8
		2.1.1 Covarianza e Correlazioni Campionarie	8
		2.1.2 Th. Retta di Regressione ai Minimi Quadrati	9
3	$\mathbf{Pro}$	babilità	10
	3.1	Modello Uniforme	11
		3.1.1 Accenno di Combinatoria	12
	3.2	Probabilità Condizionata	12
		3.2.1 Regole Prodotto, Catena, Fattorizzazione della Probab. Condizionata	12
	3.3	Formula di Bayes	13
	3.4	Indipendenza	13
		3.4.1 Schema di Bernoulli	14
4	Var	iabili Aleatorie	14
	4.1	Variabili Aleatorie Discrete	14
	4.2	Variabili Aleatorie Notevoli	15
	4.3	Variabili Aleatorie con Densità	16
		4.3.1 Variabili Aleatorie con Densità Note	17
	4.4	Funzione di Ripartizione e Quantili	18
	4.5	Variabili Aleatorie Gaussiane	19
		4.5.1 Trasformazioni di Variabili Aleatorie con Densità	19
		4.5.2 Variabile Aleatoria Gaussiana Generale	19
	4.6	Valore Atteso e Momenti	20
		4.6.1 Calcolo e Proprietà del Valore Atteso	21

# 1 Statistica Descrittiva

Questo ramo della statistica cerca di raccogliere dati per descrivere degli oggetti. Elenchiamo delle definizioni standard:

- 1. Popolazione: Insieme di oggetti da studiare.
- 2. Carattere: Caratteristiche degli oggetti della popolazione.
  - (a) Colore di una biglia, altezza di un individuo.

Ricordiamo che un carattere può essere sia **qualitativo** (es. colore) sia **quantitativo** (es. altezza).

- 3. Modalità: Possibili valori che il carattere può assumere.
  - (a) Colore biglia istanziato: rosso, blu. Lancio moneta istanziato: testa/croce.
- 4. Campione Statistico (Sample): Sottoinsieme della popolazione scelto a rappresentarla.
- 5. Dati: Esiti delle misure del carattere sugli individui del campione.
  - (a) Lanci moneta:  $T, C, T, T, T, C, \dots$
- 6. Taglia Campione: Numero di elementi nel campione.

# 1.1 Frequenze e Campioni

Abbiamo due tipi di frequenze:

 Frequenza Assoluta: Corrisponde al numero di volte in cui la modalità appare nei dati:

$$\#\{\ i \mid x_i = a\ \}$$

 Frequenza Relativa: Corrisponde al numero di volte in cui la modalità appare nei dati fratto il numero dei dati stessi:

frequenza   
 relativa = 
$$\frac{\text{frequenza assoluta di } a}{\text{taglia campione}}$$

## 1.2 Caratteri e Rappresentazioni Grafiche

La rappresentazione dei dati dipende fortemente dal tipo di carattere:

- 1. Carattere Discreto: Quantità piccola e finita di modalità assumibili.
  - (a) Lancio di un dado, esiti di un sondaggio.

In questo caso per le rappresentazioni si utilizzano diagrammi a barre.

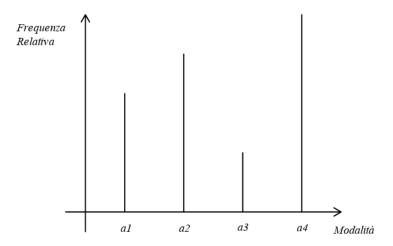


Figure 1: Esempio di diagramma a barre.

## 2. Carattere Continuo: Quantità assumibili in un intervallo continuo.

## (a) Altezza della popolazione.

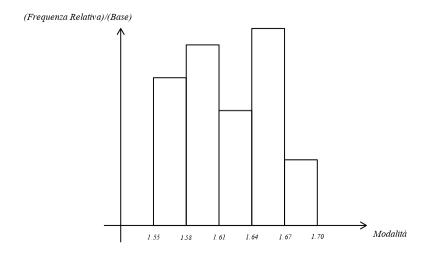


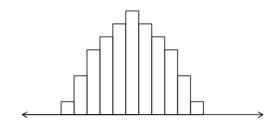
Figure 2: Esempio di istogramma.

La scelta di mettere sull'asse y il rapporto tra freq. relativa e base non è casuale, infatti se scegliessimo intervalli di ampiezza diversa si andrebbe in contro ad un errore di rappresentazione.

