Leonardo Sasso
Enrico Zoli

Colori della Matematica

GeoGebra

COMPETENZE

DIDATTICA INCLUSIVA

EDIZIONE VERDE

STATISTICA E CALCOLO DELLE PROBABILITÀ

DeA SCUOLA | Petrini

Leonardo Sasso
Enrico Zoli

Colori della Matematica

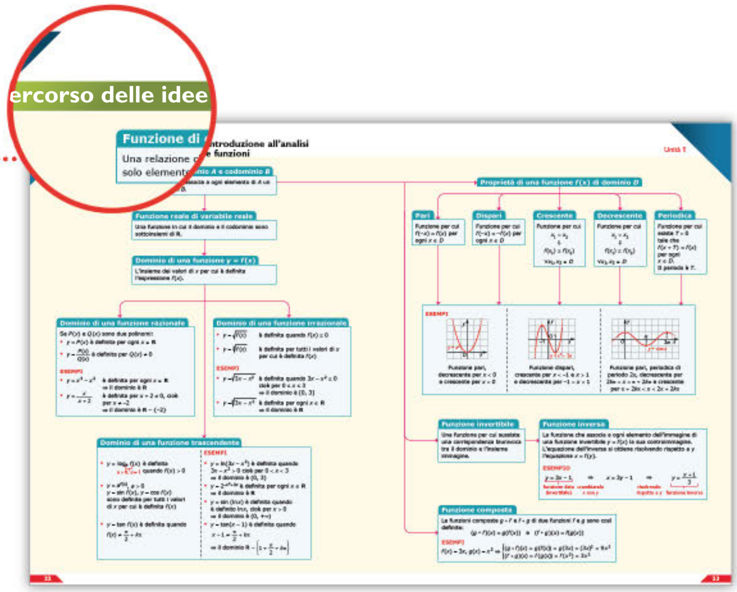
EDIZIONE VERDE

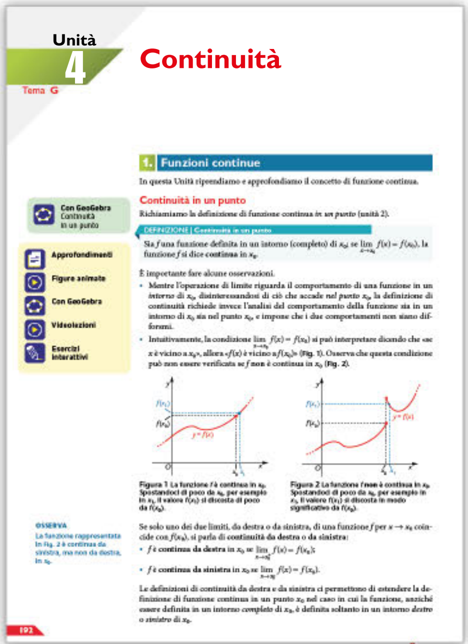
STATISTICA E CALCOLO DELLE PROBABILITÀ

DeA SCUOLA
Petrini

Colori della Matematica

Modellizzazione, visualizzazione, inclusività



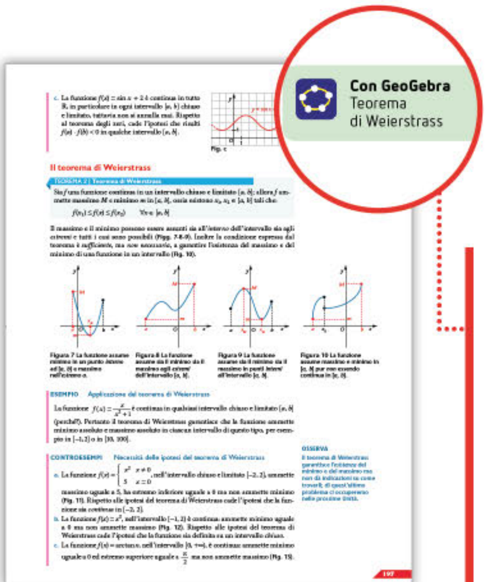


Didattica inclusiva e videolezioni

I **Percorsi delle idee**, realizzati in font ad **alta leggibilità**, presentano in modo visuale i **concetti fondamentali** di ciascuna unità.
Le **videolezioni** hanno una efficace funzione di tutoraggio.

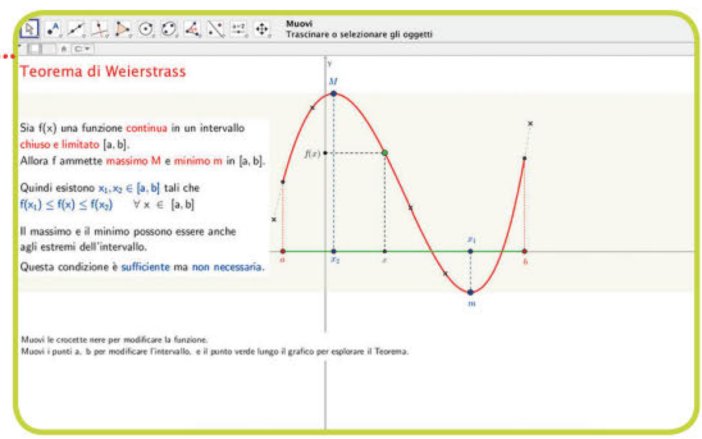
Visualizzazioni e metodo di studio

La teoria si avvale di schemi e tabelle per **visualizzare** e collegare, applica il **problem solving** e la **modellizzazione** per comprendere, propone strumenti per acquisire un corretto **metodo di studio** e per tracciare costanti **collegamenti con la realtà**.



GeoGebra

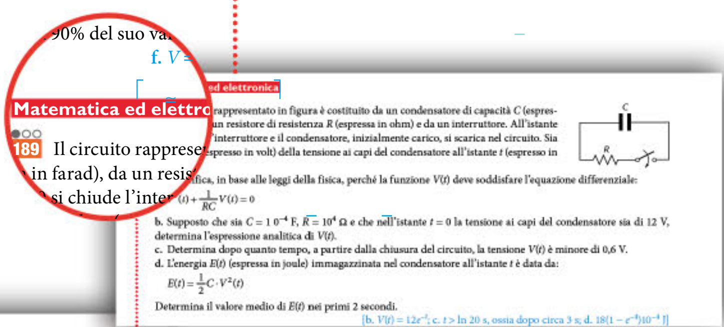
La **visualizzazione** e la comprensione dei concetti è facilitata dalla presenza costante di **animazioni in GeoGebra** di figure geometriche e grafici di funzione.

