

# Matematica discreta e Probabilità

Leonardo Mengozzi

Le voci dell'indice, "←Indice" e i nomi delle dimostrazioni sono dei link alle rispettive sezioni.

## Indice

<b>1 Combinatoria</b>	<b>5</b>
Insiemi . . . . .	5
Definizione 1.1 (Cardinalità) . . . . .	5
Definizione 1.2 (Insieme finito) . . . . .	5
Definizione 1.3 (Cardinalità insiemi infiniti) . . . . .	5
Definizione 1.4 (Insieme numerabile) . . . . .	5
Definizione 1.5 (Insieme Discreto) . . . . .	6
Definizione 1.6 (Combinatoria di base) . . . . .	6
Definizione 1.7 (Prodotto Cartesiano) . . . . .	6
Definizione 1.8 (Sequenza) . . . . .	6
Definizione 1.9 (Insieme delle parti) . . . . .	6
Teorema 1.1 (Insieme delle parti) . . . . .	6
Principi di base . . . . .	7
Definizione 1.10 (Prodotto Condizionato) . . . . .	7
Definizione 1.11 (Disposizioni) . . . . .	8
Proposizione 1.1 (Fattoriale Discendente) . . . . .	8
Definizione 1.12 (Permutazioni) . . . . .	8
Definizione 1.13 (Combinazioni) . . . . .	8
Proposizione 1.2 (Numero Combinazioni) . . . . .	8
Definizione 1.14 (Combinazione di 1...n) . . . . .	9
Definizione 1.15 (Combinazione tipo (a,b)) . . . . .	10
Definizione 1.16 (Combinazione tipo (a,b,c)) . . . . .	10
Definizione 1.17 (Anagramma) . . . . .	10
Proposizione 1.3 (Binomio Combinazioni tipo (a,b,c)) . . . . .	10
Proposizione 1.4 (Somma binomi) . . . . .	11
Cammino Reticolare . . . . .	11
Numeri di Fibonacci . . . . .	12

Definizione 1.18 (Principio di inclusione esclusione) . . . . .	13
Definizione 1.19 (Sequenze k-piene) . . . . .	13
Definizione 1.20 (Scombussolamenti) . . . . .	14
Definizione 1.21 (Partizione) . . . . .	14
Definizione 1.22 (Numero di Bell) . . . . .	14
Triangolo di Bell . . . . .	15
Definizione 1.23 (Numero di Stirling) . . . . .	15
Triangolo di Stirling . . . . .	16
<b>2 Statistica Descrittiva</b>	<b>17</b>
Definizione 2.1 (Popolazione) . . . . .	17
Definizione 2.2 (Carattere) . . . . .	17
Definizione 2.3 (Unità (Statistica)) . . . . .	17
Definizione 2.4 (Campione) . . . . .	17
Definizione 2.5 (Modalità) . . . . .	17
Definizione 2.6 (Classi) . . . . .	17
Definizione 2.7 (Frequenza (assoluta)) . . . . .	18
Definizione 2.8 (Frequenza relativa) . . . . .	18
Definizione 2.9 (Istogramma) . . . . .	18
Definizione 2.10 (Media Campionaria) . . . . .	18
Definizione 2.11 (Linearità della media) . . . . .	19
Definizione 2.12 (Media Ponderata) . . . . .	19
Definizione 2.13 (Varianza) . . . . .	19
Proprietà 2.1 (Prima proprietà varianza) . . . . .	20
Proprietà 2.2 (Seconda proprietà varianza) . . . . .	20
Definizione 2.14 (Correlazione Positiva) . . . . .	21
Definizione 2.15 (Correlazione Negativa) . . . . .	21
Definizione 2.16 (Covarianza Campionaria) . . . . .	21
Definizione 2.17 (Retta ai minimi quadrati) . . . . .	21
Definizione 2.18 (Media geometrica) . . . . .	22
<b>3 Probabilità Base</b>	<b>23</b>
Definizione 3.1 (Fenomeno) . . . . .	23
Definizione 3.2 (Evento) . . . . .	23
Definizione 3.3 (Valutazione di probabilità) . . . . .	23
Definizione 3.4 (Evento coerente) . . . . .	23
Definizione 3.5 (Probabilità uniforme) . . . . .	24
Definizione 3.6 (Probabilità eventi) . . . . .	24
Definizione 3.7 (Non-esempio) . . . . .	24
Considerazioni elementari . . . . .	24
Definizione 3.8 (Probabilità condizionata) . . . . .	25

Definizione 3.9 (Formula delle probabilità totali) . . . . .	25
Definizione 3.10 (Formula di Bayes) . . . . .	25
Definizione 3.11 (Eventi indipendenti) . . . . .	26
<b>4 Probabilità Discreta</b>	<b>27</b>
Definizione 4.1 (Variabili Aleatorie) . . . . .	27
Definizione 4.2 (Funzione di ripartizione) . . . . .	27
Definizione 4.3 (Densità concreta discreta) . . . . .	27
Definizione 4.4 (Densità astratta discreta) . . . . .	28
Approfondimento Serie Geometrica . . . . .	28
Definizione 4.5 (Densità uniforme) . . . . .	28
Definizione 4.6 (Densità di Bernulli) . . . . .	28
Definizione 4.7 (Schema successo insuccesso) . . . . .	29
Definizione 4.8 (Ssi a prove indipendenti) . . . . .	29
Definizione 4.9 (Ssi senza rimpiazzo) . . . . .	29
Definizione 4.10 (Variabili tempo di primo successo) . . . . .	30
Definizione 4.11 (Variabili geometrice non modificate) . . . . .	30
Proprietà 4.1 (Mancanza di memoria V.g.n.M.) . . . . .	30
Definizione 4.12 (Variabili di Poisson) . . . . .	31
Teorema 4.1 (Variabile di Poisson) . . . . .	31
Definizione 4.13 (Variabili aleatorie multidimensionali) . . . . .	32
Definizione 4.14 (Variabile aleatoria bidimensionale) . . . . .	32
Definizione 4.15 (Variabili aleatorie indipendenti) . . . . .	32
Definizione 4.16 (Trasformazioni di variabili) . . . . .	33
Proprietà 4.2 (F.r. di v.a. indipendenti) . . . . .	33
Proprietà 4.3 (Complementare f.r. di v.a. indipendenti) . . . . .	33
Definizione 4.17 (Densità di una Trasformazione) . . . . .	34
Definizione 4.18 (Valore atteso) . . . . .	34
Proprietà 4.4 (Valore atteso) . . . . .	34
Teorema 4.2 (Linearità Valore Atteso) . . . . .	34
Corollario 4.1 (1 - Linearità valore atteso) . . . . .	35
Corollario 4.2 (2 - Linearità valore atteso) . . . . .	35
Corollario 4.3 (3 - Linearità valore atteso) . . . . .	36
Corollario 4.4 (4 - Linearità valore atteso) . . . . .	36
Corollario 4.5 (5 - Linearità valore atteso) . . . . .	36
Definizione 4.19 (Prodotto valore atteso di v.a. indipendenti) . . . . .	36
Definizione 4.20 (Varianza del valore atteso) . . . . .	37
Proprietà 4.5 (Semplificazione Varianza v.a.) . . . . .	37
Proprietà 4.6 (Prima della Varianza) . . . . .	37
Proprietà 4.7 (Seconda della Varianza) . . . . .	38
Corollario 4.6 (Varianza v.a. Binomiale) . . . . .	38

Teorema 4.3 (Disuguaglianza di Chebyshev) . . . . .	39
Definizione 4.21 (Convergenza) . . . . .	39
Teorema 4.4 (Legge dei grandi numeri) . . . . .	39
<b>5 Probabilità Continua</b>	<b>41</b>
Definizione 5.1 (Variabili aleatorie continue) . . . . .	41
Definizione 5.2 (Densità di variabile continua) . . . . .	41
Definizione 5.3 (Densità continua uniforme) . . . . .	42
Confronto Probabilità discreta e continua . . . . .	42
Definizione 5.4 (Funzione di ripartizione astratta) . . . . .	43
Definizione 5.5 (Densità continua) . . . . .	43
Definizione 5.6 (Densità continua astratta) . . . . .	43
Definizione 5.7 (Valore atteso di V.A. Continua) . . . . .	44
Proprietà 5.1 (Valore Atteso V.A. Continua) . . . . .	44
Definizione 5.8 (Varianza V.A. Continua) . . . . .	44
Proprietà 5.2 (Varianza V.A. Continua) . . . . .	45
Definizione 5.9 (Variabile esponenziale) . . . . .	45
Definizione 5.10 (Trasformazioni v.a. Continue) . . . . .	46
Definizione 5.11 (Variabili normali (Gaussiane)) . . . . .	46
Somma opposti . . . . .	46
Valore atteso e Varianza Variabile Normale . . . . .	47
Definizione 5.12 (Variabili Normali non Standard) . . . . .	47
Teorema 5.1 (Teorema Centrale del limite) . . . . .	47
Definizione 5.13 (Correzione di Continuità) . . . . .	49
<b>6 Statistica Inferenziale</b>	<b>50</b>
6.1 Significatività d'un test . . . . .	50
<b>7 Curva normale standardizzata</b>	<b>51</b>

# 1 Combinatoria

”Serve a contare quanti elementi ci sono in un insieme. (risponde alla domanda ”Quante sono?”).”

## Insiemi

Negli insiemi non conta l'ordine (infatti si usano le ”{}”, se contava l'ordine si usavano le ”()”) e gli elementi ripetuti. Insieme finito  $\{1, \dots, n\}$ . Insieme infinito  $\{1, \dots\}$ .

### Definizione 1.1 (*Cardinalità*)

Un insieme  $A$  ha cardinalità  $n$  se contiene esattamente  $n$  elementi, o equivalentemente se  $\exists$  una corrispondenza biunivoca  $A \longleftrightarrow \{1, 2, \dots, n\}$ .

La cardinalità si indica  $|A| = n$ .

### Definizione 1.2 (*Insieme finito*)

Un insieme  $A$  si dice finito se  $\exists n \in \mathbb{N}$  t.c.  $A$  contiene esattamente  $n$  elementi distinti.

Osservazione: L'insieme vuoto è l'unico insieme finito di cardinalità 0.

### Definizione 1.3 (*Cardinalità insiemi infiniti*)

Due insiemi infiniti hanno la stessa cardinalità se  $\exists$  una biezione (Corrispondenza biunivoca) tra di loro.

### Definizione 1.4 (*Insieme numerabile*)

Un insieme  $A$  si dice numerabile se ha la stessa cardinalità di  $\{1, 2, 3, \dots\}$ .

In altre parole un insieme è numerabile se i suoi elementi possono essere messi in una fila infinità.

Insiemi numerabili sono  $\mathbb{N}$  (anche se più grande di  $\{1, 2, 3, \dots\}$ ),  $\mathbb{Z} = \{0, 1, -1, 2, -2, 3, \dots\}$ ,  $\mathbb{Q}_{>0}$  dalla dimostrazione classica di Cantor. Un insieme non numerabile sono delle sequenze infinite di bit 0/1.

**Definizione 1.5 (*Insieme Discreto*)**

*Un insieme è discreto se è finito o numerabile.*

**Definizione 1.6 (*Combinatoria di base*)**

*Costruisce schemi "complessi" partendo da schemi semplici riuscendo a controllarne la cardinalità. (si opera solo con insiemi finiti).*

**Definizione 1.7 (*Prodotto Cartesiano*)**

*Siano  $A, B$  insiemi il cui prodotto cartesiano  $A \times B$  è l'insieme delle coppie ordinate  $(a, b), a \in A, b \in B$ . Generalizzando:*

$$A_1 \times A_2 \times \cdots \times A_k = \{(a_1, \dots, a_k) \mid a_i \in A_i, \forall i \in 1, \dots, k\}$$

Definiamo  $n$ -esima potenza cartesiana di  $n$  come  $A^n = A \times \cdots \times A$ .

**Definizione 1.8 (*Sequenza*)**

*Una sequenza (o lista) finita di lunghezza  $n$  di elementi di  $A$  è un elemento  $(a_1, \dots, a_n)$  del prodotto cartesiano  $A^n$ .*

Le *successioni* sono delle sequenze di lunghezza  $\infty$ , tipo  $\{a_1, \dots\}$ .