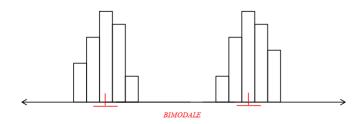
# 1.2.1 Classi di grafici

Elenchiamo qualche classificazione di rappresentazioni grafiche:

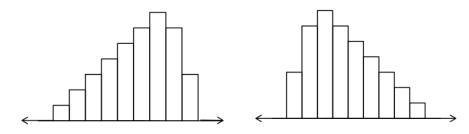
1. Normale: Simile ad una campana simmetrica:



2. Uni/Bi/Tri Modale: Si concentra attorno ad un numero k di colonne più alte:



(a) **Modale Asimmetrico Sx/Dx**: Si concentrano attorno ad una colonna più alta in maniera asimmetrica:



#### 1.3 Indici

Gli indici statistici sono quantità numeriche che riassumono proprietà significative sulla distribuizione dei dati.

## 1.3.1 Indici di Centralità - (Media, Mediana, Moda)

Descriviamo tre tipi di indici di centralità:

1. Media Campionaria: Descriviamo questo indice:

$$\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

ossia semplicemente la media aritmetica dei dati. Un modo **alternativo** è rappresentarlo così:

che formalmente si esprime così:

$$\overline{x} = \sum_{j=1}^{M} a_j \, p(a_j)$$

dove  $a_j$  sta per modalità e  $p(a_j)$  sta per frequenza relativa della modalità.

Sensibilità ai Valori Estremi Una delle caratteristiche della media campionaria è quella di essere molto sensibile ai valori estremi del campione.

Caratteristiche Riesce a vedere tutti i dati del campione e gode di alcune proprietà matematiche come la linearità.

- 2. Mediana Campionaria: Il dato  $x_i$  sarà centrale, dunque avrà metà dei dati a sinistra e metà a destra. La calcoliamo dunque in due modi:
  - (a) Numero dispari di modalità: Dato centrale.

$$mediana = x(\frac{n+1}{2})$$

(b) Numero pari di modalità: Media tra i due dati centrali.

mediana = 
$$\frac{1}{2}(x_{\frac{n}{2}} + x_{\frac{n}{2}+1})$$

6

Caratteristiche Più robusta rispetto ai valori estremi.

3. Moda: Modalità più frequente tra i dati.

#### 1.3.2 Indici di Dispersione - (Varianza, Deviazione Standard)

Gli indici di dispersione ci permettono di stabilire quanto i valori della distribuzione si allontanino da un valore centrale scelto come riferimento. Elenchiamoli:

1. Varianza Campionaria/Empirica: Permette di

CAMPIONARIA: 
$$Var(x) = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x})^2$$

EMPIRICA: 
$$Var_e(x) = \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^2 = (\frac{1}{n} \sum_{i=1}^{n} x_i^2) - \overline{x}^2$$

E' possibile calcolare la varianza anche con le frequenze relative:

$$Var_e(x) = (\sum_{j=1}^{M} a_j^2 * p(a_j)) - \overline{x}^2$$

2. Scarto Quadratico Medio: Indice basato sulla varianza.

$$\sigma(x) = \sqrt{Var(x)}$$

3. Indice Campionario di Asimmetria: Un indice che permette di stabilire se una distribuzione sia o meno asimmetrica:

$$b = \frac{1}{b^3} \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})^3$$

- (a) b > 0: Distribuzione Asimmetrica a destra.
- (b) b < 0: Distribuzione Asimmetrica a sinistra.

# 1.4 Funzione di Ripartizione Empirica (FDR/ECDF)

Dati  $x_1, x_2, \dots, x_n$  dati quantitativi definiamo  $F_e : \mathbb{R} \to \mathbb{R}$  data da

$$F_e(t) = \frac{\#\{\ i \mid x_i \le t\ \}}{n}$$

## 1.5 Quantili, Percentili, BoxPlot, Outlier

Dati  $x_1, x_2, ..., x_n$  dati quantitativi, k un numero naturale tra 0 e 100 allora, avendo un  $\beta = \frac{k}{100} \in (0, 1)$ :

7

**Definizione** k-esimo percentile o  $\beta$  – quartile è il dato  $r_{\beta} = x_i$  tale che:

- 1. Almeno il k% dei dati sia inferiore o uguale ad  $x_i$ .
- 2. Almeno il (100 k)% dei dati sia superiore o uguale ad  $x_i$ .

Se i due dati soddisfano questa condizione, si prende come k-esimo percentile la loro media aritmetica.

Outlier Valore anomalo che differisce in modo significativo dalla maggioranza dei dati.

In questo caso l'Outlier risulta essere 1001 e può essere causato da errori di misurazione o da dati molto rari. Gli outlier influenzano in modo significativo alcuni indici statistici: media, varianza. Altri indici sono invece poco influenzati: mediana, percentili.