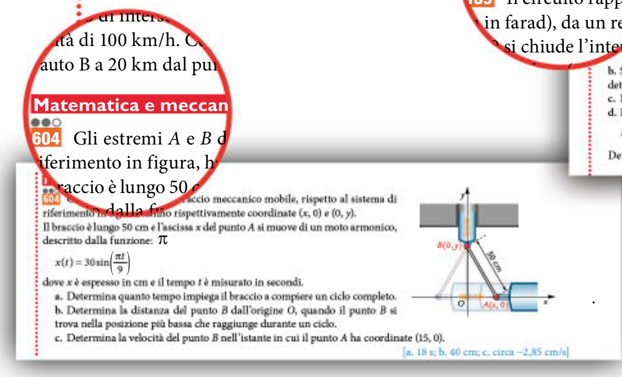


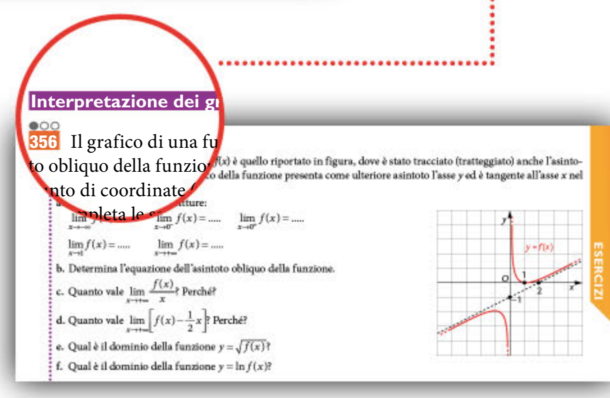
Didattica per competenze

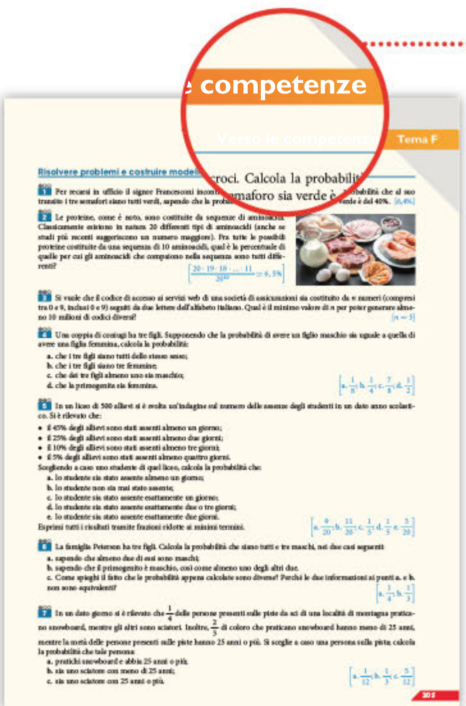
Esercizi motivanti

Numerose le **tipologie di esercizi**: “realtà e modelli”, “a mente”, “interpretazione di grafici”, “matematica e…”, “collegamenti”, “giustificare e argomentare”, “zoom sull’enunciato”, “dal giornale”, “dalle gare”… Per ogni esercizio è indicato il **grado di difficoltà**. Molte attività sono orientate ad **applicazioni specifiche** in ambito tecnico.









Competenze e Invalsi

Le **schede di fine tema** allenano all’**acquisizione delle competenze** e alla prova **Invalsi**.

Compiti di realtà

I **Compiti di realtà** consolidano le competenze in contesti diversi, perlopiù in **ambito lavorativo**.

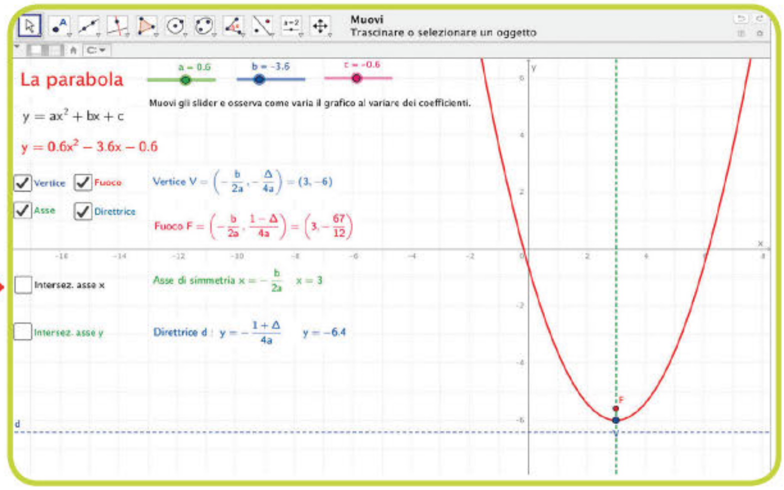
L'ambiente educativo digitale di Colori della Matematica

Le risorse digitali dell'eBook

In tutte le pagine dell'eBook si trovano icone che segnalano la presenza di una o più risorse digitali.
Queste risorse espandono il libro di testo e sono di diverse tipologie, ciascuna caratterizzata da un'icona.

Con GeoGebra

Costruzioni e animazioni in GeoGebra per esplorare dinamicamente grafici di funzione e figure geometriche.



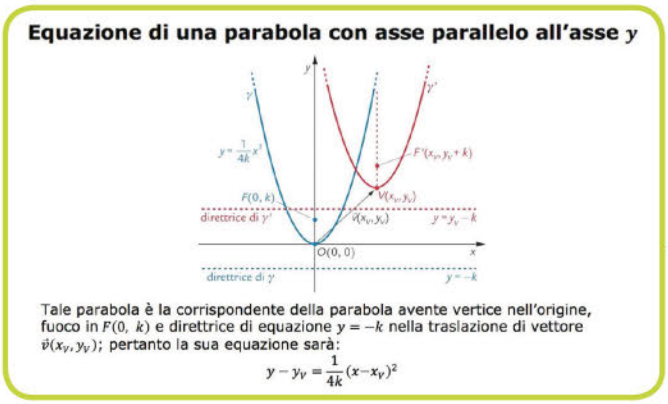
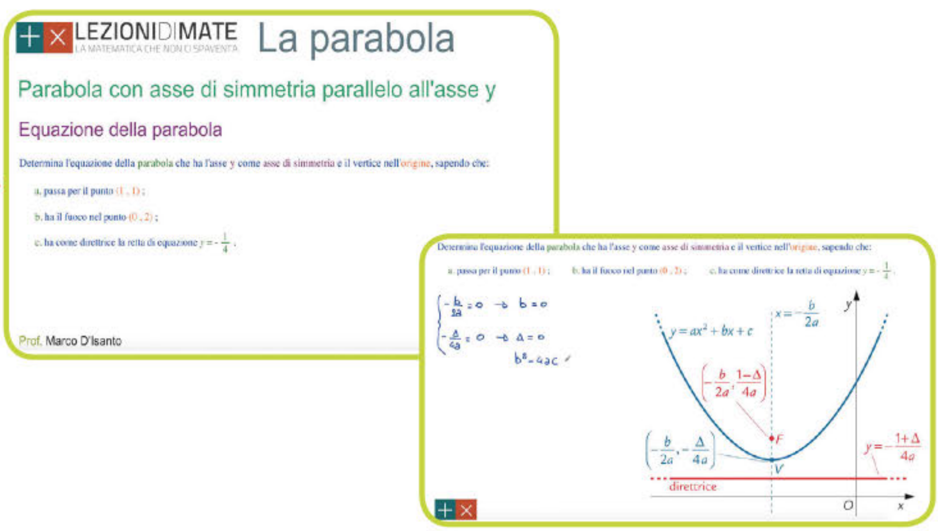


Figure animate

Brevi **filmati** in cui viene rappresentata una proprietà matematica o una figura geometrica.

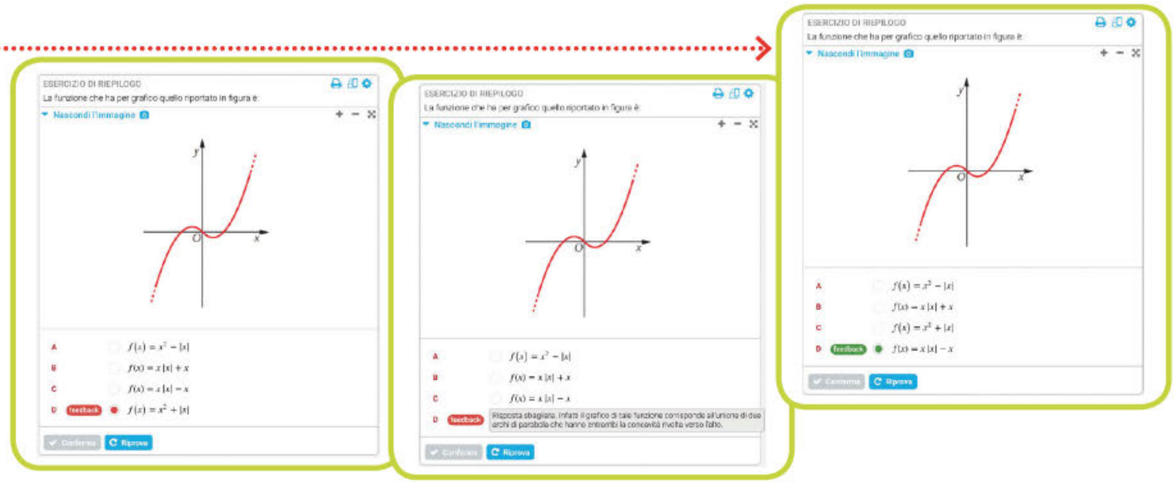
Videolezioni

Svolgimenti commentati di esercizi-modello con richiami alle regole teoriche da applicare.