**Definizione 1.9 (*Insieme delle parti*)**

*Sia  $A$  un insieme, l'insieme delle parti  $\mathcal{P}(A)$  è l'insieme i cui elementi sono tutti i sotto insiemi di  $A$ , inclusi l'insieme vuoto  $\emptyset$  e  $A$  stesso. (insieme i cui elementi sono insiemi).*

La cardinalità dell'insieme delle parti è  $|\mathcal{P}(A)| = 2^{|A|}$ .

**Teorema 1.1 (*Insieme delle parti*)**

*Sia  $A$  un insieme di  $|A| = n$ , allora  $\exists$  una corrispondenza biunivoca tra  $\mathcal{P}(A)$  e  $\{0, 1\}^n$ .*

**Dimostrazione Teorema 1.1 :** Vediamo il caso particolare  $A = \{1, 2, 3\}$ .

$\{1\}$	$\{2\}$	$\{3\}$	$\{1, 2\}$	$\{1, 3\}$	$\{2, 3\}$	$\emptyset$	$A$
↑	↑	↑	↑	↑	↑	↑	↑
100	010	001	110	101	011	000	111

Per il caso generale  $|A| = n, A = \{a_1, \dots, a_n\}$  procediamo come prima facendo corrispondere ad'un sotto insieme  $S \subseteq A$  la sequenza binaria  $B$  il cui  $i$ -esimo bit è 1  $\iff a_i \in S$ . ■

## Principi di base

1. **Principio di ugualanza:** Siano  $A, B$  insiemi qualunque in corrispondenza biunivoca allora questi hanno lo stesso numero di elementi.

Generalizzazione del Principio di ugualanza:  $F : A \rightarrow B$  si dice  $k : 1$  ( $k$  a 1) se è surriettiva e a ogni elemento di  $B$  corrispondono esattamente  $k$  elementi di  $A$ . In questo caso  $|A| = k|B|$ . Il principio di ugualanza corrisponde al caso  $k = 1$ .

2. **Principio della somma:** Siano  $A, B$  insiemi qualunque disgiunti (non hanno elementi in comune), allora  $|A \cup B| = |A| + |B|$ .

Ricorda: Si dice Distinti se gli insiemi sono diversi per almeno un elemento.

3. **Principio del prodotto:** Siano  $A, B$  insiemi qualunque, allora  $|A \times B| = |A| \cdot |B|$ .

### Definizione 1.10 (*Prodotto Condizionato*)

Siano  $A, B$  insiemi.  $C \subseteq A \times B$  è sotto insieme del prodotto cartesiano e si dice prodotto condizionato di tipo  $(n, m)$  se sono soddisfatte queste condizioni:

1.  $\exists n$  elementi di  $A$  che compaiono come 1° coefficiente in un elemento di  $C$ .
2. Fissata la 1° coordinata di un elemento di  $C$ ,  $\exists m$  elementi di  $B$  che possono essere aggiunti come 2° coordinata.

In altri termini possiamo scegliere la 1° coordinata in  $n$  modi e fissata questa scelta possiamo scegliere la 2° in  $m$  modi.

Se per tutti i primi  $n$  coefficienti si hanno  $m$  secondi coefficienti si ha il prodotto condizionato, altrimenti no.

Nota: Una coordinata può non essere scelta.

**Definizione 1.11 (*Disposizioni*)**

Sia  $A = \{1, \dots, n\}$ , una disposizione di lunghezza  $k$  in  $A$  è una sequenza  $(a_1, \dots, a_k)$  di  $A$  t.c.  $a_i \neq a_j, \forall i \neq j$ .

**Proposizione 1.1 (*Fattoriale Discendente*)**

Le disposizioni di lunghezza  $k$  in  $\{1, \dots, n\}$  sono

$$\underbrace{n(n-1)\cdots(n-k+1)}_k = (n)_k$$

**Dimostrazione Proposizione 1.1:** Abbiamo  $n$  scelte per la prima coordinata. Fissata la prima coordinata ho  $n-1$  scelte per la seconda coordinata. Fissate le prime due coordinate ho  $n-2$  scelte per la terza coordinata, e così via. Abbiamo quindi un prodotto condizionato di tipo  $(n, n-1, \dots, n-k+1)$ . ■

**Definizione 1.12 (*Permutazioni*)**

Una permutazione di lunghezza  $n$  è una disposizione di lunghezza  $n$  in  $\{1, \dots, n\}$ .

Conseguenza: Dal Fattoriale Discendente, il numero di permutazioni di lunghezza  $n$  è  $(n)_n = 1 \cdot 2 \dots n = n!$ .

**Definizione 1.13 (*Combinazioni*)**

Sia  $A = \{1, \dots, n\}$ , una combinazione è un sotto insieme di  $A$  con  $a$  elementi. (l'ordine non importa).

**Proposizione 1.2 (*Numero Combinazioni*)**

Siano  $A, B, N$  tre insiemi di cardinalità  $a, b, n$  con  $a+b=n$ , allora questi 3 insiemi sono in biezione tra loro, ossia:

1.  $A$  sotto insieme di  $\{1, \dots, n\}$  con  $a$  elementi.
2.  $B$  sotto insieme di  $\{1, \dots, n\}$  con  $b$  elementi.
3. Si hanno sequenze binarie di lunghezza  $n$  con  $a$  volte 1 e  $b$  volte 0.

Questi insiemi hanno esattamente  $\frac{(n)_a}{a!} = \frac{(n)_b}{b!} = \frac{n!}{a!b!}$  elementi. Questo numero  $\frac{(n)_a}{a!} = \binom{n}{a} = \binom{n}{a,b}$  è il coefficiente binomiale.

Coefficiente binomiale visto diversamente:

$$(x+y)^n = \sum_{a=0}^n \binom{n}{a} x^a y^{n-a} = \sum_{a,b|a+b=n} \binom{n}{a,b} x^a y^b$$

*Osservazione:* Nel coefficiente binomiale considerare il numero sopra con "l'insieme di elementi" e quello sotto "quante posizioni in cui metterli".

**Dimostrazione Proposizione 1.2:** (*Parte1*) La corrispondenza tra 1 e 3 la si dimostra con il seguente esempio:  $n = 5, a = 3$

$$\begin{aligned}\{1, 2, 3\} &\leftrightarrow 11100 \\ \{1, 2, 4\} &\leftrightarrow 11010 \\ \{1, 2, 5\} &\leftrightarrow 11001 \\ \{1, 3, 4\} &\leftrightarrow 10110\end{aligned}$$

Tutte le sequenze hanno 3 "1" e 2 "0". Poi punto1 e punto2 della definizione sono in biezione data dall'complementare. (*Parte2*) Pensiamo di avere una funzione fatta così: Nel dominio abbiamo le disposizioni di lunghezza  $a$  in  $\{1, \dots, n\}$  e nel codominio abbiamo i sotto insiemi di  $\{1, \dots, n\}$  con  $a$  elementi. Per esempio:

$$\begin{aligned}(2, 5, 1) &\mapsto \{1, 2, 5\} \\ (3, 5, 4) &\mapsto \{3, 4, 5\} \\ (3, 4, 5) &\mapsto \{3, 4, 5\}\end{aligned}$$

Non è in biezione perché non è ignettiva. (Nel caso generale abbiamo una funzione  $a! : 1$ ).

Quindi le disposizioni sono  $A!$  volte i sotto insiemi e quindi i sotto insiemi con  $a$  elementi sono  $\frac{\binom{n}{a}}{a!} = \frac{n(n-1)\dots(b+1)}{a!} = \frac{n(n-1)\dots3\cdot2\cdot1}{a!b!} = \frac{n!}{a!b!}$ . ■

#### Definizione 1.14 (*Combinazione di $1\dots n$* )

Una combinazione di lunghezza  $a$  in  $\{1, \dots, n\}$  è un sotto insieme di  $\{1, \dots, n\}$  con  $a$  elementi.

**Definizione 1.15 (*Combinazione tipo (a,b)*)**

Una combinazione di tipo  $(a, b)$  con  $(a + b = n)$  è una coppia ordinata  $(A, B)$  di sotto insiemi di  $\{1, \dots, n\}$  con  $|A| = a, |B| = b, A \cup B = \{1, \dots, n\}$ .

**Definizione 1.16 (*Combinazione tipo (a,b,c)*)**

Una combinazione di tipo  $(a, b, c)$  con  $(a + b + c = n)$  è una terna ordinata  $(A, B, C)$  di sotto insiemi di  $\{1, \dots, n\}$  con  $|A| = a, |B| = b, |C| = c, A \cup B \cup C = \{1, \dots, n\}$ .

Esempio di riferimento (terna):

$$\begin{aligned} (\{1, 2\}, \{3\}, \{4\}) &\leftrightarrow (1, 1, 2, 3) \\ (\{1, 2\}, \{4\}, \{3\}) &\leftrightarrow (1, 1, 3, 2) \\ (\{1, 3\}, \{2\}, \{4\}) &\leftrightarrow (1, 2, 1, 3) \\ (\{2, 4\}, \{1\}, \{3\}) &\leftrightarrow (2, 1, 3, 1) \end{aligned}$$

**Definizione 1.17 (*Anagramma*)**

Sequenze binarie, ternarie, quaternarie, ecc, ordinate di numeri  $\{1, \dots, n\}$ , con  $n = a + b + \dots$ , di tipo  $(a, b, c, \dots)$  in biezione con le sequenze in cui compare  $a$  volte "1",  $b$  volte "2",  $c$  volte "3", ecc.

**Proposizione 1.3 (*Binomio Combinazioni tipo (a,b,c)*)**

Le combinazioni di tipo  $(a, b, c)$  con  $(a + b + c = n)$  sono  $\binom{n}{a, b, c} = \frac{n!}{a!b!c!}$ .

**Dimostrazione Proposizione 1.3:** Posso scegliere  $A$  in  $\binom{n}{a}$  modi. Fissato  $A$  posso scegliere  $B$  in  $\binom{n-a}{b}$  modi.  $C$  è univocamente determinato da  $A, B$ .

Il numero di combinazioni di tipo  $(a, b, c)$  è  $\binom{n}{a} \cdot \binom{n-a}{b} = \frac{n!}{a!(n-a)!} \cdot \frac{(n-a)!}{b!(n-a-b)!} = \frac{n!}{a!b!c!}$ . ■

Osservazione:  $\binom{n}{a, b, c} = \binom{n}{b, c, a} = \dots$ , in particolare con due indici abbiamo  $\binom{n}{a} = \binom{n}{b}$  se  $(a + b = n)$ .

**Proposizione 1.4 (*Somma binomi*)**

$$\left| \begin{array}{l} \binom{n}{a,b} = \binom{n-1}{a-1,b} + \binom{n-1}{a,b-1} \text{ che è uguale a } \binom{n}{a} = \binom{n-1}{a-1} + \binom{n-1}{a} \\ \binom{n}{a,b,c} = \binom{n-1}{a-1,b,c} + \binom{n-1}{a,b-1,c} + \binom{n-1}{a,b,c-1} \end{array} \right.$$

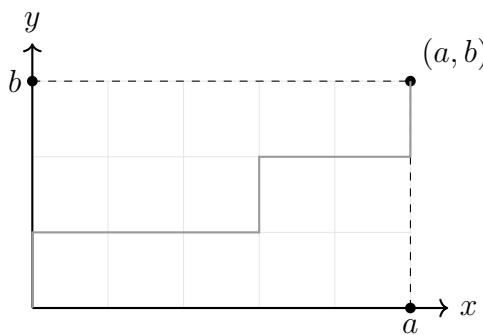
**Dimostrazione Proposizione 1.4:** Le combinazioni di tipo  $(a, b, c)$  sono in biezione con le combinazioni di tipo:

- $\{(A, B, C) \mid n \in A\}$  sono  $\binom{n-1}{a-1,b,c}$ .
- $\{(A, B, C) \mid n \in B\}$  sono  $\binom{n-1}{a,b-1,c}$ .
- $\{(A, B, C) \mid n \in C\}$  sono  $\binom{n-1}{a,b,c-1}$ . ■

**Cammino Reticolare**

Il cammino reticolare nel piano cartesiano è il percorso che congiunge l'origine degli assi a un punto  $(a, b)$  tramite *passi* verticali o orizzontali di 1.