# 2 Dati Bivariati

Consideriamo coppie di dati (due caratteri di un individuo di un campione), analizzando se esistono eventuali relazioni tra i dati  $x_i$  ed  $y_i$ .

# 2.1 Scatterplot e Regressione

In uno scatterplot sono riportati i dati bivariati  $(x_i, y_i)$  sul piano cartesiano.

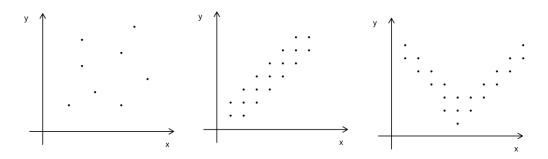


Figure 3: Da sinistra verso destra: Scatterplot Generico, Scatterplot Relazione Lineare, Scatterplot Relazione Non Lineare

**Regressione** Dato un insieme di punti  $(x_i, y_i)$  come possiamo determinare quale curva nel piano approssima meglio quell'insieme? Ci occuperemo solo di relazioni lineari, la cui classe di curve corrisponde alle rette.

#### 2.1.1 Covarianza e Correlazioni Campionarie

Definiamo prima la notazione necessaria:

$$\overline{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \qquad Var(x) = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \overline{x}) \qquad \overline{\sigma(x) = \sqrt{Var(x)}}$$

$$\overline{y} = \frac{1}{n} \sum_{i=1}^{n} y_i \qquad Var(y) = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \overline{y}) \qquad \overline{\sigma(y) = \sqrt{Var(y)}}$$

Covarianza Campionaria Empirica tra x ed y

$$Cov_e(x) = \frac{1}{n} \sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})$$

$$Cov(x) = \frac{n}{n-1}Cov_e(x,y)$$

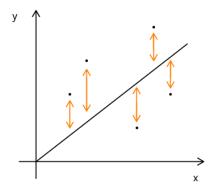
Coefficiente di Correlazione Campionario tra  $\mathbf{x}$  e  $\mathbf{y}$  Supponiamo  $\sigma(x) \neq 0$  e  $\sigma(y) \neq 0$  allora:

$$r(x,y) = \frac{Cov(x,y)}{\sigma(x) \sigma(y)}$$

## 2.1.2 Th. Retta di Regressione ai Minimi Quadrati

Vogliamo trovare la retta di equazione y = a + bx che meglio approssima un insieme di dati bivariati  $(x_i, y_i)$ . Risulta quindi necessario trovare a e b in modo tale da minimizzare

$$min_{a,b\in\mathbb{R}}\sum_{i=1}^{n}(y_i-(a+bx_i))^2$$



**Teorema** Ricaviamo dunque a, b della funzione y = ax + b:

$$b = b^* = \frac{Cov(x, y)}{Varx}$$
  $a = a^* = \overline{y} - b^* \overline{x}$ 

Dunque sostituendo nella funzione da minimizzare definita sopra:

$$min_{a,b \in \mathbb{R}} \sum_{i=1}^{n} (y_i - (a + bx_i))^2 = \sum_{i=1}^{n} (y_i - \overline{y})^2 (1 - r(x, y)^2)$$

La retta di regressione è quindi la retta che meglio approssima l'insieme di dati bivariati, la bontà dell'approssimazione lineare è tanto migliore tanto è più piccolo  $1-r^2$ .

# 3 Probabilità

Elenchiamo delle definizioni fondamentali per lo studio delle probabilità, ossia la descrizione degli esiti di un esperimento aleatorio:

- 1. Spazio Campionario: Insieme  $\Omega$  di tutti i possibili esiti dell'esperimento.
- 2. Esperimento Aleatorio: Fenomeno il cui esito non è determinato con certezza a priori.
- 3. Evento: Sottoinsieme dello Spazio Campionario, rappresentano affermazioni sull'esito dell'esperimento.

$$\Omega = \{1,2,3,4,5,6\}$$
 "Esce numero pari" =  $A = \{2,4,6\}$ 

- 4. Relazioni logiche ed operazioni insiemistiche: Elenchiamo alcune delle operazioni eseguibili:
  - (a) Accade A o B  $\mapsto$   $A \cup B$
  - (b) Accade A e B  $\mapsto$   $A \cap B$
  - (c) Non accade  $A \mapsto \overline{A} = \Omega A$
  - (d) Accade A ma non B  $\mapsto A B$
  - (e) Se accade A allora accade B  $\mapsto$   $A \subseteq B$
  - (f) A e B non possono accadere contemporaneamente  $\mapsto A \cap B = \emptyset$
- 5. **Definizioni Storiche di Probabilità**: Esistono diverse interpretazioni di probabilità:
  - (a) **Definizione Classica**: Viene così definita:

$$P(A) = \frac{\text{casi favorevoli ad A}}{\text{casi possibili}} = \frac{\#A}{\#\Omega}$$

(b) **Definizione Frequentista**: Viene così definita:

 $P(A) = \lim_{n \to \infty}$  (frequenza relativa di A in n prove ripetute dell'esperimento)

(c) **Definizione Soggettivista**: Viene così definita:

P(A) = grado di fiducia che accada A da parte di un soggetto razionale

L'obiettivo è dunque quello di associare un valore  $P(A) \in [0,1]$  ad un evento in modo tale da esprimere quanto sia probabile l'evento A in questione.