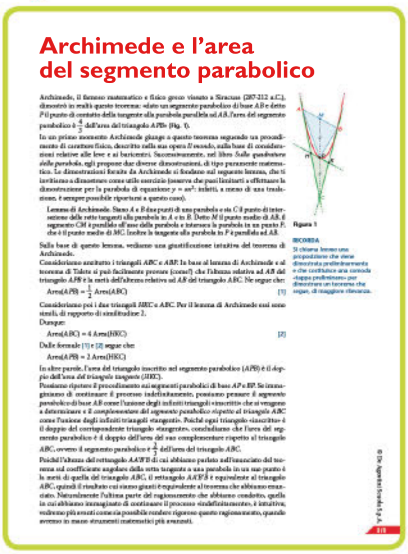
Esercizi interattivi

Esercizi autocorrettivi con feedback delle risposte sbagliate e guida alla risoluzione corretta.



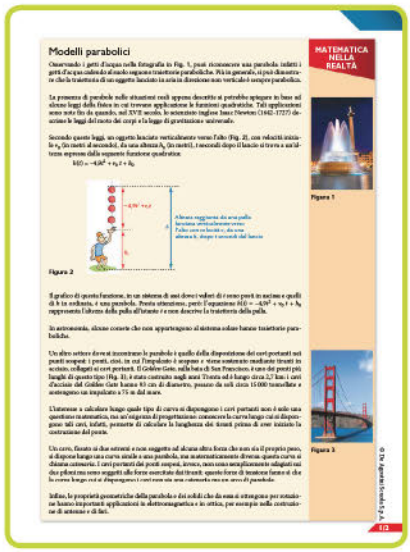
Approfondimenti

Schede in pdf (scaricabili e stampabili) che integrano e approfondiscono i concetti trattati nel libro.



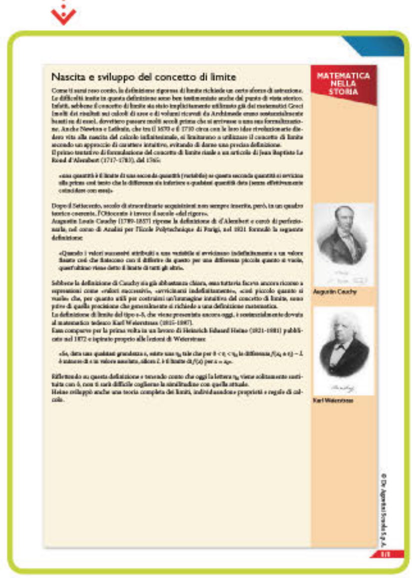
Matematica nella realtà

Schede in pdf (scaricabili e stampabili) di collegamento con la realtà.



Matematica nella storia

Schede in pdf (scaricabili e stampabili) dedicate alla storia della matematica.



In più

SITO DI PRODOTTO

SITO SASSO VERDE

Tutte le risorse digitali del libro, organizzate per indice e tipologia, sono disponibili all'indirizzo **sassoverde.deascuola.it**.

PERCORSO INVALSI

PERCORSO INVALSI

Ambiente dedicato alla preparazione della Prova INVALSI di Matematica con palestre e simulazioni.

INDICE

Unità

4 Calcolo delle probabilità 138

Approfondimenti

Con GeoGebra

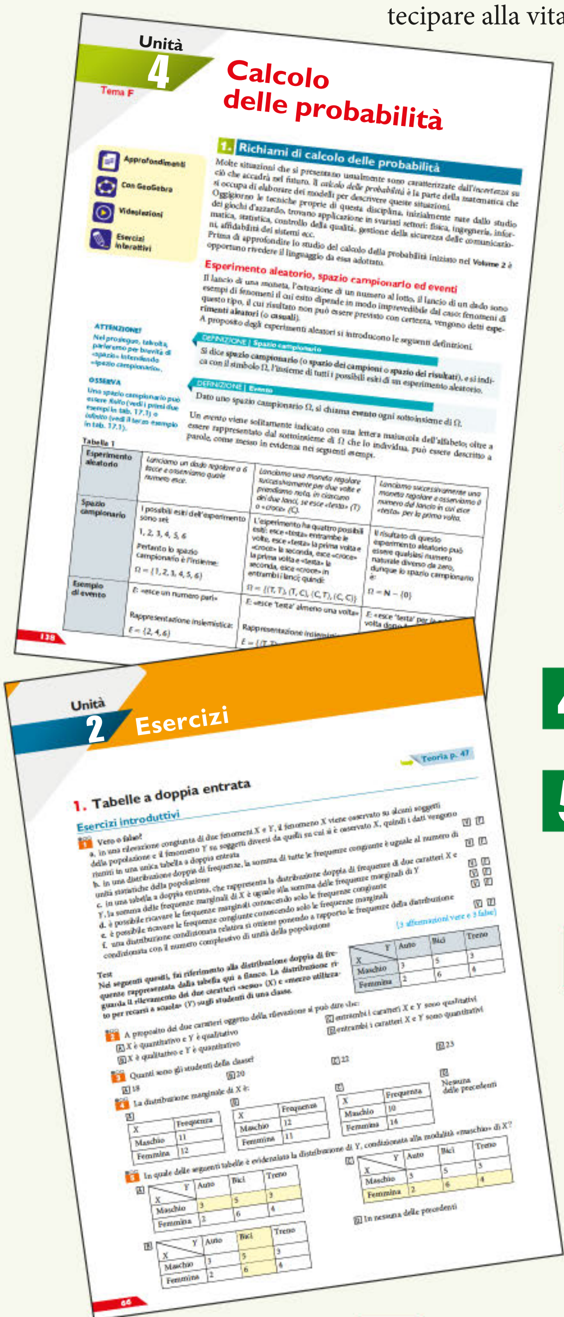
Videolezioni

Esercizi interattivi

Come studiare la Matematica e utilizzare questo libro

Lo studio della Matematica e questo libro in particolare hanno diversi scopi:

- continuare lo sviluppo delle competenze matematiche che hai acquisito nei corsi precedenti;
- farti scoprire alcune applicazioni della matematica nel mondo in cui viviamo;
- contribuire a farti acquisire quegli strumenti scientifici sempre più essenziali per partecipare alla vita sociale con consapevolezza e capacità critica.



Per raggiungere questi scopi, ti diamo qualche consiglio su come studiare matematica.

1 Lo studio della matematica, come hai già avuto modo di constatare, richiede **impegno e partecipazione**. Non puoi imparare molto limitandoti ad assistere alle lezioni: devi partecipare, porti domande e confrontarti, anche da solo, con problemi ed esercizi.