Il cammino reticolare da un'ottima interpretazione grafica di una combinazione di tipo  $(a, b)$ .



**Quanti sono i cammini reticolari da  $(0,0)$  a  $(a,b)$ ?** Possiamo codificare i passi verticali con "1" e i passi orizzontali con "0". Così i cammini grafici diventano sequenze binarie. La lunghezza di tali sequenze è  $a + b$ , con  $a$  volte "0" e  $b$  volte "1". Abbiamo ottenuto un anagramma di tipo  $(a, b)$ . Osserviamo che abbiamo una biezione tra cammini e anagrammi e i cammini sono quindi  $\binom{a+b}{a,b}$ .

## Numeri di Fibonacci

Conta i sotto insiemi di  $\{1, \dots, n\}$  che non hanno 2 numeri consecutivi.

- $F_n$  = numero sotto insiemi di  $\{1, \dots, n\}$  senza consecutivi.
- $F_{n,k}$  = numero sotto insiemi di  $\{1, \dots, n\}$  di cardinalità  $k$  senza consecutivi.

$$F_n = \sum_{k=0}^n F_{n,k} = \sum_{k=0}^n \binom{n-k+1}{k}$$

**Dimostrazione:** Arriviamo alla formula con un esempio,  $F_{8,3}$ . Vediamo questi insiemi come sequenze di lunghezza 8 con 3 "1" non consecutivi (astrazione simile Camminio Reticolare). Devo riuscire ad'usare il principio d'equivalenza con un insieme di cui so la cardinalità. Sò che dopo il primo e secondo "1" c'è sempre uno "0", quindi cancello questo "0" così ho sequenze da 6. Ora "1" possono essere consecutivi ma posso tornare univocamente alla precedente rappresentazione, quindi è biunivoco. Deduciamo che  $F_{8,3} = \binom{6}{3} = 20$ .

In generale  $F_{n,k}$  sono le sequenze di lunghezza  $n$  con  $k$  volte "1", non consecutivi. Cancellando gli "0" che seguono i primi  $k-1$  "1" otteniamo una sequenza binaria lunga  $n-k+1$ , con  $k$  "1".

Tali sequenze sono  $F_{n,k} = \binom{n-k+1}{k}$  e quindi  $F_n = \sum_{k=0}^n \binom{n-k+1}{k}$ . A  $n$  non ci arrivo perchè il coefficiente binomiale da 0 quando  $k > n-k+1$ . ■

Versione ricorsiva di  $F_n$ ,  $\forall n \geq 3$  si ha  $F_n = F_{n-1} + F_{n-2}$ .

**Dimostrazione:** Suddividiamo i sotto insiemi di  $\{1, \dots, n\}$  senza consecutivi in due gruppi.

1.  $A_n$  quelli che contengono  $n$ .
2.  $B_n$  quelli che non contengono  $n$ .

Notare che  $F_n = F_{n-1} + F_{n-2}$ . In generale gli elementi di  $B_n$  non contengono  $n$ , quindi  $B_n = F_{n-1}$  e togliendo  $n$  agli elementi di  $A_n$  rimane un qualunque sotto insieme di  $\{1, \dots, n-2\}$  senza consecutivi. Quindi  $A_n = F_{n-2}$ . ■

**Definizione 1.18 (*Principio di inclusione esclusione*)**

*La cardinalità dell'unione di due insiemi:*

$$|A \cup B| = |A| + |B| - |A \cap B|$$

*La cardinalità dell'unione di tre insiemi:*

$$\begin{aligned} |A \cup B \cup C| &= |A| + |B| + |C| \\ &\quad - |A \cap B| - |A \cap C| - |B \cap C| \\ &\quad + |A \cap B \cap C| \end{aligned}$$

*La cardinalità dell'unione di  $n$  insiemi:*

$$|A_1 \cup \dots \cup A_n| = \sum_{\emptyset \neq I \subseteq \{1, \dots, n\}}^n (-1)^{|I|-1} |\bigcap_{i \in I} A_i|$$

Definiamo l'insieme complementare con  $(A_1, \dots, A_n)^C$ . Permette di risolvere tanti quesiti più facilmente. Per convenzione definiamo  $\bigcap_{i \in \emptyset} A_i = U$ , l'intero insieme. Quindi otteniamo che

$$|(A_1, \dots, A_n)^C| = \left| \bigcap_{i \in \emptyset} A_i \right| - \sum_{\emptyset \neq I \subseteq \{1, \dots, n\}}^n (-1)^{|I|-1} \left| \bigcap_{i \in I} A_i \right|$$

che è come dire  $\sum_{I \subseteq \{1, \dots, n\}}^n (-1)^{|I|} \left| \bigcap_{i \in I} A_i \right|$ .

**Definizione 1.19 (*Sequenze  $k$ -piene*)**

*Una sequenza  $k$ -piena è una sequenza in cui ogni numero da 1 a  $k$  compare almeno una volta. (non compaiono altri numeri).*

Usando il principio di inclusione esclusione scopro che il numero di sequenze  $k$ -piene lunghe  $n$  e con coefficienti in  $\{1, \dots, k\}$  è

$$|S| = \sum_{j=0}^k (-1)^j \binom{k}{j} (k-j)^n$$

**Definizione 1.20 (*Scombussolamenti*)**

Uno scombussolamento è una permutazione di  $\{1, \dots, n\}$  in cui per ogni  $i$  il numero non sta al posto  $i$ .

Usando il principio di inclusione esclusione scopro che il numero di scombussolamenti di  $n$  coefficienti è

$$|S| = \sum_{j=0}^n (-1)^j (n)_{n-j}$$

**Definizione 1.21 (*Partizione*)**

Una partizione di  $A$  è una suddivisione in sotto insiemi di  $A$ , cioè un insieme di blocchi della partizione  $\{S_1, S_2, S_3, \dots\}$  tale che:

1.  $S_i \cap S_j = \emptyset, \forall i \neq j.$
2.  $\bigcup_i S_i = A.$

(La partizione è la suddivisione e i blocchi sono i pezzi della partizione).

Nota: Le partizioni possono essere usate per "raggruppare" le permutazioni degli stessi elementi. Non è una fuzione biunivoca.

**Definizione 1.22 (*Numero di Bell*)**

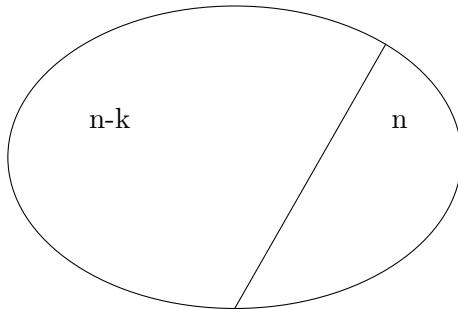
Il numero di partizioni di  $\{1, \dots, n\}$  si dice numero di Bell e si indica con  $B_n$ . Si calcola con la formula ricorsiva  $\forall n \geq 2$ :

$$B_n = \sum_{k=0}^{n-1} \binom{n-1}{k} B_k$$

Si assume  $B_0 = 1$  per convenzione.

**Dimostrazione Definizione 1.22:** Poniamo  $B_{n,k}$  come il numero di partizioni di  $\{1, \dots, n\}$  in cui  $n$  compare in un blocco con  $k$  elementi.

Calcoliamo  $B_{n,k}$ :



I numeri nel blocco che contiene  $n$  li possiamo scegliere in  $\binom{n-1}{k-1}$  modi.

Gli altri  $n-k$  numeri li possiamo partizionare in  $B_{n-k}$  modi. Implicando  $B_{n,k} = \binom{n-1}{k-1} B_{n-k}$ . Ottenendo così  $B_n = \sum_{k=1}^n \binom{n-1}{k-1} B_{n-k}$ .

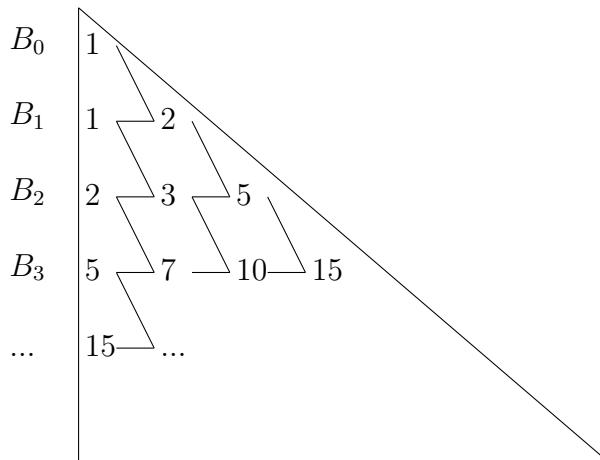
Equivalentemente a quella dell'enunciato ponendo  $h = n - k$ :

$$B_n = \sum_{h=0}^{n-1} \binom{n-1}{n-h-1} B_h \rightarrow \sum_{h=0}^{n-1} \binom{n-1}{h} B_h$$

■

## Triangolo di Bell

Trucco per calcolare il numero di Bell:



### Definizione 1.23 (*Numero di Stirling*)

Indichiamo con  $S_{n,k}$  il numero di partizioni di  $\{1, \dots, n\}$  in  $k$  blocchi. Si calcola con la formula ricorsiva:

$$S_{n,k} = S_{n-1,k-1} + kS_{n-1,k}$$

Si assume  $S_{n,n} = 1$  e  $S_{n,1} = 1$ .

**Dimostrazione Definizione 1.23:** Contiamo le partizioni in  $k$  blocchi in cui  $n$  è da solo. Gli altri  $n-1$  numeri devono essere partizionati in  $k-1$  blocchi:  $S_{n-1,k-1}$  scelte. Contiamo ora quelle in cui  $n$  non è in blocco da solo (procedendo per scelte successive):

- Partizioniamo i numeri da 1 a  $n-1$  in  $k$  blocchi:  $S_{n-1,k}$  scelte.
- Scegliamo in quale blocco inserire  $n$ :  $k$  scelte.

■

## Triangolo di Stirling

Trucco per calcolare i numeri di Stirling:

$$S_{1,1}$$

$$S_{2,1}$$

$$S_{2,2}$$

$$S_{3,1}$$

$$S_{3,2}$$

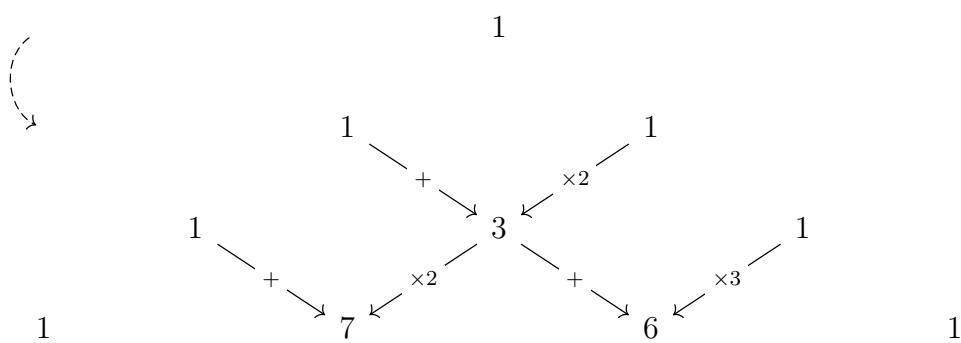
$$S_{3,3}$$

$$S_{4,1}$$

$$S_{4,2}$$

$$S_{4,3}$$

$$S_{4,4}$$



Ogni numero è la somma di quello in alto a sinistra più quello in alto a destra moltiplicato per il numero di diagonale.

## 2 Statistica Descrittiva

”Si occupa di presentare/descrivere i dati raccolti in un’indagine nel modo migliore possibile (sintetico, comunicativo).”

### Definizione 2.1 (*Popolazione*)

*Insieme su cui vogliamo effettuare un’indagine.*

### Definizione 2.2 (*Carattere*)

*Un carattere è quello che vogliamo studiare dalla popolazione.*

- **Carattere Quantitativo** se assumiamo valori numerici che esprimono una misura.
- **Carattere Qualitativo** se non è quantitativo.