10

Regole Fondamentali di Probabilità Abbiamo visto che esistono varie interpretazioni di probabilità, di conseguenza risulta necessario trovare dei principi minimi che devono essere rispettati per definire una probabilità. Assumiamo un  $\Omega$  spazio campionario ed una funzione  $P: \mathbb{P}(\Omega) \to \mathbb{R}$  dove  $\mathbb{P}(\Omega)$  rappresenta l'insieme delle parti di  $\Omega$ :

- 1.  $0 \le P(A) \le 1 \quad \forall A \subseteq \Omega$
- 2.  $P(\Omega) = 1$
- 3. Se  $A_i$  è una successione di eventi a due a due disgiunti (ossia  $A_i \cap A_j = \emptyset \ \forall i \neq j$ ) allora:

$$P(A_1 \cup A_2 \cup ... \cup A_n) = \sum_{i=1}^{n} P(A_i)$$

Definiamo grazie a questo uno **spazio di probabilità**, ossia la coppia  $(\Omega, P)$ .

Evento Trascurabile ed Evento Quasi Certo : Definiamo due specifici tipi di eventi:

- 1. Evento Trascurabile: P(A) = 0.
- 2. Evento Quasi Certo: P(A) = 1.

**Proprietà di Probabilità** Elenchiamo delle proprietà a disposizione di qualunque probabilità:

- 1.  $P(\overline{A}) = 1 P(A)$
- 2.  $B \subseteq A \Rightarrow P(A B) = P(A) P(B)$
- 3. Dati  $A, B \subseteq \Omega$  non necessariamente disgiunti, allora:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

#### 3.1 Modello Uniforme

Un modello uniforme è uno spazio di probabilità  $(\Omega, P)$  tale che  $\Omega$  è finito ed ogni esito  $\omega \in \Omega$  risulta essere equiprobabile, ossia che  $P(\omega)$  risulta essere la stessa per ogni omega.

**Probabilità in Modelli** Se  $(\Omega, P)$  è un modello uniforme, allora:

$$P(A) = \frac{\#A}{\#\Omega} \quad \forall A \subseteq \Omega$$

#### 3.1.1 Accenno di Combinatoria

Elenchiamo diversi tipi di sequenze dato un insieme di n oggetti:

1. Numero di sequenze ordinate con ripetizione di k oggetti:

$$\#\{(a_1, a_2, ..., a_n)\} \mid a_j \in \{1, 2, ..., n\} = n^k$$

2. Numero di sequenze ordinate senza ripetizione di k oggetti:

$$\#\{(a_1, a_2, ..., a_n)\} \mid a_j \in \{1, 2, ..., n\} = \frac{n!}{(n-k)!}$$

3. Numero di sottoinsiemi di  $\{1, 2, 3, ..., n\}$  formati da k elementi  $(k \le n)$ :

numero sottoinsiemi = 
$$\frac{n!}{k!(n-k)!}$$

## 3.2 Probabilità Condizionata

Dato  $(\Omega, P)$  spazio di probabilità, B evento non trascurabile, A evento di probabilità condizionata di A dato B:

$$P(A|B) = \frac{P(A|B)}{P(B)}$$

P(A|B) corrisponde dunque alla probabilità che avvenga A sapendo che si verifica B.

## 3.2.1 Regole Prodotto, Catena, Fattorizzazione della Probab. Condizionata

Elenchiamo le regole della Probabilità Condizionata:

- 1. Regola Prodotto:  $P(A \cap B) = P(A|B) P(B)$
- 2. **Regola Catena**: E' possibile applicare la regola del prodotto *n* volte. In questo modo di definisce la catena.
- 3. **Regola Fattorizzazione**: La probabilità dell'unione dei rami di una rappresentazione ad albero<sup>1</sup> corrisponde alla somma delle probabilità dei rami.