2 È importante che studi matematica **con regolarità**: potrai così assimilare più agevolmente i concetti e il tuo insegnante potrà aiutarti più facilmente a superare le difficoltà.

3 Dovresti **leggere** le lezioni di questo libro e cercare di capire ciò che hai letto. A questo proposito ti diamo alcuni suggerimenti:

- leggi lentamente, prestando attenzione a ogni parola e ai simboli;
- rileggi le parti che non ti risultano chiare;
- prova a rifare da solo gli **esempi** che compaiono svolti nel testo.

4 Risolvi gli esercizi che trovi al termine di ciascuna Unità, suddivisi in paragrafi, con l'aiuto degli **esercizi svolti** e **guidati**.

5 Alla fine di ogni **tema** trovi una serie di esercizi sulle **competenze** che devi acquisire relativamente agli argomenti trattati nel tema stesso; cerca di risolvere anche gli esercizi **verso la prova INVALSI**, simili a quelli che dovrai affrontare in questa prova.

6 Sfrutta le **risorse digitali** disponibili nell'**eBook** e nel **sito** dedicato al libro: potrai trovare costruzioni e animazioni in GeoGebra per affrontare dinamicamente i principali argomenti del libro, figure animate, esercizi interattivi autocorrettivi, videolezioni sulla risoluzione degli esercizi, approfondimenti e schede dedicate alla storia della matematica e all'applicazione della matematica nella vita quotidiana.

7 Quando risolvi un problema, non limitarti a scrivere la tua soluzione: sforzati di **illustrare ciò che stai facendo** e di **giustificare i vari passaggi**, con spiegazioni sintetiche ma esaurienti.

8 Se non riesci a rispondere a una domanda o a risolvere un esercizio immediatamente, non preoccuparti! Rileggi la lezione e gli esempi. Se puoi, abbandona momentaneamente la questione e affrontala in un secondo tempo. Quando qualcosa non ti è chiaro, **poni domande** e parlane con altri.

9 Cerca di studiare con **spirito critico**: la matematica non è solo calcolo, ma è soprattutto una *forma di pensiero*. Nell'epoca di innovazioni tecnologiche in cui viviamo, questo secondo aspetto è sempre più essenziale: spesso i calcoli si possono demandare alle macchine, mentre è essenziale saper ragionare in modo corretto, risolvere e porsi problemi, unire fantasia e razionalità.

TEMA
E

Statistica

La statistica è una disciplina all'ordine del giorno. Basta sfogliare un giornale per trovare dati statistici sui prezzi dei prodotti in commercio, sul gradimento dei programmi televisivi e anche su risultati politici o economici. In medicina, questa disciplina è uno strumento di fondamentale importanza per classificare e analizzare la diffusione delle patologie e i loro legami con i fattori che le determinano. Inoltre, tutti i settori della scienza impiegano i metodi della statistica per ordinare e analizzare i dati numerici ottenuti negli esperimenti.

Studiando i metodi di presentazione dei dati della statistica scopriremo come, mediante opportuni grafici, sia possibile rendere intuitivi dati che altrimenti non sarebbero altro che una sterile successione di numeri. Però proprio su questo punto bisogna prestare attenzione: infatti, in alcuni casi chi trae delle conclusioni con grafici o percentuali lo fa in maniera non corretta, scegliendo la rappresentazione dei dati in base al risultato che vuole suggerire. Studiando le Unità di questo Tema capiremo perciò quanto sia importante saper leggere i dati con metodi corretti, senza dover dipendere dall'interpretazione di altri.

Unità 1
Richiami e complementi di statistica in una variabile

Unità 2
Statistica bivariata, correlazione e regressione

PREREQUISITI

- Proprietà dei numeri reali
- Rapporti e percentuali
- Proprietà dei radicali
- Approssimazioni e stime

COMPETENZE

- Utilizzare il linguaggio e i metodi propri della matematica per organizzare e valutare adeguatamente informazioni qualitative e quantitative

Unità 1

Tema E

Richiami e complementi di statistica in una variabile

- Approfondimenti
- Matematica nella storia
- Con GeoGebra
- Videolezioni
- Esercizi interattivi

1. Richiami di statistica descrittiva

Il primo passo per riprendere lo studio della statistica iniziato nel **Volume 1** è conoscere il significato di alcuni termini specifici.

DEFINIZIONE | Popolazione e unità statistica

L'insieme degli individui oggetto di un'indagine statistica si chiama **popolazione** (o **universo** o **collettivo**); ciascun individuo facente parte della popolazione viene chiamato anche **unità statistica**.

Alcune indagini statistiche consentono di interpellare *tutti* i membri della popolazione; in altri casi, per motivi di costi e di tempi, bisogna limitarsi a sottoporre le domande o le richieste di informazioni solo a *una parte* della popolazione, che viene chiamata **campione**.

DEFINIZIONE | Carattere

Si chiama **carattere** la proprietà che è oggetto di studio in un'indagine statistica.

Matematica nella storia
La nascita e gli sviluppi della statistica

Per esempio, il *peso* di una persona può essere un possibile carattere di un'indagine. In corrispondenza di ogni individuo della popolazione, il carattere oggetto di studio assume una determinata *modalità*; per esempio, il carattere «peso» può assumere in corrispondenza di un dato individuo la modalità 72 kg, in corrispondenza di un altro la modalità 80 kg e così via.

DEFINIZIONE | Modalità

Si chiama **modalità** ciascuna delle varianti con cui un carattere può presentarsi; le modalità osservate si chiamano **dati.**

I caratteri si classificano secondo le seguenti definizioni.

MODI DI DIRE
«Variabile» e «mutabile» sono, in questo contesto, sostantivi. Si usa talvolta parlare di **valori** di una variabile, anziché di **modalità** di una variabile.

DEFINIZIONI | Caratteri quantitativi e caratteri qualitativi

Un carattere le cui modalità sono espresse da numeri è detto **carattere quantitativo** (o **variabile**), un carattere le cui modalità **non** sono espresse da numeri è detto **carattere qualitativo** (o **mutabile**).

ESEMPI

Caratteri quantitativi	Caratteri qualitativi
• I giorni di assenza di uno studente in un anno scolastico • La quantità, in kilogrammi, di mele vendute da un negozio in un giorno • Il numero di multe dato in un giorno dai vigili di una città	• Il gusto di gelato preferito • Il tipo di alimentazione (benzina, diesel o elettrica) delle macchine che ci sono in un parcheggio • Il colore delle automobili vendute da un concessionario in una settimana

Approfondimento
Le fasi di un'indagine statistica

Le *variabili* (ovvero i caratteri *quantitativi*) si classificano ulteriormente a seconda del tipo di valori che possono assumere.

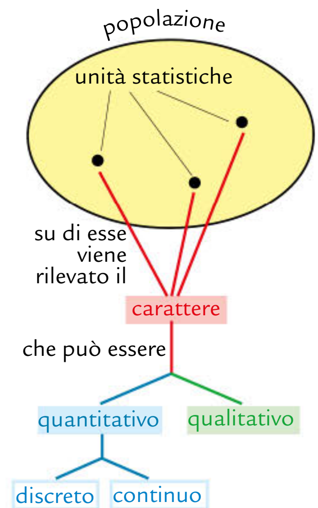
DEFINIZIONI | **Variabili discrete e variabili continue**

Una variabile si dice **discreta** quando le sue modalità sono esprimibili mediante numeri interi; si dice **continua** quando può assumere (almeno teoricamente) tutti i valori reali di un determinato intervallo.

Tipicamente, le variabili discrete sono quelle che si rilevano *contando* (per esempio il numero dei figli in una famiglia o il numero dei dipendenti di un'azienda), mentre le variabili continue sono quelle che si rilevano mediante *misurazioni* (per esempio il peso di un bambino a una certa età o la temperatura massima giornaliera registrata a Milano in un dato giorno).