I caratteri sarebbero distinguibili anche in **discreti e continui**.

### Definizione 2.3 (*Unità (Statistica)*)

*Un’unità (statistica) è un elemento della popolazione.*

### Definizione 2.4 (*Campione*)

*Un campione è un sotto insieme (rappresentativo) della popolazione del quale possiamo determinare il valore.*

### Definizione 2.5 (*Modalità*)

*Le modalità sono i valori che può assumere il carattere.*

### Definizione 2.6 (*Classi*)

*Una classe, solitamente, è determinata dal **confine inferiore e superiore**. Il **valore centrale** di una classe è la media dei confini.*

*Se abbiamo un solo confine chi fa l’indagine può decidere un valore rappresentativo come valore centrale.*

Se le modalità sono molto numerose (o infinite) è conveniente raggrupparle in **classi**.

**Definizione 2.7 (*Frequenza (assoluta)*)**

*La frequenza (assoluta) di una modalità è il numero di volte in cui compare nel campione.*

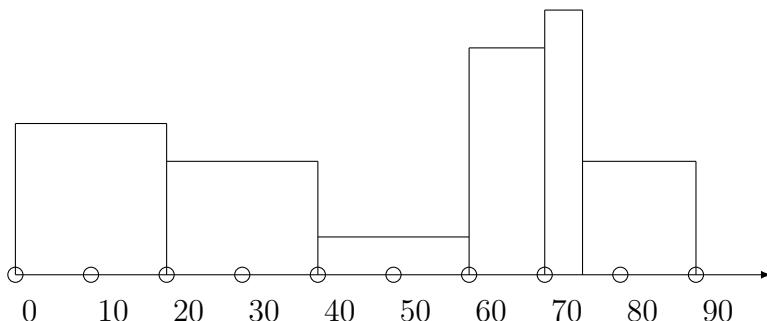
**Definizione 2.8 (*Frequenza relativa*)**

*La frequenza relativa è il rapporto fra la frequenza assoluta e la cardinalità del campione.*

$$F_R = \frac{F_A}{|campiono|}$$

**Definizione 2.9 (*Iistogramma*)**

*L'istogramma è un grafico che rappresenta i risultati di un'indagine. Ad'ogni modalità (classe) è associato un rettangolo con base proporzionale all'ampiezza e area proporzionale alla frequenza.*

**Definizione 2.10 (*Media Campionaria*)**

*La media campionaria si effettua su un carattere quantitativo su un campione di  $n$  elementi.*

$$\bar{x} = \bar{x}_n = \frac{1}{n}(x_1 + \dots + x_n)$$

Ricordiamo delle proprietà delle sommatorie:

- $\sum_{i=1}^n (a_i + b_i) = \sum_{i=1}^n a_i + \sum_{i=1}^n b_i$
- $\sum_{i=1}^n c = nc$
- $\sum_{i=1}^n (ca_i) = c \sum_{i=1}^n a_i$
- $\sum_{i=1}^n (\sum_{j=1}^n (a_i b_j)) = (\sum_{i=1}^n a_i) \cdot (\sum_{i=1}^n b_i)$

**Definizione 2.11 (*Linearità della media*)**

Siano  $x, y, z$  tre caratteri legati dalla relazione  $z = ax + by$  con costanti  $a, b$ . Allora

$$\bar{z} = a\bar{x} + b\bar{y}$$

Caso particolare è con  $y = 1$  ottenendo così  $z = ax + b$  e quindi  $\bar{z} = a\bar{x} + b$ . (Può essere usato per semplificare i conti).

**Dimostrazione Definizione 2.11 :**

$$\bar{z} = \frac{1}{n} \sum_{i=1}^n z_i = \frac{1}{n} \sum_{i=1}^n (ax_i + by_i) = \frac{1}{n} a \sum_{i=1}^n x_i + \frac{1}{n} b \sum_{i=1}^n y_i = a\bar{x} + b\bar{y}$$
■

**Definizione 2.12 (*Media Ponderata*)**

La media ponderata si usa per dare diversa importanza (peso) ai vari elementi di un campione.

$$\overline{x_w} = \frac{x_1 w_1 + \cdots + x_n w_n}{w_1 + \cdots + w_n}$$

*Osservazione:* La media campionaria si ottiene dando peso 1 a tutti gli elementi.

**Definizione 2.13 (*Varianza*)**

La varianza di  $x$  è la media dei quadrati della distanza da  $\bar{x}$ .

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

In altre parole indica la dispersione dei valori, ovvero da una misura di quanto i dati sono lontani tra loro.

**Dimostrazione Definizione 2.13 :** Ricordiamo la formula della parabola  $at^2 + bt + c = 0, a > 0$  e con vertice d'ascissa  $-b/2a$ . Possiamo fissare un punto a caso  $t$ . Consideriamo la media dei quadrati delle distanze date come:

$$\frac{(x_1 - t)^2 + \cdots + (x_n - t)^2}{n}$$

Per quale  $t$  questa quantità è minima?

$$\begin{aligned}\frac{1}{n} \sum_{i=1}^n (x_i - t)^2 &= \frac{1}{n} \left( \sum_{i=1}^n (t^2 - 2x_i t + x_i^2) \right) \\ &= t^2 - \frac{2t}{n} \sum_{i=1}^n x_i + \frac{1}{n} \sum_{i=1}^n x_i^2 \\ &= t^2 - 2\bar{x}t + \frac{1}{n} \sum_{i=1}^n x_i^2\end{aligned}$$

Il valore minimo lo otteniamo per  $t = \frac{-(-2\bar{x})}{2 \cdot 1} = \bar{x}$ . ■

*Osservazione:* Dal calcolo precedente con  $\bar{x}, \sigma_x^2 = \bar{x}^2 - 2\bar{x}\bar{x} - \bar{x}^2 = \bar{x}^2 - \bar{x}^2$ .

*Osservazione:* La varianza è sempre positiva.

### Proprietà 2.1 (*Prima proprietà varianza*)

*La varianza rimane invariata per traslazioni di costanti.*

$$\sigma_y^2 = \sigma_x^2$$

**Dimostrazione Proprietà 4.6:** Sia  $y = x + c$  con  $c$  costante. Otteniamo  $\bar{y} = \bar{x} + c$  ovvero  $\bar{y}^2 = \bar{x}^2 + 2c\bar{x} + c^2 = \bar{x}^2 + 2c\bar{x} + c^2$  che ci da

$$\sigma_y^2 = \bar{y}^2 - \bar{y}^2 = \bar{x}^2 + 2c\bar{x} + c^2 - (\bar{x}^2 + 2c\bar{x} + c^2) = \sigma_x^2$$

■

### Proprietà 2.2 (*Seconda proprietà varianza*)

*La varianza aumenta per prodotto con costanti.*

$$\sigma_y^2 = a^2 \sigma_x^2$$

**Dimostrazione Proprietà 4.7:** Sia  $y = ax$  con  $a$  costante. Otteniamo  $\bar{y} = a\bar{x}$  ovvero  $\bar{y}^2 = a^2\bar{x}^2 = a^2\bar{x}^2$  che ci da  $\sigma_y^2 = \bar{y}^2 - \bar{y}^2 = a^2\bar{x}^2 - a^2\bar{x}^2 = a^2\sigma_x^2$ . ■

**Definizione 2.14 (*Correlazione Positiva*)**

Dati due caratteri  $x$  e  $y$  diciamo che sono positivamente correlati se al crescere di uno ci aspettiamo che cresca anche l'altro.

**Definizione 2.15 (*Correlazione Negativa*)**

Dati due caratteri  $x$  e  $y$  diciamo che sono negativamente correlati se al crescere di uno ci aspettiamo che l'altro diminuisca.

**Definizione 2.16 (*Covarianza Campionaria*)**

La covarianza campionaria di  $x$  e  $y$  è

$$\sigma_{x,y} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \bar{xy} - \bar{x} \cdot \bar{y}$$

In altre parole se  $x$  e  $y$  sono positivamente correlate allora  $(x_i - \bar{x})$  e  $(y_i - \bar{y})$ . Ci aspettiamo che siano concordi e quindi che il prodotto sia maggiore di zero.

Osservazione:  $\sigma_{x,x} = \sigma_x^2$ .

**Dimostrazione Definizione 2.16 :**

$$\begin{aligned} \sigma_{x,y} &= \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \\ &= \frac{1}{n} \sum_{i=1}^n (x_i y_i - x_i \bar{y} - \bar{x} y_i + \bar{x} \cdot \bar{y}) \\ &= \bar{xy} - \cancel{\bar{x}}\cancel{y} - \bar{x} \cdot \bar{y} + \cancel{\bar{x}}\cancel{y} \\ &= \bar{xy} - \bar{x} \cdot \bar{y} \end{aligned}$$

■

**Definizione 2.17 (*Retta ai minimi quadrati*)**

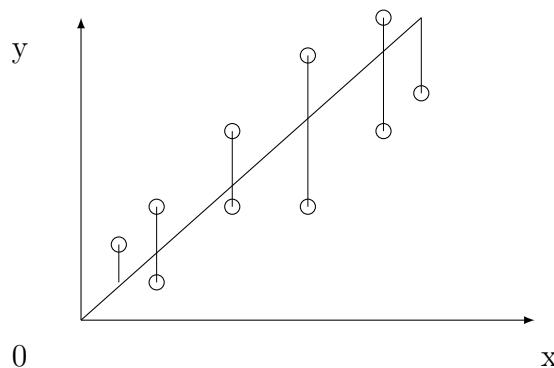
La retta ai minimi quadratici è la retta che minizza i quadrati degli errori (cioè delle lunghezze dei segmenti verticali).

La somma dei quadrati degli errori è

$$S(m, q) = \sum_{i=1}^n (y_i - mx_i - q)^2$$

vogliamo  $m$  e  $q$  in modo che questa quantità sia minima.

$$S(m, q) = \sigma_y^2 + m^2 \sigma_x^2 - em\sigma_{x,y} + (\bar{y} - q - m\bar{x})^2$$



*Osservazione:* Sia  $\bar{y} = m\bar{x} + q$ , il punto  $(\bar{x}, \bar{y})$  è retta è  $S(m, q) = \sigma_y^2 + m^2\sigma_x^2 - 2m\sigma_{x,y} \implies m = \frac{-b}{2a} = -\sigma_{x,y} \frac{1}{2\sigma_x^2} = \frac{\sigma_{x,y}}{\sigma_x^2}$ .

### Definizione 2.18 (*Media geometrica*)

La media geometrica permette di ottenere la variazione percentile media dei fattori moltiplicativi:

$$\bar{x}_g = \sqrt[n]{x_1 \cdot \dots \cdot x_n}$$

$$\bar{x}_g^n = x_1 \cdot \dots \cdot x_n$$

*Osservazione:* Ogni variazione percentile corrisponde ad'una moltiplicazione. La **media armonica** è l'inverso della media degli inversi.

### 3 Probabilità Base

”Si occupa di prevedere quanto è facile/possibile che qualcosa accada. Consiste in passaggi logici rigorosi partendo da un modello fisso (spazio di probabilità).”

#### Definizione 3.1 (*Fenomeno*)

*Un fenomeno è qualcosa che accade e che porta ad'un esito o risultato.*

Un fenomeno può essere:

- **Deterministico** se il risultato può essere predetto con esattezza.
- **Aleatorio** se il risultato è imprevedibile.

#### Definizione 3.2 (*Evento*)

*Un evento è un insieme di possibili risultati.*

#### Definizione 3.3 (*Valutazione di probabilità*)

*La valutazione di probabilità è una funzione che ad'ogni evento associa un numero tanto più grande quanto riteniamo che l'evento possa accadere.*

#### Definizione 3.4 (*Evento coerente*)

*Sia  $\Omega$  l'insieme dei risultati. Un evento è un sotto insieme di  $\Omega$  di cui ha senso calcolare la valutazione di probabilità, ossia:*

1. Se  $A$  e  $B$  sono eventi allora  $A \cap B$  è un evento.
2. Se  $A_1, \dots$  sono eventi allora  $\bigcup_i A_i$  è un evento.
3. Se  $A$  è un evento allora  $A^C$  è un evento.
4.  $\Omega$  è un evento.