Formula di Fattorizzazione Definiamo prima un sistema di alternative

(eventi  $B_1, B_2, ..., B_n$ ) caratterizzato da:

- 1. Eventi a due a due disgiunti.
- $2. \ \Omega = \bigcup_{i=1}^n B_i$
- 3.  $P(B_i) > 0 \quad \forall i$

Date quindi  $B_1, B_2, ..., B_n$  sistema di alternative, allora  $\forall A \subseteq \Omega$ :

$$P(A) = \sum_{i=1}^{n} P(A|B_i) P(B_i)$$

 $<sup>^{1}</sup>$ Vedi lezione 3.3, 20/02/2025.

## 3.3 Formula di Bayes

Siano A e B eventi non trascurabili, allora:

$$P(B|A) = \frac{P(A|B) P(B)}{P(A)}$$

In particolare se  $B_1, B_2, ..., B_n$  sistema di alternative, allora:

$$P(B_i|A) = \frac{P(A|B_i) P(B_i)}{\sum_{j=1}^{n} P(A|B_j) P(B_j)} \quad \forall i = 1, ..., n$$

**Applicazione e Funzionamento** : Solitamente questo teorema viene utilizzato per invertire il condizionamento, tipicamente in due casi:

- 1. Accade un evento A riferito ad un osservabile.
- 2. Vogliamo calcolare la probabilità di una possibile causa  $B_i$ .

## 3.4 Indipendenza

Vogliamo definire formalmente in concetto che la probabilità di un determinato evento A non cambia sapendo che accade B. Per A e B non trascurabili accade:

$$P(A) = P(A|B) = \frac{P(A \cap B)}{P(B)} \iff P(A)P(B) = P(A \cap B)$$

Dunque due **eventi** sono detti **indipendenti**<sup>2</sup> se:

$$P(A \cap B) = P(A) P(B)$$

Indipendenza a 3 o più Eventi Dati tre eventi A, B, C questi sono indipendenti se:

1. Sono a due a due indipendenti, ossia:

$$P(A \cap B) = P(A) P(B), P(A \cap C) = P(A) P(C), P(B \cap C) = P(B) P(C)$$

2. Vale che  $P(A \cap B \cap C) = P(A) P(B) P(C)$ 

Questo informalmente vuol dire che avere informazioni riguardo alcuni degli eventi non cambia la probabilità relativa degli altri eventi, quindi l'indipendenza a due a due non implica indipendenza globale.

**Definizione Indipendenza n-esima Generica** Dati  $A_1, ..., A_n$  eventi indipendenti, la probabilità di ogni possibile intersezione è il prodotto delle probabilità degli  $A_i$  coinvolti nell'intersezione:

$$P(A_{i1} \cap A_{i2} \cap ... \cap A_{ik}) = P(A_{i1}) P(A_{i2}) ... P(A_{ik}) \quad \forall k \leq n, \ \forall 1 \leq i_1 \leq i_2 \leq ... \leq i_k \leq n$$

<sup>&</sup>lt;sup>2</sup>Se due eventi sono disgiunti, allora non sono indipendenti, dato che se accade ad esempio il primo allora non può accadere il secondo.

#### 3.4.1 Schema di Bernoulli

Date n prove ripetute, in cui ciascuna prova può avere successo o insuccesso, la formulazione viene così definita:

$$\Omega = \{(w_1, w_2, ..., w_n) \mid w_i \in \{0, 1\}\} = \{0, 1\}^n$$

f = probabilità del successo della singola prova

$$P\{(w_1, w_2, ..., w_n)\} = f^{\#successi} (1 - f)^{\#insuccessi} = f^{\sum_{i=1}^{n} w_i} (1 - f)^{n - \sum_{i=1}^{n} w_i}$$

Da questo possiamo osservare che l'indipendenza non implica l'assenza di causalità e viceversa.

# 4 Variabili Aleatorie

Elenchiamo delle definizioni base:

- 1. Variabile Aleatoria (v.a): Carattere quantitativo dell'esperimento in esame. Ad esempio in n lanci di moneta X = #teste. Nello specifico la definiamo come una funzione  $X : \Omega \Rightarrow \mathbb{R}$ .
- 2. Legge/Distribuzione di X: Distribuzione di probabilità della caratteristica X in esame:

$$P^X(A) = P(X \in A) = P(\# \text{teste in A}) \quad A \subseteq \{0, 1, ..., n\}$$

Nello specifico possiamo anche definirla come probabilità  $P^X$  su  $\mathbb R$  data da

$$\{\omega \in \Omega \mid X(\omega) \in A\}$$

3. Equidistribuzione: Due variabili aleatore X ed Y sono dette equidistribuite se

$$P^X = P^Y$$

# 4.1 Variabili Aleatorie Discrete

Una variabile aleatoria X è detta **discreta** se può assumere un numero finito o numerabile di modalità. Ad esempio nei lanci di moneta, se

$$X = \text{teste di n lanci di moneta}, \quad \text{modalità} = \{0, 1, 2, ..., n\}$$

Questo è un esempio di **variabile discreta**. Risulta necessario ricordare che parliamo di possibili modalità e non di dati di uno specifico campione.