ESEMPI RIASSUNTIVI

Fenomeno studiato	**Popolazione**	**Carattere**	**Modalità**	**Tipo di carattere**
Il colore degli occhi degli italiani	Tutti gli italiani	Il colore degli occhi	Verdi, azzurri, marroni ecc.	Qualitativo
Altezza (misurata in metri) degli studenti di una classe	Gli studenti della classe	La misura dell'altezza	1,72 m; 1,85 m; 1,78 m	Quantitativo continuo
L'anno di nascita degli iscritti a una palestra	Tutti gli iscritti alla palestra	L'anno di nascita	..., 1970, ..., 1981, 1965, ...	Quantitativo discreto



Distribuzioni di frequenze

Ricordiamo anzitutto alcune definizioni.

Termine	**Significato**
Frequenza (assoluta) di una modalità	Il numero di volte in cui una modalità è stata osservata
Frequenza relativa di una modalità	Il rapporto tra la frequenza assoluta della modalità e il numero di individui della popolazione
Frequenza percentuale di una modalità	La rappresentazione in percentuale della frequenza relativa
Frequenza cumulata di una modalità di un carattere quantitativo	La somma delle frequenze di tutte le modalità minori o uguali a quella considerata

Una prima forma di elaborazione dei dati, volta a ottenere una maggiore *sintesi*, consiste nel costruire una tabella in cui riportare, per ciascuna delle *modalità* differenti $x_1, x_2, ..., x_k$ osservate, la rispettiva frequenza assoluta; questa tabella, in cui si aggiunge di solito un'ultima riga in cui si riporta il totale n delle frequenze (**Tab. 1**), è detta **distribuzione di frequenze** del carattere esaminato. Il numero n rappresenta il numero complessivo di unità della popolazione. Nella costruzione della tabella bisogna ricordare, se il carattere è quantitativo, di ordinare le modalità in senso crescente.
In modo del tutto analogo si possono costruire le **distribuzioni di frequenze relative o percentuali**, sostituendo le frequenze assolute con quelle relative o percentuali.

Tabella 1

X	**Frequenze**
x_1	f_1
x_2	f_2
.....	.....
x_k	f_k
Totale	n

ESEMPIO Distribuzione di frequenze

Supponiamo di avere rilevato, in una classe di una scuola, il colore degli occhi degli allievi. Il risultato della rilevazione fornisce i cosiddetti **dati grezzi**, che abbiamo raccolto nella seguente tabella in cui a ogni *unità statistica* (cioè a ogni studente) è stata associata la *modalità* del carattere osservata, cioè il colore nero (N), marrone (M), azzurro (A) o verde (V).

Studente	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Colore degli occhi	N	M	A	N	N	M	A	V	V	M	M	N	N	V	V	N	M	N

Associando a ogni modalità la sua *frequenza* assoluta, la tabella dei dati grezzi si sintetizza nella seguente, che rappresenta la distribuzione di frequenze del carattere esaminato.

Colore degli occhi	Numero di studenti
Nero	7
Marrone	5
Azzurro	2
Verde	4
Totale	18

In alcuni casi, prima di costruire la tabella che rappresenta la distribuzione di frequenze, è utile *accorpare* le modalità in *intervalli* tra loro *disgiunti*, detti **classi**.

ESEMPIO Distribuzione di frequenze suddivisa per classi

Nella stessa classe in cui prima abbiamo rilevato il *colore degli occhi* degli studenti, rileviamo ora la loro *statura*. Il risultato della rilevazione fornisce i dati grezzi riassunti nella seguente tabella.

Studente	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Altezza (in cm)	173	164	174	180	182	176	184	185	170	172	186	167	188	183	168	176	184	178

Suddividendo le possibili altezze degli studenti (misurate in centimetri) nei seguenti intervalli:

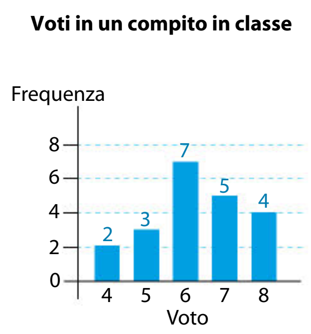
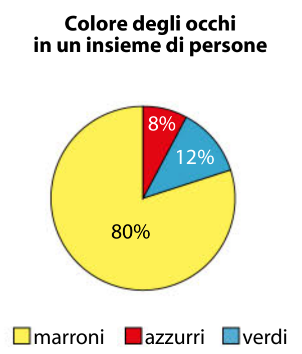
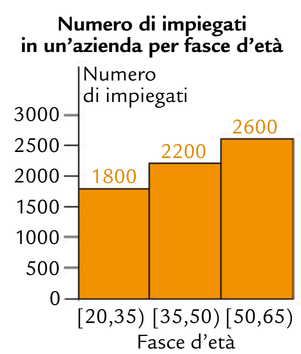
(160, 165]; (165, 170]; (170, 175]; (175, 180]; (180, 185]; (185, 190]

otteniamo la distribuzione di frequenze rappresentata nella tabella seguente.

Altezza (in cm)	Frequenza
(160, 165]	1
(165, 170]	3
(170, 175]	3
(175, 180]	4
(180, 185]	5
(185, 190]	2

Principali rappresentazioni grafiche

Esiste una grande varietà di grafici utilizzati in statistica, ma i più importanti si possono ricondurre alle quattro categorie di cui puoi vedere alcuni esempi nella tabella seguente: diagrammi a barre, diagrammi circolari, diagrammi cartesiani e istogrammi.

Diagramma a barre (rettangoli distanziati)	Diagramma circolare	Diagramma cartesiano	Istogramma (rettangoli affiancati)
			

Esercizi p. 24

2. Richiami sugli indici di posizione e di variabilità

Con GeoGebra
Indici di posizione

Il caso in cui è data una distribuzione di dati grezzi

Ci riferiamo a un carattere *quantitativo* X, di cui sono stati osservati i valori $x_1, x_2, ..., x_n$.

Termine	Definizione	Esempio
Media aritmetica (indicata con $\overline{x}$ o con μ)	$\mu = \overline{x} = \dfrac{x_1 + x_2 + ... + x_n}{n}$	La media dei tre numeri 2, 4 e 6 è: $\overline{x} = \dfrac{2+4+6}{3} = 4$
Mediana	Ordinati i numeri $x_1, x_2, ..., x_n$ in senso crescente (o decrescente) la loro mediana è: – il numero che occupa la posizione centrale, se n è *dispari*; – la media aritmetica dei due numeri che occupano le posizioni centrali, se n è *pari*.	La mediana dei tre numeri: 4, 5, 6 **$n = 3$ (dispari)** è il numero 5. La mediana dei quattro numeri: 4, 5, 6, 7 **$n = 4$ (pari)** è la media aritmetica dei due numeri che occupano le posizioni centrali, quindi è $\dfrac{5+6}{2} = \dfrac{11}{2}$
Moda	Il dato (o i dati) che hanno la massima frequenza.	Dati i numeri 1, 2, 3, 4, 3 la moda è il numero 3, che compare con frequenza massima, uguale a 2.
Varianza (indicata con V o con σ^2)	$\sigma^2 = V = \dfrac{(x_1 - \overline{x})^2 + ... + (x_n - \overline{x})^2}{n}$ oppure $\sigma^2 = V = \dfrac{x_1^2 + ... + x_n^2}{n} - \overline{x}^2$	Consideriamo i due numeri 4 e 8. La loro media aritmetica è: $\overline{x} = \dfrac{4+8}{2} = 6$ La loro varianza, in base alla seconda formula, è: $\sigma^2 = V = \dfrac{4^2 + 8^2}{2} - 6^2 = 4$
Deviazione standard (o **Scarto quadratico medio**)	$\sigma = \sqrt{V}$	La deviazione standard di 4 e 8, in base alla varianza poc'anzi calcolata, è $\sigma = \sqrt{4} = 2$.