*Osservazione:* Una collezione di sotto insiemi di  $\Omega$  che soddisfa tutti i presupposti si dice **famiglia coerente d'eventi**, con simbolo  $\mathcal{A}$ , ( $\mathcal{A} \subseteq \Omega$ ).

La valutazione di probabilità,  $P : \mathcal{A} \rightarrow \mathbb{R}_{\geq 0}$ , deve soddisfare le proprietà:

1.  $P(\Omega) = 1$ .
2.  $P(A) \geq 0, \forall$  evento  $A$ .
3. Se  $A_1, \dots$  sono eventi disgiunti allora

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

*Osservazione:*  $P$  non ha dominio  $\Omega$  ma l'insieme degli eventi. Non calcoliamo la  $P$  di un risultato ma di un evento.

*Osservazione:* Non tutti i sotto insiemi di  $\Omega$  sono eventi.

### Definizione 3.5 (*Probabilità uniforme*)

La probabilità è uniforme:

1. Se  $\Omega$  è finito con  $|\Omega| = n$ .
2. Se ogni sotto insieme di  $\Omega$  è un evento.
3. Se  $\omega_1$  e  $\omega_2$  sono due risultati allora  $P(\{\omega_1\}) = P(\{\omega_2\})$ .

Dalle proprietà della valutazione di probabilità deduciamo anche

$$P(\omega_1) + \cdots + P(\omega_n) = P(\Omega) = 1$$

In generale abbiamo quindi  $P(\omega) = \frac{1}{n}, \forall \omega \in \Omega$ .

### Definizione 3.6 (*Probabilità eventi*)

Sia  $A \subseteq \Omega, A = \{\omega_1, \dots, \omega_k\}$

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\#\text{risultati favorevoli}}{\#\text{risultati possibili}}$$

Ragionamento: Se  $|A| = k$ , allora

$$\begin{aligned} P(A) &= P(\omega_1) + \cdots + P(\omega_k) \\ &= \frac{1}{n} + \cdots + \frac{1}{n} = \frac{k}{n} \end{aligned}$$

*Osservazione:* Vale uno spazio con proprietà uniforme.

### Definizione 3.7 (*Non-esempio*)

Un non-esempio è qualcosa che non funziona.

## Considerazioni elementari

- Se  $E$  è un evento allora  $E^C$  è un evento.
- $E \cup E^C = \Omega$  è un'unione disgiunta.
- Se  $P(E) + P(E^C) = 1$  allora  $P(E^C) = 1 - P(E)$ .

*Osservazione:* Il principio di inclusione-esclusione vale anche per l'unione di più di due probabilità.

### Definizione 3.8 (*Probabilità condizionata*)

Chiamiamo probabilità condizionata la probabilità che accada l'evento  $B$  sapendo che accade l'evento  $A$  prima di  $B$ .

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

In altre parole riparametrizziamo la probabilità di  $B$  da  $\Omega$  al sotto insieme codiviso con  $A$ .

*Osservazione:* Ponendo  $P'(B) = P(B | A)$  allora  $P'$  soddisfa le proprietà delle valutazioni di probabilità:

- $P'(\Omega) = P(\Omega | A) = \frac{P(\Omega \cap A)}{P(A)} = 1$ .
- $P'(A_1 \cup A_2) = P'(A_1) + P'(A_2)$  se  $A_1, A_2$  sono disgiunti.

*Osservazione:*  $P(A \cap B) = P(B | A)P(A)$ .

### Definizione 3.9 (*Formula delle probabilità totali*)

Sia  $\Omega$  partizionato in  $\{A_1, A_2, \dots\}$ . Per sapere la probabilità d'un evento  $B$  condizionato da un qualsiasi altro evento.

$$P(B) = \sum_{i=1}^n P(B \cap A_i) = \sum_{i=1}^n P(B | A_i)P(A_i)$$

### Definizione 3.10 (*Formula di Bayes*)

Siano  $A, B$  eventi, allora la formula di Bayes

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

inverte le probabilità condizionate,  $P(B | A) \leftrightarrow P(A | B)$ .

**Dimostrazione Definizione 3.10:** Siano  $P(A | B) = \frac{P(A \cap B)}{P(B)}$  e  $P(B | A) = \frac{P(A \cap B)}{P(A)}$ . Allora  $P(A \cap B) = P(A | B)P(B) = P(B | A)P(A) = P(A \cap B)$ . ■

**Definizione 3.11 (*Eventi indipendenti*)**

*A e B sono eventi indipendenti se sapere che accade uno dei due non cambia la probabilità che accada l'altro.*

$$P(B | A) = P(B)$$

$$P(A | B) = P(A)$$

Osservazione:  $P(A \cap B) = P(A)P(B)$  dato che  $P(A | B) = \frac{P(A \cap B)}{P(B)} = P(A)$ .

## 4 Probabilità Discreta

### Definizione 4.1 (*Variabili Aleatorie*)

Sia  $\Omega$  uno spazio di probabilità,  $A$  un insieme coerente di eventi e  $P$  la valutazione di probabilità.

Una variabile aleatoria è una funzione  $X : \Omega \rightarrow \mathbb{R}$  tale che  $\forall a \in \mathbb{R}, \{\omega \in \Omega \mid X(\omega) \leq a\}$ , oppure  $X \leq a$ , e deve essere un evento. (si scrive anche  $\chi_A$ ).

Per una variabile aleatoria non ha senso scrivere  $P(X)$ , ma ha senso scrivere  $P(X \leq a)$ . Una variabile aleatoria può essere continua o discreta.

### Definizione 4.2 (*Funzione di ripartizione*)

Sia  $X$  una variabile aleatoria. La sua funzione di ripartizione  $F_X$  è una funzione  $F_X : \mathbb{R} \rightarrow \mathbb{R}$  definita come

$$F_X(t) = P(X \leq t)$$

### Definizione 4.3 (*Densità concreta discreta*)

Sia  $X$  una variabile aleatoria discreta. La densità concreta discreta di  $X$  è  $d_X : \mathbb{R} \rightarrow \mathbb{R}$  con

$$d_X(h) = P(X = h) = \begin{cases} a_1 & \text{se } h = 0, \dots \\ \dots & \\ 0 & \text{altrimenti} \end{cases}$$

La densità concreta discreta soddisfa le proprietà:

1.  $d_X(h) \geq 0, \forall h$ .
2.  $d_X(h) \neq 0$  per una quantità finita o numerabile di valori  $h$ .
3.  $\sum_h d_X(h) = 1$ .

Osservazione: Una densità è valida se soddisfa le tre proprietà.

**Definizione 4.4 (*Densità astratta discreta*)**

La densità astratta discreta è  $d : \mathbb{R} \rightarrow \mathbb{R}$  e soddisfa le proprietà:

1.  $d(h) \geq 0, \forall h$ .
2.  $d(h) \neq 0$  per una quantità discreta di valori  $h$ .
3.  $\sum_h d(h) = 1$ .

**Approfondimento Serie Geometrica**

Sia  $0 < a < 1$  allora  $1 - a^n = (1 - a)(1 + \dots + a^{n-1})$  sapendo che  $(1 + \dots + a^{n-1}) = \frac{1-a^n}{1-a}$  possiamo dire che

$$\lim_{h \rightarrow \infty} (1 + \dots + a^{n-1}) = \lim_{h \rightarrow \infty} \frac{1 - a^n}{1 - a} = \frac{1}{1 - a} = \sum_{h=0}^{\infty} a^h$$

**Definizione 4.5 (*Densità uniforme*)**

Sia  $X$  una v.a., si dice uniforme nell'insieme  $A = \{a_1, \dots, a_n\} \subseteq \mathbb{R}$  se

$$d_X(h) = \begin{cases} \frac{1}{n} & \text{se } h = a_1, \dots, a_n \\ 0 & \text{altrimenti} \end{cases}$$

Si scrive  $X \sim U(\{\dots\})$ .

**Definizione 4.6 (*Densità di Bernulli*)**

Si dice densità di Bernulli se  $X$  assume solo i valori 0 e 1.

$$d_X(h) = \begin{cases} p & \text{se } h = 1 \\ 1 - p & \text{se } h = 0 \\ 0 & \text{altrimenti} \end{cases}$$

Si scrive  $X \sim B(1, p)$ .

*Osservazione:* Si usa quando qualcosa può accadere o non accadere.

**Definizione 4.7 (*Schema successo insuccesso*)**

Lo schema successo insuccesso è una sequenza di  $n$  fenomeni aleatori ognuno dei quali da un successo oppure un insuccesso.

**Definizione 4.8 (*Ssi a prove indipendenti*)**

Uno schema successo insuccesso si dice a prove indipendenti se:

1. Ogni prova ha la stessa probabilità di dare successo.
2. L'esito di ogni tentativo è indipendente da quello degl'altri.

La variabile  $X = \#\text{successi}$  la chiamiamo binomiale. Scriviamo  $X \sim B(n, p)$  con  $n$  il numero di tentativi e  $p$  la probabilità che ogni tentativo dia successo.

$$d_X(k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & \text{se } k = 1, \dots, n \\ 0 & \text{altrimenti} \end{cases}$$

**Definizione 4.9 (*Ssi senza rimpiazzo*)**

Uno schema successo insuccesso si dice senza rimpiazzo se:

1. L'ordine delle estrazioni non è importante.
2. Ogni prova non ha la stessa probabilità di dare successo.
3. L'esito di ogni tentativo non è indipendente da quello degli altri.

Sia  $\Omega$  le combinazioni, con probabilità uniforme, di  $n$  palline tra blu e rosse,  $|\Omega| = \binom{b+r}{n}$ . La variabile  $X = \#\text{successi}$  (estrazione pallina blu) la chiamiamo ipergeometrica. Scriviamo  $X \sim H(h, b, r)$  con  $h$  il numero di tentativi.

$$P(X = k) = d_X(k) = \begin{cases} \frac{\binom{b}{k} \cdot \binom{r}{h-k}}{\binom{b+r}{h}} & \text{se } k = 0, \dots, h \\ 0 & \text{altrimenti} \end{cases}$$

I risultati favorevoli a " $X = k$ " sono  $\binom{b}{k} \cdot \binom{r}{h-k}$  ossia la scelta di  $k$  palline blue per la scelta di  $h - k$  palline rosse.

**Definizione 4.10 (Variabili tempo di primo successo)**

Sia  $\Omega$  le combinazioni, con probabilità uniforme, di  $n$  palline tra blu e rosse,  $|\Omega| = \binom{b+r}{n}$ . Si parla di variabile tempo di primo successo quando estraiamo senza rimpiazzo fino ad'ottenere un successo (blu).

Definiamo  $T = \#\text{tentativi effettuati}$ . Il tempo di primo successo in uno s.s.i. a prove indipendenti definisce  $T$  come tempo primo successo che assume  $\infty$  valori e  $p$  come probabilità di avere successo ad'ogni tentativo.

$$p(T = k) = (1 - p)^{k-1}p, d_T(k) = \begin{cases} (1 - p)^{k-1}p & \text{se } k = 1, \dots \\ 0 & \text{altrimenti} \end{cases}$$

Diciamo  $T \sim \tilde{G}(p)$ , ossia  $T$  ha densità geometrica modificata di parametro  $p$ . Siccome  $d_T$  è una densità abbiamo:  $\sum_{k=1}^{\infty} (1 - p)^{k-1}p = 1$ .

**Definizione 4.11 (Variabili geometrice non modificate)**

In uno s.s.i. a prove indipendenti la variabile geometrica  $X$  da il numero di fallimenti prima di un successo.  $X + 1 = T$  è una variabile geometrica modificata con  $X$  che assume valori da 0 a  $+\infty$ . Quindi

$$\begin{aligned} d_X(k) &= P(X = k) = P(T - 1 = k) = P(T = k + 1) \\ &= d_T(k + 1) = \begin{cases} p(1 - p)^k & \text{se } k = 0, \dots \\ 0 & \text{altrimenti} \end{cases} \end{aligned}$$

Scriviamo  $X \sim G(p)$ .