Funzione di Massa (Densità Discreta) Data una variabile aleatoria X definiamo pX:  $\{a_1, a_2, ..., a_n\} \to \mathbb{R}$  definita da:

$$p^X(a_i) = P\{X = a_i\} = P^X(\{a_i\})$$

Da questo possiamo ricavare un altra caratteristica:

Calcolo delle Probabilità Relative ad X Dagli statement definiti sopra otteniamo:

$$P(X \in A) = \sum_{a_i \in A} P(X = a_i)$$

**Proprietà delle Funzioni di Massa** Data una funzione di massa  $p^X$ , vale che:

- 1.  $p^X(a_i) \ge 0 \quad \forall a_i$
- 2.  $\sum_{i} p^{X}(a_{i}) = 1$

Viceversa, se una funzione q soddisfa queste due proprietà allora esiste una variabile aleatoria X che ha q come funzione di massa, e la legge di X è determinata univocamente da  $q = p^X$ .

Interpretazione delle Definizioni Date Mostriamo un esempio di estrazione da una popolazione, elencando tutti gli elementi in analisi:

- 1.  $\Omega = \{\text{popolazione}\}, \text{ assumiamo di avere } P \text{ uniforme, dunque ad esempio } \Omega = \{\text{italiani}\}, X = \#figli$
- 2.  $X(\omega) = \text{carattere quantitativo di } \omega$ .
- 3.  $p^X(a_i) = P(X = a_i) = \frac{\#\{\text{individui con } X = a_i\}}{\#\text{popolazione}} = \text{frequenza relativa di } \{X = a_i\}$  su tutta la popolazione.

#### 4.2 Variabili Aleatorie Notevoli

Questo sottocapitolo elencherà specifiche variabili aleatorie caratterizzate da funzioni di massa predefinite.

Elenchiamo alcune variabili aleatorie notevoli<sup>3</sup>:

1. Variabile Aleatoria Binomiale: Variabile Aleatoria Discreta di parametri  $n \in \mathbb{N}^+$ ,  $f \in [0,1]$  con funzione di massa data da:

$$p^{X}(k) = P(X = k) = \begin{pmatrix} x \\ y \end{pmatrix} f^{k} (1 - f)^{n-k} \quad k \in \{0, 1, 2, ..., n\}$$

Mostriamo un esempio di utilizzo a pagina successiva.

 $<sup>^3</sup>$ Capitoli 4.3, 4.4 delle note del corso 24/25.

Singola prova: lancio di dado, 5 prove ripetute, il successo corrisponde all'esito "6", di conseguenza la probabilità è di  $\frac{1}{6}$ . Quanto vale la probabilità che almeno due lanci siano uguali a "6"?

$$P(\text{almeno 2 lanci "6"}) = P(X \ge 2) = 1 - P(X < 2) =$$

$$= 1 - P(X = 0) - P(X = 1) =$$

$$= 1 - \left(\begin{array}{c} 5 \\ 0 \end{array}\right) \, \left(\frac{1}{6}\right)^0 \, \left(1 - \frac{1}{6}\right)^{5-0} - \left(\begin{array}{c} 5 \\ 1 \end{array}\right) \, \left(\frac{1}{6}\right)^1 \, \left(1 - \frac{1}{6}\right)^{5-1} \, \dots$$

2. **Variabile Aleatoria Geometrica**: Variabile Aleatoria Geometrica di parametro *p* con funzione di massa:

$$p^T(k) =$$

 $= f(1-f)^{k-1}$ , dove T (istante) = num. prova primo successo =  $min\{x \in \mathbb{N}^+ | \omega_n = 1\}$ 

3. Variabile Aleatoria Poisson: X Variabile Aleatoria di Poisson  $(Poisson(\lambda))$  di parametro  $\lambda > 0$ : variabile discreta, a valori a valori in  $\mathbb{N} = \{0, 1, 2, ...\}$ , con funzione di massa:

$$p^X(k) = p\{X = k\} = \frac{\lambda^k}{k!}e^{-k}$$

**Interpretazione** La variabile aleatoria X di Poisson conta il numero di eventi rari, ossia in numero di successi in n prove ripetute con n molto grande ed f probabilità di successo molto piccola (ossia in successo è raro), con  $\lambda = nf > 0$ .