La media, la moda e la mediana sono detti indici di *posizione*, mentre la varianza e la deviazione standard sono detti indici di *variabilità*. Gli indici di posizione e di variabilità consentono di sintetizzare in pochi numeri significativi le principali caratteristiche del fenomeno indagato. In particolare, gli indici di *variabilità* forniscono informazioni sull'attitudine di un fenomeno a manifestarsi sulle varie unità statistiche con modalità diverse e distanti tra loro.

Il caso in cui è data una distribuzione di frequenze

Nel caso che sia assegnata una distribuzione di frequenze, occorre tenere presente che:

- per il calcolo della *media* e della *varianza* si utilizzano le formule seguenti:

$$\overline{x} = \frac{x_1 f_1 + x_2 f_2 + \ldots + x_k f_k}{f_1 + f_2 + \ldots + f_k} \qquad \sigma^2 = \frac{x_1^2 \cdot f_1 + \ldots + x_k^2 \cdot f_k}{f_1 + \ldots + f_k} - \overline{x}^2$$

  dove x_1, x_2, ..., x_k sono i valori osservati, rispettivamente con frequenze f_1, f_2, ..., f_k, del carattere (quantitativo) X in esame; indicando con f_{r_1}, f_{r_2}, ..., f_{r_k} rispettivamente le frequenze relative di x_1, x_2, ..., x_k, le formule precedenti possono essere espresse anche nelle forme equivalenti:

$$\overline{x} = \underbrace{x_1 f_{r_1} + x_2 f_{r_2} + \ldots + x_k f_{r_k}}_{\text{utilizzando le frequenze relative}} \qquad \sigma^2 = \underbrace{x_1^2 \cdot f_{r_1} + \ldots + x_k^2 \cdot f_{r_k}}_{\text{utilizzando le frequenze relative}} - \overline{x}^2$$

- per il calcolo della *mediana* conviene ricorrere al calcolo delle frequenze cumulate, come spiegato nel prossimo esempio.

ESEMPIO Valori medi nel caso di una distribuzione di frequenze

Un'indagine effettuata su un campione di famiglie ha prodotto la distribuzione di frequenze rappresentata nella tabella qui sotto. Determiniamo la media, la mediana, la moda e la deviazione standard della distribuzione.

Numero di figli per famiglia	Frequenza
0	9
1	27
2	40
3	20
4	3
5	1

- Il numero medio di figli per famiglia è dato dalla formula:

$$\overline{x} = \frac{0 \cdot 9 + 1 \cdot 27 + 2 \cdot 40 + 3 \cdot 20 + 4 \cdot 3 + 5 \cdot 1}{9 + 27 + 40 + 20 + 3 + 1} = \frac{184}{100} = 1{,}84$$

- Per il calcolo della mediana determiniamo preliminarmente le frequenze cumulate:

Numero di figli per famiglia	Frequenza	Frequenza cumulata
0	9	9
1	27	$9 + 27 = 36$
2	40	$36 + 40 = 76$
3	20	$76 + 20 = 96$
4	3	$96 + 3 = 99$
5	1	$99 + 1 = 100$

Il numero complessivo di famiglie intervistate è $n = 100$ (ultima frequenza cumulata). Le due famiglie che occupano le posizioni centrali sono la cinquantesima e la cinquantunesima. Dalla colonna delle frequenze cumulate deduciamo che le famiglie dalla numero 37 alla numero 76 hanno 2 figli, quindi in particolare hanno 2 figli le famiglie corrispondenti alle due posizioni centrali. La mediana è, per definizione, la media fra i figli di queste due famiglie, dunque è 2.

- La moda della distribuzione è chiaramente «2», che corrisponde alla massima frequenza (uguale a 40).
- La varianza è data dalla formula:

$$\sigma^2 = \frac{0^2 \cdot 9 + 1^2 \cdot 27 + 2^2 \cdot 40 + 3^2 \cdot 20 + 4^2 \cdot 3 + 5^2 \cdot 1}{9 + 27 + 40 + 20 + 3 + 1} - (1{,}84)^2 = 1{,}0144$$

e la deviazione standard è quindi:

$$\sigma = \sqrt{1{,}0144} \simeq 1{,}007$$

RIFLETTI

1. Il significato della mediana è il seguente: almeno il 50% delle famiglie hanno un numero di figli maggiore o uguale a 2 e almeno il 50% delle famiglie hanno un numero di figli minore o uguale a 2.

2. Nell'esempio qui a fianco la mediana e la moda coincidono. In generale non è detto che ciò avvenga.

Il caso in cui è data una distribuzione suddivisa per classi

Se è data una distribuzione di frequenze **suddivisa per classi**, non è possibile determinare i valori *esatti* di media e mediana (perché non conosciamo esattamente i valori osservati all'interno di ciascuna classe); tuttavia è possibile calcolare delle loro approssimazioni, come mostriamo nel prossimo esempio, *assumendo che tutti i valori osservati all'interno di una classe siano uguali al valore centrale della classe stessa*.
Nel caso di distribuzioni di frequenze suddivise per classi, si definisce inoltre come *classe modale*:

- quella (o quelle) con *frequenza* massima, se le classi hanno la stessa ampiezza;
- quella (o quelle) con la massima *densità di frequenza*, se le classi hanno ampiezza diversa (in questo caso, infatti, bisogna «depurare» la frequenza dall'influenza dovuta all'ampiezza e per farlo si pongono a rapporto frequenza e ampiezza: difatti, più ampio è l'intervallo, più è facile che la frequenza sia elevata per effetto dell'ampiezza dell'intervallo).

RICORDA

Si chiama **densità di frequenza** il rapporto tra la frequenza della classe e la sua ampiezza.

Anche per il calcolo della varianza (e quindi quello della deviazione standard) si sostituisce ogni classe con il corrispondente valore centrale, in modo da ricondursi a una distribuzione di frequenze.

ESEMPIO Valori medi nel caso di una distribuzione suddivisa in classi

Un'indagine effettuata su un campione di individui ha prodotto la distribuzione di frequenze in **Tab. 1**. Determiniamo:

a. la media b. la mediana c. la classe modale d. la deviazione standard

della distribuzione.

Tabella 1

Peso (in kg)	$40 \le p < 50$	$50 \le p < 60$	$60 \le p < 70$	$70 \le p < 80$	$80 \le p < 90$
Frequenza	16	48	45	8	3

a. Assumendo che i dati osservati all'interno di ciascuna classe coincidano con il suo valore centrale, siamo ricondotti al calcolo della media della seguente distribuzione di frequenze in **Tab. 2**:

Tabella 2

Peso (in kg)	45	55	65	75	85
Frequenza	16	48	45	8	3

$\underbrace{\frac{40+50}{2} = 45}_{\text{valore centrale della classe } 40 \le p < 50}$ $\underbrace{\frac{50+60}{2} = 55}_{\text{valore centrale della classe } 50 \le p < 60}$ $\underbrace{\frac{60+70}{2} = 65}_{\text{valore centrale della classe } 60 \le p < 70}$ $\underbrace{\frac{70+80}{2} = 75}_{\text{valore centrale della classe } 70 \le p < 80}$ $\underbrace{\frac{80+90}{2} = 85}_{\text{valore centrale della classe } 80 \le p < 90}$

ATTENZIONE!