**Proprietà 4.1 (Mancanza di memoria V.g.n.M.)**

$$P(X \geq k + m \mid X \geq k) = P(X \geq m)$$

**Dimostrazione Proprietà 4.1:** Sia  $X \sim G(p)$ . Perchè  $P(X \geq k)$  è uguale a  $(1 - p)^k$ , ossia i primi  $k$  tentativi vanno male?

$$\begin{aligned} P(X \geq k + m \mid X \geq k) &= \frac{P(X \geq k + m \cap X \geq k)}{P(X \geq k)} \\ &= \frac{(1 - p)^{k+m}}{(1 - p)^k} = (1 - p)^m = P(X \geq m) \quad \blacksquare \end{aligned}$$

**Definizione 4.12 (Variabili di Poisson)**

Una variabile aleatoria  $X$  si dice variabile di Poisson di parametro  $\lambda > 0$  se ha la seguente densità:

$$d(k) = \begin{cases} e^{-\lambda} \frac{\lambda^k}{k!} & k = 0, \dots \\ 0 & \text{altrimenti} \end{cases}$$

Scriviamo  $X \sim P(\lambda)$ .

Osservazione:  $d(k)$  è una densità discreta astratta:

- $d(k) \geq 0, \forall k$ .
- $d(k) \neq 0, \forall k \in \mathbb{N}$ .
- $\sum_{k=0}^{\infty} d(k) = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1$ .

Ricorda:  $e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}, \forall x \in \mathbb{R}$ .

**Teorema 4.1 (Variabile di Poisson)**

Sia  $X$  una variabile binomiale,  $X \sim B(n, p)$ , e poniamo  $\lambda = np$  ottenendo  $X \sim B(n, \lambda/n)$ . Per  $n \rightarrow \infty$  la densità di  $X$  tende alla densità di Poisson  $P(\lambda)$ .

**Dimostrazione Teorema 4.1 :** La densità di  $X$  è

$$\begin{aligned} d_X(k) &= \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\ &= \frac{n(n-1)(n-k+1)}{k!} \frac{\lambda^k}{n^k} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-k} \\ &= \underbrace{\frac{\lambda^k}{k!}}_{e^\lambda} \underbrace{\frac{n(n-1)\dots(n-k+1)}{n^k}}_1 \underbrace{\left(1 - \frac{\lambda}{n}\right)^n}_{e^{-\lambda}} \underbrace{\left(1 - \frac{\lambda}{n}\right)^{-k}}_1 \\ &= e^\lambda e^{-\lambda} = 1 \end{aligned}$$

■

**Definizione 4.13 (Variabili aleatorie multidimensionali)**

Una v.a.  $n$ -dimensionale è data da  $(X_1, \dots, X_n)$  dove  $X_1, \dots, X_n$  sono variabili aleatorie.

**Definizione 4.14 (Variabile aleatoria bidimensionale)**

Una variabile aleatoria bidimensionale  $(X, Y)$  ha densità  $d_{X,Y} : \mathbb{R}^2 \rightarrow \mathbb{R}$ .

$$d_{X,Y}(h, k) = P(X = h, Y = k)$$

La densità bidimensionale  $d_{X,Y}$  si chiama **densità congiunta**. Le densità delle variabili  $X$  e  $Y$ ,  $d_X$  e  $d_Y$ , si chiamano **densità marginali**.

$$\begin{array}{c|ccc} X \setminus Y & 0 & \dots & n \\ \hline 0 & n_{00} & \dots & n_{0n} \\ \vdots & \vdots & \ddots & \vdots \\ n & n_{n0} & \dots & n_{nn} \end{array} = P(X = 0) = P(X = \dots) = P(X = n)$$

Le marginali sono univocamente determinabili dalla congiunta:

$$d_X(h) = \sum_k d_{X,Y}(h, k) \quad d_Y(k) = \sum_h d_{X,Y}(h, k)$$

Le densità marginali non determinano la densità congiunta.

Osservazione:  $d_{X,Y}(a, b)d_{X,Y}(c, d) = d_{X,Y}(a, d)d_{X,Y}(c, b)$ .

**Definizione 4.15 (Variabili aleatorie indipendenti)**

Due variabili aleatorie  $X, Y$  si dicono indipendenti se

$$P(X = h, Y = k) = P(X = h)P(Y = k)$$

Ossia  $d_{X,Y}(h, k) = d_X(h)d_Y(k), \forall h, k \in \mathbb{R}$ .

In altre parole  $X$  e  $Y$  sono indipendenti se gli eventi " $X = h$ " e " $Y = k$ " sono indipendenti per ogni  $h, k$ .

**Definizione 4.16 (*Trasformazioni di variabili*)**

Una trasformazione di una o più variabili aleatorie è data da

$$\Phi(X_1, X_2), \Phi : \mathbb{R}^2 \rightarrow \mathbb{R} \text{ oppure } \Phi(X), \Phi : \mathbb{R} \rightarrow \mathbb{R}$$

**Proprietà 4.2 (*F.r. di v.a. indipendenti*)**

Sia  $X$  e  $Y$  variabili indipendenti e  $Z$  una trasformazione di variabile allora

$$F_Z = F_X \cdot F_Y$$

Ricorda: La funzione di ripartizione della variabile aleatoria  $X$  è  $F_X(h) = P(X \leq h)$ .

**Dimostrazione Proprietà 4.2:** Sia  $Z = \max(X, Y)$

$$\begin{aligned} F_Z(k) &= P(Z \leq k) = P(\max(X, Y) \leq k) \\ &= P(X \leq k, Y \leq k) \end{aligned}$$

$$\begin{aligned} (\text{indipendenza}) &= P(X \leq k)P(Y \leq k) \\ &= F_X(k)F_Y(k) \end{aligned}$$

■

**Proprietà 4.3 (*Complementare f.r. di v.a. indipendenti*)**

Sia  $X$  e  $Y$  variabili indipendenti e  $Z$  una trasformazione di variabile allora

$$(1 - F_Z(k)) = (1 - F_X(k))(1 - F_Y(k))$$

**Dimostrazione Proprietà 4.3:** Sia  $Z = \min(X, Y)$

$$\begin{aligned} 1 - F_Z(k) &= 1 - P(Z \leq k) \\ &= P(Z > k) = P(\min(X, Y) > k) \\ &= P(X > k, Y > k) \end{aligned}$$

$$\begin{aligned} (\text{indipendenza}) &= P(X > k)P(Y > k) \\ &= (1 - F_X(k))(1 - F_Y(k)) \end{aligned}$$

■

**Definizione 4.17 (*Densità di una Trasformazione*)**

Sia  $X$  una variabile aleatoria e  $\Phi(X), \Phi : \mathbb{R} \rightarrow \mathbb{R}$  allora la densità della trasformazione è

$$d_{\Phi(X)}(k) = \sum_{h|\Phi(h)=k} d_X(h)$$

per una variabile bidimensionale

$$d_{\Phi(X,Y)}(k) = \sum_{i,j|\Phi(i,j)=k} d_{X,Y}(i,j)$$

**Definizione 4.18 (*Valore atteso*)**

Il valore atteso di  $X$  è

$$E[X] = \sum_h d_X(h)h$$

detto anche media o speranza matematica.

**Proprietà 4.4 (*Valore atteso*)**

Sia  $X$  variabile aleatoria e  $\Phi : \mathbb{R} \rightarrow \mathbb{R}$  allora

$$E[\Phi(X)] = \sum_h \Phi(h)d_X(h)$$

Analogamente  $E[\Phi(X, Y)] = \sum_{h,k} \Phi(h, k)d_{X,Y}(h, k)$ .

**Teorema 4.2 (*Linearità Valore Atteso*)**

Il valore atteso è lineare

$$E[aX + bY] = aE[X] + bE[Y]$$

**Dimostrazione Teorema 4.2:** Sia  $\Phi(X, Y) = aX + bY$  per la formula analoga della proprietà del valore atteso.

$$\begin{aligned}
E[aX + bY] &= \sum_{h,k} (ah + bk)d_{X,Y}(h,k) \\
&= \sum_{h,k} ah d_{X,Y}(h,k) + \sum_{h,k} bk d_{X,Y}(h,k) \\
&= a \sum_h h \underbrace{\sum_k d_{X,Y}(h,k)}_{d_X(h)} + b \sum_k k \underbrace{\sum_h d_{X,Y}(h,k)}_{d_Y(k)} \\
&= a \sum_h h d_X(h) + b \sum_k k d_Y(k) \\
&= aE[X] + bE[Y]
\end{aligned}$$

■

**Corollario 4.1 (1 - Linearità valore atteso)**

Sia $X \sim B(n, p)$	$E[X] = np$
----------------------	-------------

**Dimostrazione Corollario 4.1:** Sia  $X = X_1 + \dots + X_n$  con  $X_i \sim B(1, p), \forall i \in \{1, \dots, n\}$ . Per esempio

$$X_1 = \begin{cases} 1 & \text{se esistono tentativi che danno successo} \\ 0 & \text{se da insuccesso} \end{cases}$$

per linearità  $E[X] = E[X_1] + \dots + E[X_n] = p + \dots + p = np$ .

■

**Corollario 4.2 (2 - Linearità valore atteso)**

Sia $X \sim H(n, b, r)$	$E[X] = n \frac{b}{b+r}$
-------------------------	--------------------------

**Dimostrazione Corollario 4.2:** Sia  $X = X_1 + \dots + X_n$  con  $X_i \sim B(1, p), \forall i \in \{1, \dots, n\}$ . Per esempio

$$X_1 = \begin{cases} 1 & \text{se i-esima è blu} \\ 0 & \text{se i-esima è rossa} \end{cases}$$

Per linearità  $E[X] = E[X_1] + \dots + E[X_n] = \frac{b}{b+r} + \dots + \frac{b}{b+r} = n \frac{b}{b+r}$ .

■

**Corollario 4.3 (3 - Linearità valore atteso)**

Sia $X \sim P(\lambda)$	$E[X] = \lambda$
-------------------------	------------------

**Corollario 4.4 (4 - Linearità valore atteso)**

Sia $X \sim \tilde{G}(p)$	$E[X] = \frac{1}{p}$
---------------------------	----------------------

**Dimostrazione Corollario 4.4:** Sapendo che  $d_X(k) = p(1-p)^{k-1}$  e usando il "trucchetto" ( $0 < x < 1$ )

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x} \xrightarrow{\text{derivo}} \sum_{k=1}^{\infty} kx^{k-1} = \frac{1}{(1-x)^2}$$

Ponendo  $x = 1 - p$  ottengo

$$E[X] = \sum_{k=1}^{\infty} kp(1-p)^{k-1} = p \sum_{k=1}^{\infty} k(1-p)^{k-1} = p \frac{1}{(1-(1+p))^2} = \frac{1}{p} \quad \blacksquare$$

**Corollario 4.5 (5 - Linearità valore atteso)**

Sia $X \sim G(p)$ ossia $X + 1 \sim \tilde{G}(p)$	$E[X + 1] = E[X] + 1 = \frac{1-p}{p}$
---	---------------------------------------

**Definizione 4.19 (Prodotto valore atteso di v.a. indipendenti)**

Siano  $X$  e  $Y$  sono indipendenti allora

$$E[XY] = E[X]E[Y]$$

Se  $X \sim B(1, p)$  ho  $E[XY] = E[X^2] = E[X] = p$ . Perchè  $X$  vale solo 0 o 1.

**Dimostrazione Definizione 4.19:** Sia  $\Phi(X, Y) = XY$ .

$$\begin{aligned} E[XY] &= \sum_{h,k} \Phi(h, k) d_{X,Y}(h, k) = \sum_{h,k} h k d_X(h) d_Y(k) \\ &= \sum_h h d_X(h) \sum_k k d_Y(k) \\ &= E[X]E[Y] \end{aligned} \quad \blacksquare$$

Attenzione non vale il viceversa. La definizione può valere anche se le v.a. non sono indipendenti ma **incorrelate**.

#### **Definizione 4.20 (Varianza del valore atteso)**

*La varianza del valore atteso è il valore atteso dei quadrati degli scarti rispetto al valore atteso di  $X$ .*

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

#### **Proprietà 4.5 (Semplificazione Varianza v.a.)**

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

**Dimostrazione Proprietà 4.5 :**

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

■

#### **Proprietà 4.6 (Prima della Varianza)**

$$\text{Var}(aX) = a^2\text{Var}(X)$$

**Dimostrazione Proprietà 4.6 :**

$$\begin{aligned} \text{Var}(aX) &= E[(aX)^2] - E[aX]^2 \\ &= a^2E[X^2] - a^2E[X]^2 \\ &= a^2(E[X^2] - E[X]^2) \\ &= a^2\text{Var}(X) \end{aligned}$$

■

**Proprietà 4.7 (*Seconda della Varianza*)**

Siano  $X$  e  $Y$  variabili aleatorie indipendenti.