## 4.3 Variabili Aleatorie con Densità

Le variabili aleatorie X con densità sono caratterizzate appunto da una funzione  $f: \mathbb{R} \to \mathbb{R}$  dove P(X) corrisponde a:

$$P(X \in A) = \int_A f(x)dx$$

Quindi P(X) corrisponde all'area sottesa alla funzione. Questo implica che ogni caratteristica sugli estremi dell'integrale influenzerà la densità ed il suo calcolo.

Interpretazione La densità calcolata sulla funzione f corrisponde alla distribuzione di X per unità di lunghezza. Questo tipo di variabili si utilizzano per caratteri continui.

Proprietà delle Densità Elenchiamo le proprietà delle densità:

1. 
$$f(x) \le 0 \quad \forall x \in \mathbb{R}$$

$$2. \int_{+\infty}^{-\infty} f(x)dx = 1$$

Di conseguenza se f soddisfa le due proprietà allora esiste almeno una variabile aleatoria X con densità f e la sua distribuizione è univocamente determinata da f.

#### 4.3.1 Variabili Aleatorie con Densità Note

Elenchiamo e descriviamo alcune variabili con densità note:

1. Variabili Aleatorie Uniformi: X variabile uniforme su un intervallo  $(\alpha, \beta)$  dato  $(U(\alpha, \beta))$  con densità:

$$f(x) = \begin{cases} \frac{1}{\beta - \alpha} & \text{se } x \in (\alpha, \beta) \\ 0 & \text{se } x \notin (\alpha, \beta) \end{cases}$$

La parte a tratti viene "fattorizzata" come un ulteriore funzione, ossia:

$$1_A(x) = \begin{cases} 1 & \text{se } x \in (\alpha, \beta) \\ 0 & \text{se } x \notin (\alpha, \beta) \end{cases}$$

Di conseguenza la funzione viene così definita:

$$f(x) = \frac{1}{\beta - \alpha} \, 1_{\alpha, \beta}(x)$$

**Interpretazione** Viene scelto un punto a caso nell'intervallo dato senza alcuna preferenza.

2. Variabili Aleatorie Esponenziali: X variabile esponenziale di parametro  $\lambda > 0$  e con densità:

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{se } x > 0 \\ 0 & \text{se } x \le 0 \end{cases}$$

Fattorizzando con la funzione 1(x) vista prima otteniamo  $f(x) = \lambda e^{-\lambda x} 1_{0,+\infty}(x)$ .

Interpretazione Tempo di attesa tra due eventi rari.

# 4.4 Funzione di Ripartizione e Quantili

Data  $X: \Omega \to \mathbb{R}$  variabile aleatoria,  $P_X$  legge di X e  $P_X(A) = P(X \in A)$  con  $A \in \mathbb{R}$ , definiamo la funzione di ripartizione (FdR) di X con:

$$F_X: \mathbb{R} \to [0,1]$$
 data da  $F_X(x) = P(X \le x) = P_X((-\infty, x]), \ x \in \mathbb{R}$ 

Proprietà di FdR Elenchiamo le proprieta della funzione di ripartizione:

- 1. Non decrescente.
- 2. Continua a destra.
- 3.  $\lim_{x\to+\infty} F(x) = 1$ ,  $\lim_{x\to-\infty} F(x) = 0$

Legame tra FdR e Densità Esiste un legame tra FdR e Densità:

$$F_X(x) = \int_{-\infty}^x f(y)dy$$

Questo accade se  $F_X$  è continua, viceversa se f è continua a tratti:

$$f(x) = F_X'(x)$$

Calcolo Probabilità Intervalli Calcoliamo la probabilità di un intervallo:

$$P(a \le X \le b) = F_X(b) - F_X(a)$$

**Probabilità e Beta-Quantili** Data X variabile aleatoria, dato  $\beta \in (0,1)$  allora  $\beta$ -quantile:

numero 
$$r_{\beta}$$
 tale che  $P(X \leq r_{\beta}) \geq \beta$  e  $P(X \geq r_{\beta}) \geq 1 - \beta$ 

#### 4.5 Variabili Aleatorie Gaussiane

Definiamo Z o N(0,1) variabile gaussiana (o normale) standard, nello specifico corrisponde ad una variabile aleatoria con densità:

$$y(x) = \frac{1}{\sqrt{2\pi}}e^{\frac{-x^2}{2}}$$

Funzione di Ripartizione e Beta Quantili di N

$$\Phi(x) = P(Z \le x) = \int_{x}^{-\infty} \frac{1}{\sqrt{2\pi}} e^{\frac{-y^2}{2}} dy \quad (\text{FdR})$$

$$q_{\beta} = \Phi^{-1}(\beta)$$
 (Beta-Quantile)

**Esempio di Utilizzo FdR** Mostriamo un esempio in cui risulta necessario utilizzare la FdR:

1. Immaginiamo di voler calcolare questa probabilità:

$$P(a < z < b) = \int_{a}^{b} y(x)dx$$

Non esiste un espressione esplicita per questo integrale.