Il metodo di calcolo approssimato della mediana per un carattere suddiviso in classi esposto si basa come detto sull'assunzione che tutti i valori osservati (ignoti) che cadono all'interno della classe mediana siano uguali al *valore centrale* della classe. Questa ipotesi non è tuttavia l'unica possibile. Per esempio, un'altra ipotesi è quella della distribuzione *uniforme*: in mancanza di informazioni su dove cadono i valori osservati all'interno di una classe, si suppone che essi siano distribuiti in modo uniforme ed equidistante lungo tutto l'intervallo. Noi abbiamo scelto l'ipotesi del valore centrale (che peraltro è quella adottata dalla maggior parte dei software statistici) per la sua semplicità, ma è bene sapere che, adottando ipotesi diverse, si può giungere a valori della mediana leggermente differenti.

A questo punto il peso medio medio $\overline{p}$ può essere ricavato con una media aritmetica ponderata:

$$\overline{p} = \frac{45 \cdot 16 + 55 \cdot 48 + 65 \cdot 45 + 75 \cdot 8 + 85 \cdot 3}{16 + 48 + 45 + 8 + 3} = \frac{119}{2} \simeq 59{,}5 \text{ (kg)}$$

b. Per individuare la mediana della distribuzione, è utile calcolare le *frequenze cumulate* (**Tab. 3**):

Tabella 3

Peso (in kg)	$40 \leq p < 50$	$50 \leq p < 60$	$60 \leq p < 70$	$70 \leq p < 80$	$80 \leq p < 90$
Frequenza	16	48	45	8	3
Frequenza cumulata	16	64	109	117	120

Il collettivo è composto complessivamente da 120 individui (*pari*); la mediana è data perciò dalla media fra il sessantesimo e il sessantunesimo peso osservato. Dalla riga delle frequenze cumulate si deduce che i pesi osservati dal numero 17 al numero 64 appartengono alla classe $50 \leq p < 60$. Assumiamo pertanto che il sessantesimo e il sessantunesimo peso osservato siano uguali a 55 (valore centrale della classe $50 \leq p < 60$) e concludiamo che il peso mediano è 55 kg.

c. Dal momento che le classi hanno la stessa ampiezza (uguale a 10), la classe modale è quella che ha maggiore frequenza, ossia $50 \leq p < 60$.

d. La varianza è data dalla formula:

$$\sigma^2 = \frac{45^2 \cdot 16 + 55^2 \cdot 48 + 65^2 \cdot 45 + 75^2 \cdot 8 + 85^2 \cdot 3}{16 + 48 + 45 + 8 + 3} - (59{,}5)^2 \simeq 108{,}1$$

e la deviazione standard è quindi $\sigma \simeq \sqrt{108{,}1} \simeq 10{,}4$.

Proprietà degli indici di posizione e di variabilità

Lasciamo a te le dimostrazioni dei prossimi teoremi, che esprimono alcune proprietà della media aritmetica, della varianza e della deviazione standard.

TEOREMA 1 | Proprietà della media aritmetica

a. La media aritmetica $\overline{x}$ di n numeri $x_1, x_2, \ldots, x_n$ è quel valore che, sostituito al posto di ciascuno di essi, lascia invariata la loro somma complessiva:

$$x_1 + x_2 + \ldots + x_n = \underbrace{\overline{x} + \ldots + \overline{x}}_{n \text{ volte}} = n \cdot \overline{x}$$

b. La somma delle differenze tra ciascuno degli n numeri $x_1, x_2, \ldots, x_n$ e la loro media aritmetica $\overline{x}$ è uguale a zero:

$$(x_1 - \overline{x}) + (x_2 - \overline{x}) + \ldots + (x_n - \overline{x}) = 0$$

c. Supponiamo che un insieme X di numeri venga suddiviso in due sottoinsiemi disgiunti A e B, formati rispettivamente da n ed m elementi, tali che $A \cup B = X$. Se gli elementi di A hanno media aritmetica $\overline{a}$ e gli elementi di B hanno media aritmetica $\overline{b}$, allora la media aritmetica $\overline{x}$ degli elementi di X è la media aritmetica ponderata di $\overline{a}$ e $\overline{b}$, con pesi uguali a n ed m:

$$\overline{x} = \frac{\overline{a} \cdot n + \overline{b} \cdot m}{n + m}$$

d. Sia k un numero reale. Se $\overline{x}$ è la media aritmetica dei numeri $x_1, x_2, \ldots, x_n$, allora:

- la media aritmetica dei numeri $x_1 + k, x_2 + k, \ldots, x_n + k$ è $\overline{x} + k$;
- la media aritmetica dei numeri $kx_1, kx_2, \ldots, kx_n$ è $k\overline{x}$.

TEOREMA 2 | Proprietà della varianza e della deviazione standard

a. Sia k un numero reale; se a tutti i numeri $x_1, x_2, ..., x_n$ si aggiunge (o si toglie) k, la varianza e la deviazione standard restano invariate.

b. Sia k un numero reale; se tutti i numeri $x_1, x_2, ..., x_n$ vengono moltiplicati per k, allora:
- la varianza dei nuovi numeri ottenuti risulta moltiplicata per k^2 rispetto alla varianza dei numeri originari;
- la deviazione standard dei nuovi numeri ottenuti risulta moltiplicata per $|k|$ rispetto alla deviazione standard dei numeri originari.

ESEMPI

a. La varianza e la deviazione standard dei numeri 105, 101, 102, 103 sono rispettivamente uguali alla varianza e alla deviazione standard dei numeri 5, 1, 2, 3 (essendo questi ultimi ottenuti sottraendo 100 da ciascuno dei precedenti).

b. Sapendo che la varianza e la deviazione standard dei due numeri 4 e 8 valgono $\sigma^2 = 4$ e $\sigma = 2$, possiamo immediatamente concludere che la varianza e la deviazione standard dei numeri 12 e 24 (ottenuti moltiplicando per 3 i numeri originari) sono:

$$(\sigma')^2 = 3^2 \cdot 4 = 36 \qquad \sigma' = 3 \cdot 2 = 6$$

La deviazione standard è un indice particolarmente importante, perché tramite esso è possibile dedurre informazioni sulla *distribuzione dei dati intorno alla media*. Si può infatti dimostrare che, qualsiasi sia la distribuzione dei dati, detta $\overline{x}$ la loro media e σ la loro deviazione standard:
- almeno il 75% dei dati cade nell'intervallo $(\overline{x} - 2\sigma, \overline{x} + 2\sigma)$;
- almeno l'89% dei dati cade nell'intervallo $(\overline{x} - 3\sigma, \overline{x} + 3\sigma)$.

Puoi quindi comprendere che, quanto più la deviazione standard è piccola, tanto più la maggior parte dei dati è «vicina» alla media. Vedremo nel **Paragrafo 5** che, se la distribuzione delle frequenze dei dati è approssimativamente «a forma di campana», è possibile stabilire relazioni ancora più precise circa la quantità di dati della distribuzione che cadono a una distanza di una, due o tre deviazioni standard dalla media.