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

**Dimostrazione Proprietà 4.7:**

$$\begin{aligned} \text{Var}(X + Y) &= E[(X + Y)^2] - E[X + Y]^2 \\ &= E[X^2 + 2XY + Y^2] - (E[X] + E[Y])^2 \\ &= \underline{E[X^2]} + \underbrace{2E[XY]}_{2E[X]E[Y]} + \underline{E[Y^2]} - \underline{E[X]^2} - \cancel{2E[X]E[Y]} - \cancel{E[Y]^2} \\ &= \underline{\text{Var}(X)} + \underline{\text{Var}(Y)} \end{aligned}$$

■

**Corollario 4.6 (*Varianza v.a. Binomiale*)**

Sia  $X \sim B(n, p)$ .

$$\text{Var}(X) = np(1 - p)$$

**Dimostrazione Corollario 4.6:** Sia  $X = X_1 + \dots + X_n$  con  $X_i \sim B(1, p), \forall i \in 1, \dots, N$  indipendenti. Sia la varianza di una variabile aleatoria di bernulli  $p(1 - p)$ .

$$\begin{aligned} \text{Var}(X) &= \text{Var}(X_1) + \dots + \text{Var}(X_n) \\ &= p(1 - p) + \dots + p(1 - p) \\ &= np(1 - p) \end{aligned}$$

■

Attenzione: Questo ragionamento non vale per  $X \sim H(n, b, r)$ . Infatti in questo caso possiamo scrivere  $X = X_1 + \dots + X_n$  con  $X \sim B(1, \frac{b}{b+r})$ . Ma le  $X_i$  non sono indipendenti.

Se  $X \sim P(\lambda)$  ossia  $X \sim B(n, \frac{\lambda}{n})$  per  $n \rightarrow \infty$  si ha  $\text{Var}(X) = \lambda$ .

**Teorema 4.3 (Disuguaglianza di Chebyshev)**

Sia  $X$  una variabile aleatoria discreta, allora  $\forall \epsilon > 0$

$$P(|X - E[X]| > \epsilon) \leq \frac{\text{Var}(X)}{\epsilon^2}$$

**Dimostrazione Teorema 4.3:** Sia  $A = "|X - E[X]| > \epsilon"$  la distanza di  $X$  dal valore atteso di  $\epsilon$ .

$$\chi_A = \begin{cases} 1 & \text{se } A \text{ accade} \\ 0 & \text{se } A \text{ non accade} \end{cases}$$

Siano  $Y = \epsilon^2 \chi_A$  e  $Z = (X - E[X])^2$ . Se  $Y = \epsilon^2$  vuol dire che  $\chi_A = 1$  e quindi  $|X - E[X]| > \epsilon$ . Quindi  $Y < Z$ . Se  $Y = 0$  e quindi  $Y \leq Z$ . In tutti i casi  $Y \leq Z$  si ha  $E[Y] \leq E[Z]$ .

$$E[Y] = \epsilon^2 P(|X - E[X]| > \epsilon)$$

$$E[Z] = \text{Var}(X)$$

■

**Definizione 4.21 (Convergenza)**

Sia  $X_1, \dots$  una successione di variabili aleatorie. Diciamo che questa successione converge in probabilità ad  $l$  se  $\forall \epsilon > 0$

$$\lim_{n \rightarrow \infty} P(|X_n - l| > \epsilon) = 0$$

**Teorema 4.4 (Legge dei grandi numeri)**

Sia  $X_1, \dots$  una successione di variabili aleatorie discrete aventi tutte la stessa densità e indipendenti. Sia  $\mu = E[X_1]$ .

Poniamo  $\bar{X}_n = \frac{X_1 + \dots + X_n}{n}$ . Allora la successione  $\bar{X}_n$  converge in probabilità a  $\mu$ .

Esempio: Lancio un dato  $n$  volte e  $k$  è il numero di volte in cui abbiamo ottenuto 6. Consideriamo  $\frac{k}{n}$ . Per  $n$  piccolo questo rapporto può variare ma se  $n$  è grande ci aspettiamo che sia vicino a  $\frac{1}{6}$ .

**Dimostrazione Teorema 4.4:** Applichiamo la disuguaglianza di Chebyshev alla variabile  $\bar{X}_n$ .

$$P(|\bar{X}_n - E[\bar{X}_n]| > \epsilon) < \frac{\text{Var}(\bar{X}_n)}{\epsilon^2}$$

Siano:

$$E[\bar{X}_n] = E\left[\frac{X_1 + \dots + X_n}{n}\right] = \frac{1}{n}(E[X_1]) + \dots + E[X_n] = \mu$$

$$\begin{aligned} \text{Var}(\bar{X}_n) &= \text{Var}\left(\frac{X_1 + \dots + X_n}{n}\right) = \frac{1}{n^2}(\text{Var}(X_1) + \dots + \text{Var}(X_n)) \\ &= \frac{1}{n^2}n\text{Var}(X_1) = \frac{\text{Var}(X_1)}{n} \end{aligned}$$

Sostituendo otteniamo

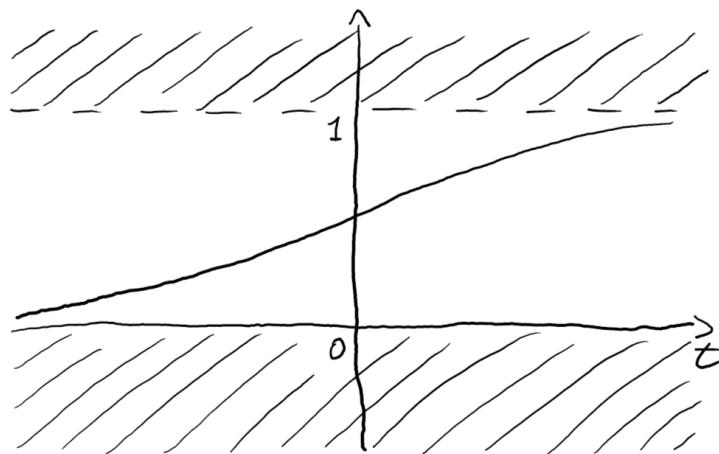
$$P(|\bar{X}_n - \mu| > \epsilon) < \frac{\text{Var}(X_1)}{n\epsilon^2} \xrightarrow{n \rightarrow \infty} 0$$

Ossia  $\lim_{n \rightarrow \infty} P(|\bar{X}_n - \mu| > \epsilon) = 0$ . ■

## 5 Probabilità Continua

**Definizione 5.1 (Variabili aleatorie continue)**

Una variabile aleatoria si dice continua se la funzione di ripartizione ( $F_X(t) = P(X \leq t)$ ) è una funzione continua.



- $F_X(t)$  è sempre debolmente monotona, ossia  $t_1 < t_2$ ,  $F_X(t_1) \leq F_X(t_2)$ .
- Non ci chiederemo la probabilità di valori esatti ma di intervalli.

$$P(a < x \leq b) = P(x \leq b) - P(x \leq a)$$

Più  $b$  è vicino ad  $a$  più la probabilità tende a 0, quindi  $P(x = t) = 0, \forall t \in \mathbb{R}$ .

**Definizione 5.2 (Densità di variabile continua)**

Sia  $f_X(s)$  densità di  $X$  se  $\forall a, b | a < b$  si ha

$$P(a < X < b) = \int_a^b f_X(s)ds$$

L'area sotto il grafico tra  $a$  e  $b$  è la probabilità.

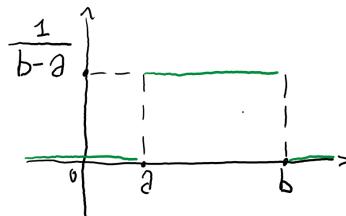
$$\int_a^b f_X(s)ds = F_X(b) - F_X(a)$$

$F$  è primitiva di  $f$ , quindi  $F'_X = f_X$ .

$$F_X(t) = \int_{-\infty}^t f_X(s)ds$$

**Definizione 5.3 (*Densità continua uniforme*)**

$X$  si dice uniforme in un intervallo  $[a, b]$  se la densità di  $X$  è costante in  $[a, b]$  e nulla al di fuori di  $[a, b]$ .



*Osservazione:* Siccome  $P(-\infty < x < +\infty) = 1$  l'area compresa tra il grafico della densità e l'asse orizzontale è sempre 1.

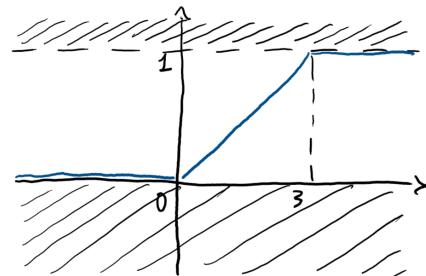
*Osservazione:*  $X$  è uniforme in  $[a, b]$  nel senso che  $\forall a_1, b_1, a_2, b_2$  con  $b_1 - a_1 = b_2 - a_2 = l$ .

$$P(a_1 < x < b_1) = P(a_2 < x < b_2) = \frac{l}{b-a}$$

**Confronto Probabilità discreta e continua**

$$X \sim B(3, \frac{1}{2})$$

$$d_X(k) = \begin{cases} \frac{1}{8} & \text{se } k = 0, 3 \\ \frac{3}{8} & \text{se } k = 1, 2 \\ 0 & \text{altrimenti} \end{cases}$$



$$P(x \leq a) = \int_{-\infty}^a f_X(s) ds = 0$$

$$P(x \leq 1) = 1$$

$$F_X(t) = P(X \leq t)$$

$$= \int_a^t \frac{1}{b-a} ds = \frac{t-a}{b-a}$$

**Definizione 5.4 (*Funzione di ripartizione astratta*)**

Una funzione  $F : \mathbb{R} \rightarrow \mathbb{R}$  si dice funzione di ripartizione astratta se:

1.  $F$  è una funzione continua.
2.  $\lim_{t \rightarrow -\infty} F(t) = 0, \lim_{t \rightarrow +\infty} F(t) = 1$ .
3.  $F_X$  è debolmente crescente, cioè  $\forall t_1 < t_2, F(t_1) \leq F(t_2)$ .

**Definizione 5.5 (*Densità continua*)**

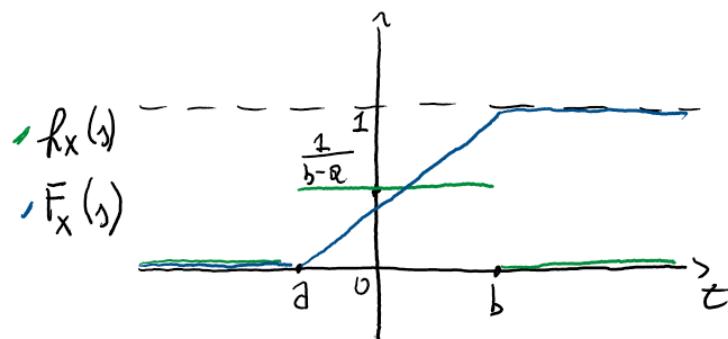
Una funzione  $f_X(s)$  si dice densità continua di  $X$  se  $\forall a, b$  con  $a < b$  si ha

$$P(a < X < b) = \int_a^b f_X(s)ds$$

**Definizione 5.6 (*Densità continua astratta*)**

Una funzione  $f : \mathbb{R} \rightarrow \mathbb{R}$  si dice densità continua astratta se:

1.  $f(s) \geq 0, \forall s \in \mathbb{R}$ .
2.  $\int_{-\infty}^{\infty} f(s)ds = 1$ .



**Definizione 5.7 ( *Valore atteso di V.A. Continua*)**

Il valore atteso di una variabile aleatoria continua  $X$  è

$$E[X] = \int_{-\infty}^{\infty} s f_X(s) ds$$

*Osservazione:* Il valore atteso di  $X \sim U\{[a, b]\}$  è  $\frac{a+b}{2}$ .