2. Si usa allora la FdR:

$$P(a < z < b) = \Phi(b) - \Phi(a)$$

#### 4.5.1 Trasformazioni di Variabili Aleatorie con Densità

Data  $X: \Omega \to \mathbb{R}$  ed  $h: \mathbb{R} \to \mathbb{R}$  allora  $h \circ X: \Omega \to \mathbb{R}$  è anch'essa una variabile aleatoria. Ma  $h \circ X$  ammette densità? In generale no, ammette densità solo se h è invertibile.

\*\*\* Proposizione sul cambio di variabile \*\*\*

## 4.5.2 Variabile Aleatoria Gaussiana Generale

Dati  $Z=N(0,1),\,m\in\mathbb{R},\,\sigma>0,\,Y=h(z)=\sigma Z+m$  con densità:

$$f(y) = \frac{1}{\sigma}y(\frac{y-m}{\sigma}) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y-m)^2}{2\sigma^2}\right)$$

**Definizione** Y variabile aleatoria gaussiana di media m e varianza  $\sigma^2$  con densità:

$$f(y) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{(y-m)^2}{2\sigma^2}\right) \quad y \in \mathbb{R}$$

In generale dunque ci riferiremo a questo tipo di variabile aleatoria con  $(N(m, \sigma^2))$ .

Standardizzazione tra due tipi di N Dato  $(N(m, \sigma^2))$  possiamo dirigerci verso Z = N(0, 1) tramite standardizzazione:

$$Z = N(0,1) \Leftrightarrow y = \sigma Z + m = N(m, \sigma^2)$$
 quindi

$$P(a < Y < b) = P(\frac{a-m}{\sigma} < Z < \frac{b-m}{\sigma}) = \Phi\left(\frac{b-m}{\sigma}\right) - \Phi\left(\frac{a-m}{\sigma}\right) \ \ (\Phi \text{ FdR di } N(0,1))$$

## 4.6 Valore Atteso e Momenti

Data X variabile aleatoria **discreta**, a valori in  $a_1, a_2, ...$ , con funzione di massa p il valore atteso di X:

$$\mathbb{E}[X] = \sum_{j} a_j \, p(a_j)$$

**Interpretazione** Può essere visto come il baricentro di P, o la media delle  $a_j$  pesata con  $p(a_j)$ .

Valore Atteso su Variabile Aleatoria con Densità Data X variabile aleatoria con densità f, il valore atteso di X:

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

Il valore atteso in queste condizioni non è sempre finito, di conseguenza non è sempre definito. Quando è definito:

1. Se X variabile aleatoria (discreta o con densità) è  $\geq 0$ , allora:

$$\exists \ \mathbb{E}[X] = \sum_{j} a_{j} p(a_{j}) \text{ oppure } \int x f(x) dx \in [0, +\infty]$$

2. Per variabile aleatoria generica, X ammette valore atteso se:

$$\mathbb{E}[X] = \sum_{i} |a_{i}| p(a_{i}) \text{ oppure } \int |x| f(x) dx < +\infty$$

rispettivamente se è un caso discreto o con densità.

- 3. Criteri sufficienti per  $\exists \mathbb{E}[X] \in \mathbb{R}$ :
  - (a) Caso Discreto: Se  $a_i$  sono in numero finito.
  - (b) Caso con Densità: Se f=0 fuori da un intervallo limitato  $[\alpha,\beta]$ .

## 4.6.1 Calcolo e Proprietà del Valore Atteso

Calcolo Valore Atteso Data una variabile aleatoria X discreta o con densità, calcoliamo il valore atteso:

1. Se X è discreta con funzione di massa p:

$$\mathbb{E}[g(X)] = \sum_{j} g(a_j) p(a_j)$$

2. Se X ha densità f:

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

**Proprietà Valore Atteso** Siano X, Y variabili aleatorie date e  $a, b \in \mathbb{R}$ :

1. Linearità:

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

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

2. Monotonia:

(a) Se 
$$X \ge 0$$
, allora  $\mathbb{E}[X] \ge 0$ 

(b) Se 
$$X \geq Y$$
, allora  $\mathbb{E}[X] \geq \mathbb{E}[Y]$ 

Momento di Ordine n-esimo Data una variabile aleatoria X:

$$\mathbb{E}[X^n] = \begin{cases} \sum_j a_j^n P(X = a_j) & \text{caso discreto} \\ \int_{-\infty}^{+\infty} x^n f(x) dx & \text{caso con densità} \end{cases}$$