**Proprietà 5.1 ( *Valore Atteso V.A. Continua*)**

1.  $E$ ’ sempre lineare

$$E[aX + bY] = aE[X] + bE[Y]$$

2. Siano  $X, Y$  indipendenti, allora

$$E[XY] = E[X]E[Y]$$

3. Sia  $\Phi : \mathbb{R} \rightarrow \mathbb{R}$  una trasformazione allora

$$E[\Phi(X)] = \int_{-\infty}^{\infty} \Phi(s) f_X(s) ds$$

**Definizione 5.8 ( *Varianza V.A. Continua*)**

Sia  $X$  una variabile aleatoria continua, allora

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

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

*Osservazione:* La varianza dipende dall’ampiezza dell’intervallo. *Osservazione:* La varianza di  $X \sim U\{[a, b]\}$  è  $(b-a)^2/12$ .

**Proprietà 5.2 (Varianza V.A. Continua)**

1.  $\text{Var}(aX) = a^2\text{Var}(X)$ .
2. Siano  $X, Y$  indipendenti, allora  $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$ .

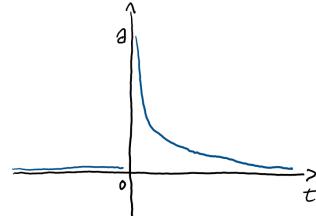
**Definizione 5.9 (Variabile esponenziale)**

Sia  $a > 0$ , allora  $X \sim \text{Exp}(a)$  si definisce

$$f(s) = \begin{cases} ae^{-as} & \text{se } s > 0 \\ 0 & \text{se } s < 0 \end{cases}$$

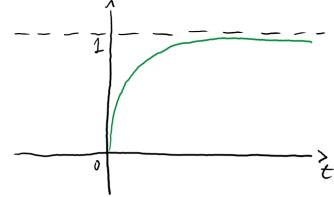
$f(s)$  è una densità continua astratta.

$$f(s) \geq 0, \forall s$$



La funzione di ripartizione è

$$\begin{aligned} F_X(t) &= \int_{-\infty}^t f(s) ds \\ &= \begin{cases} 0 & \text{se } t \leq 0 \\ 1 - e^{-at} & \text{se } t > 0 \end{cases} \end{aligned}$$



Le variabili esponenziali hanno **mancanza di memoria**,  $P(X \geq t+s | X \geq t) = P(X \geq s)$ . Sapendo che  $P(X \geq t) = 1 - P(X \leq t) = 1 - (1 - e^{-at}) = e^{-at}$  per ( $t > 0$ ). Dimostrazione:

$$\begin{aligned} P(X \geq t+s | X \geq t) &= \frac{P(X \geq t+s, X \geq t)}{P(X \geq t)} \\ &= \frac{e^{-a(t+s)}}{e^{-at}} = e^{-as} = P(X \geq s) \end{aligned}$$

Questa proprietà comporta che le variabili esponenziali modellizzano il tempo d'attesa di un fenomeno naturale. Il valore atteso è  $E[X] = \frac{1}{a}$ . La varianza è  $\text{Var}(X) = \frac{1}{a^2}$ .

**Definizione 5.10 (*Trasformazioni v.a. Continue*)**

Sia  $\Phi : \mathbb{R} \rightarrow \mathbb{R}$  una trasformazione e  $X$  una variabile aleatoria continua, allora  $\Phi(X)$  è la trasformazione della variabile aleatoria continua  $X$ . La funzione di ripartizione sarà

$$F_{\Phi(X)}(t) = P(\Phi(X) \leq t)$$

Negli esercizi si risolve la diseguaglianza e si esprime la  $F_{\Phi(X)}$  in  $F_X$ .

**Definizione 5.11 (*Variabili normali (Gaussiane)*)**

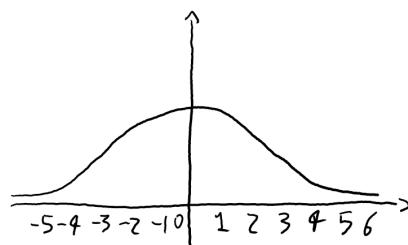
Siano l'integrale di Gauss e la rispettiva densità astratta. Non esiste una formula per  $F(s)$  (funzione di ripartizione), quindi per  $0 < t < 4$  si usano valori noti detti  $\phi(t)$ .

$$\int_{-\infty}^{\infty} e^{-\frac{s^2}{2}} ds = \sqrt{2\pi}$$

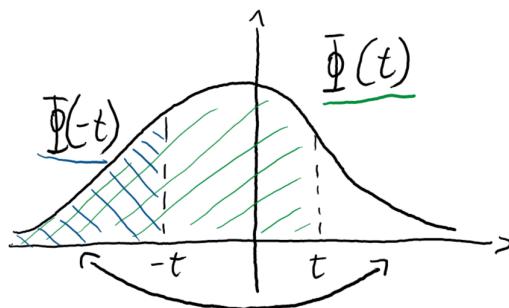
$$f(s) = \frac{1}{\sqrt{2\pi}} e^{-\frac{s^2}{2}}$$

Una variabile con questa densità la chiamiamo normale standard e la indichiamo  $\zeta_0$ . Scriviamo  $\zeta_0 \sim N(0, 1)$ .

Densità di  $\zeta_0$ :

**Somma opposti**

Chiamiamo  $\Phi(t) = F_{\zeta_0}(t)$ . Siano  $\Phi(t) \cong 1$  se  $t > 4$  e  $\Phi(t) \cong 0$  se  $t < -4$ , allora  $\Phi(t) + \Phi(-t) = 1$ .



## Valore atteso e Varianza Variabile Normale

- Per simmetria della densità si ha valore atteso 0.  $E[\zeta_0] = 0$ .
- La varianza vale 1 dato che il valore atteso del quadrato vale 1. Ossia  $\text{Var}(\zeta_0) = 1$ .

### Definizione 5.12 (*Variabili Normali non Standard*)

Dati  $\mu$  e  $\sigma > 0$  la variabile  $\zeta \sim N(\mu, \sigma^2)$  è data da  $\zeta = \mu + \sigma\zeta_0$ .  $E[\zeta] = \mu$ ,  $\text{Var}(\zeta) = \sigma^2$ .

In pratica  $\mu > 0$  slide a destra,  $\mu < 0$  slide a sinistra,  $\sigma > 1$  funzione "spaciata",  $\sigma < 0$  funzione "compatta".

**Standardizzazione:** Dato  $\zeta \sim N(a, b)$  si ottiene  $\zeta = a + \sqrt{b}\zeta_0$  di cui si calcola la probabilità  $P(\zeta \geq c) = P(\zeta_0 > \frac{c-a}{\sqrt{b}}) = 1 - \Phi(\frac{c-a}{\sqrt{b}})$ . Scopo? Le variabili normali sono quelle che si presentano solitamente quando abbiamo a che fare con misurazioni.

### Teorema 5.1 (*Teorema Centrale del limite*)

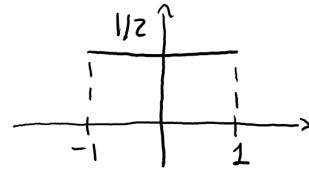
Siano  $X_1, \dots, X_n$  variabili aleatorie (discrete o continue) aventi la stessa funzione di ripartizione di valore atteso  $\mu$  e varianza  $\sigma^2$  e indipendenti. Allora  $X_1 + \dots + X_n$  è approssimativamente una variabile normale  $N(n\mu, n\sigma^2)$ . Se  $n$  è abbastanza grande.

In generale considerando  $\bar{X}_n = \frac{X_1 + \dots + X_n}{n}$  si ha  $E[\bar{X}_n] = \mu$  e  $\text{Var}(\bar{X}_n) = \frac{\sigma^2}{n}$ . Ottenendo così  $\bar{X}_n \sim N(\mu, \frac{\sigma^2}{n})$ .

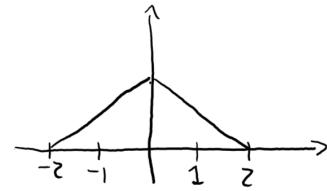
**Osservazione:** Si fa  $S_n \sim N\{n\mu, n\sigma^2\}$  invece di  $\bar{X}_n$  quando voglio la somma delle  $X$  e non la media.

**Dimostrazione Teorema 5.1 :** Tramite un esempio pratico.

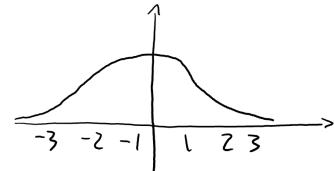
Consideriamo  $X \sim U([-1, 1])$  di densità.



Siano ora  $X_1, X_2 \sim U([-1, 1])$  indipendenti, allora  $X_1 + X_2$  assume valori tra  $[-2, 2]$ .



Analogamente con  $X_1, X_2, X_3 \sim U([-1, 1])$  indipendenti, allora  $X_1 + X_2 + X_3$  assume valori tra  $[-3, 3]$ . (3 rami di parabola).



E così facendo... ■

Dato  $c$  con  $0 < c < 1$ . La densità normale standard è data da

$$\begin{aligned} P(-Z_c < \zeta_0 < Z_c) &= \Phi(Z_c) - \Phi(-Z_c) \\ &= \Phi(Z_c) - (1 - \Phi(Z_c)) \\ &= 2\Phi(Z_c) - 1 - c \\ &= \Phi(Z_c) = \frac{c+1}{2} \end{aligned}$$

Sostituendo  $\zeta_0$  con  $\frac{\bar{X}_n - \mu}{\delta/\sqrt{n}}$  abbiamo  $P(-Z_c < \frac{\bar{X}_n - \mu}{\delta/\sqrt{n}} < Z_c) = c$ .

**Definizione 5.13 (*Correzione di Continuità*)**

Se le variabili  $X_1, \dots, X_n$  assumono valori interi, anche la loro somma darà un valore intero. In questo caso si usa la correzione di continuità in cui un numero intero  $n$  viene pensato come l'intervallo

$$\left[ n - \frac{1}{2}, n + \frac{1}{2} \right]$$

## 6 Statistica Inferenziale

”Scopo di dedurre proprietà di una popolazione dalla conoscenza dei dati di un campione.”

In generale in un’indagine statistica non conosciamo  $\mu$  o  $\sigma$  e quindi il nostro obiettivo è di stimare, in particolare  $\mu$ .

$$\begin{aligned} P\left(-Z_c \frac{\sigma}{\sqrt{n}} < \bar{X}_n - \mu < Z_c \frac{\sigma}{\sqrt{n}}\right) &= c \\ P\left(-\bar{X}_n - Z_c \frac{\sigma}{\sqrt{n}} < -\mu < -\bar{X}_n + Z_c \frac{\sigma}{\sqrt{n}}\right) &= c \\ P\left(\bar{X}_n - Z_c \frac{\sigma}{\sqrt{n}} < -\mu < \bar{X}_n + Z_c \frac{\sigma}{\sqrt{n}}\right) &= c \end{aligned}$$

$\bar{X}_n$  è la media ottenuta sul campione e la conosciamo.  $\sigma$  è la varianza della popolazione e non la conosciamo. Sostituendo  $\sigma^2$  con  $\bar{\sigma}_n$ , la varianza del campione, otteniamo

$$\text{Conf}\left(\bar{X}_n - Z_c \frac{\bar{\sigma}_n}{\sqrt{n}} < \mu < \bar{X}_n + Z_c \frac{\bar{\sigma}_n}{\sqrt{n}}\right) = c$$

### 6.1 Significatività d'un test

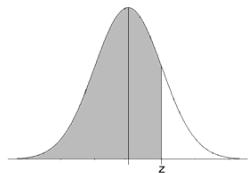
Un test vuole stabilire se un prodotto è ”buono” mediante una certa proprietà.

- Decido il criterio di valutazione.
- Definisco  $n$  il numero di tentativi da eseguire.
- Effettuo il test.

Il test è positivo se rientra nell’intervallo stimato, o negativo altrimenti, con una certa significatività (possibilità di scartare un buono sbagliando).

## 7 Curva normale standardizzata

## Tavola delle probabilità della curva normale standardizzata



### Probabilità cumulate per valori POSITIVI di z