Confronto della variabilità

La varianza e la deviazione standard sono indici che dipendono dall'*unità di misura* e dall'*ordine di grandezza* dei dati. Pertanto non avrebbe senso confrontare le deviazioni standard di due fenomeni misurati con unità di misura diverse, né fenomeni misurati con la stessa unità di misura ma tali, per esempio, che gli ordini di grandezza delle misure di un fenomeno siano molto maggiori degli ordini di grandezza delle misure dell'altro.
Per eseguire il confronto fra la variabilità di due fenomeni occorre utilizzare una misura della variabilità «depurata» dall'influenza dell'unità di misura e dell'ordine di grandezza dei dati. Questo obiettivo si può raggiungere costruendo il *rapporto* tra la *deviazione standard* e un *valore* che sintetizzi l'ordine di grandezza delle modalità del fenomeno osservato e che sia espresso nella medesima unità di misura: il valore più noto che soddisfa queste ultime proprietà è la *media aritmetica*. In definitiva, allora, si definisce il seguente indice:

DEFINIZIONE | Coefficiente di variazione

Si chiama **coefficiente di variazione** C_V di un insieme di dati, la cui media aritmetica è $\overline{x}$ e la cui deviazione standard è σ, il rapporto tra la deviazione standard e la media aritmetica stessa:

$$C_V = \frac{\sigma}{\overline{x}} \qquad \text{con } \overline{x} \neq 0$$

Tutte le volte che saremo interessati a *confrontare* la variabilità di due fenomeni, in base a quanto detto sopra, sarà opportuno confrontare non le loro deviazioni standard, bensì i loro **coefficienti di variazione**.

Con GeoGebra
Medie a confronto

APPROFONDIMENTO La media armonica e la media geometrica

La media aritmetica ha l'importante proprietà di mantenere *invariata* la *somma* dei dati, quando venga sostituita a ciascuno di essi. Ci sono però svariati problemi per risolvere i quali la media «giusta» non è quella che conserva la somma, ma qualche altro aspetto (per esempio il prodotto). Discutiamo ora alcuni di questi problemi e, prendendo spunto da essi, definiamo delle nuove medie.

La media armonica

PROBLEMA
Marco, per preparare un esame, deve studiare un testo di 300 pagine. Egli decide di studiare il libro due volte: la prima volta alla velocità di 10 pagine al giorno e la seconda alla velocità di 30 pagine al giorno. Rispettando il programma, il tempo necessario alla preparazione completa sarà di:

$$\underbrace{\frac{300}{10}}_{\text{tempo prima lettura}} + \underbrace{\frac{300}{30}}_{\text{tempo seconda lettura}} = 30 + 10 = \underbrace{40 \text{ giorni}}_{\text{tempo totale}}$$

Prima di iniziare, Marco cambia idea: vuole leggere lo stesso numero di pagine al giorno in entrambe le fasi di studio, finendo sempre in 40 giorni. Quante pagine deve studiare al giorno?

Se calcoliamo la media aritmetica tra 10 e 30 troviamo 20, ma **non** funziona: con 20 pagine al giorno lo studio viene esaurito in 30 giorni, non in 40. Per trovare la soluzione corretta, risolviamo il seguente problema più generale: indicate con v_1 e v_2 le velocità di lettura programmate inizialmente per le due fasi di studio, cerchiamo la velocità v che, applicata a entrambe le fasi, lascia invariato il tempo totale di studio. Trovare v equivale a risolvere l'equazione:

$$\underbrace{\frac{300}{v_1}}_{\text{tempo prima lettura}} + \underbrace{\frac{300}{v_2}}_{\text{tempo seconda lettura}} = \underbrace{\frac{300}{v} + \frac{300}{v}}_{\text{tempo di due letture alla stessa velocità } v}$$

Dividendo i due membri per 300 si ha l'equazione equivalente:

$$\frac{1}{v_1} + \frac{1}{v_2} = \frac{1}{v} + \frac{1}{v} \qquad [1]$$

da cui si ottiene: $v = \dfrac{1}{\frac{1}{2}\left(\frac{1}{v_1} + \frac{1}{v_2}\right)}$ [2]

Sostituendo nella [1] i dati corrispondenti al problema iniziale ($v_1 = 10$ e $v_2 = 30$) si ottiene $v = 15$: la soluzione corretta dunque è 15 pagine al giorno.
Se rifletti sulla [2] puoi renderti conto che la velocità v ha la proprietà di *lasciare invariata la somma dei reciproci* di v_1 e v_2, quando venga sostituita a v_1 e v_2. È questa la proprietà che caratterizza la media *armonica*, che si può definire come segue.

DEFINIZIONE | Media armonica

Dati n numeri $x_1, x_2, ..., x_n$, tutti *diversi da zero*, si definisce loro **media armonica**, m, il reciproco della media aritmetica dei loro reciproci, cioè si pone:

$$m = \frac{1}{\frac{1}{n}\left(\frac{1}{x_1} + \frac{1}{x_2} + ... + \frac{1}{x_n}\right)} \quad \text{ossia} \quad m = \frac{n}{\left(\frac{1}{x_1} + \frac{1}{x_2} + ... + \frac{1}{x_n}\right)}$$

ESEMPIO Applicazione della media armonica

Il percorso di una gara ciclistica è costituito da un circuito ripetuto tre volte. Un corridore percorre i tre giri alle velocità rispettive di 25 km/h, 35 km/h e 49 km/h. Qual è la velocità, uguale in tutti e tre i giri, che avrebbe consentito di giungere al traguardo nello stesso tempo?

Ragionando in modo simile al problema poc'anzi esaminato, si conclude che la velocità richiesta è la media armonica delle velocità dei tre giri:

$$v_m = \frac{3}{\frac{1}{25} + \frac{1}{35} + \frac{1}{49}} = \frac{3675}{109} \simeq 33{,}72 \text{ km/h}$$

La media geometrica

◆ **PROBLEMA**

Dato un rettangolo i cui lati misurano a e b, quale deve essere la misura m del lato di un quadrato affinché questo abbia la stessa area del rettangolo?

La condizione richiesta dal problema si traduce nell'equazione:

$$m^2 = a \cdot b \qquad \text{da cui} \qquad m = \sqrt{a \cdot b}$$

Il valore m che abbiamo trovato ha la seguente proprietà: sostituito al posto di a e b, ne *lascia invariato il prodotto*; è questa la proprietà che caratterizza la media *geometrica*, che si può definire come segue.

DEFINIZIONE | Media geometrica

Dati *n numeri positivi* $x_1, x_2, ..., x_n$, si definisce loro **media geometrica** la radice n-esima del loro prodotto:

$$\sqrt[n]{x_1 \cdot x_2 \cdot ... \cdot x_n}$$

RICORDA
Se $a \geq 0$, la radice n-esima di a, indicata con $\sqrt[n]{a}$, è l'unico numero reale non negativo che, elevato a n, dà come risultato a. Per esempio: $\sqrt[3]{27} = 3$, $\sqrt[4]{16} = 2$.

Si ricorre alla media geometrica quando si vuole conservare il prodotto di più valori.

ESEMPIO Applicazione della media geometrica

Un risparmiatore ha investito 10 000 euro in un investimento di 2 anni: alla fine del primo anno gli verrà corrisposto un tasso di interesse del 2% sul capitale investito, alla fine del secondo anno un tasso di interesse del 3% calcolato sul capitale complessivo, ottenuto come somma del capitale iniziale e dell'interesse maturato l'anno precedente. Qual è il tasso di interesse che, applicato uguale per entrambi gli anni, lascia invariato il capitale finale ottenuto allo scadere dell'investimento?

Per calcolare il capitale al termine del secondo anno occorre moltiplicare il capitale investito per i coefficienti di incremento (1 + tasso di interesse) corrispondenti a ciascuno dei due anni. Il capitale finale sarà quindi dato da:

$$\underbrace{10\,000}_{\text{capitale iniziale}} \cdot \underbrace{(1 + 0{,}02)}_{\text{coefficiente di incremento relativo al } 1^\circ \text{ anno}} \cdot \underbrace{(1 + 0{,}03)}_{\text{coefficiente di incremento relativo al } 2^\circ \text{ anno}}$$

RICORDA
Se un numero x (in questo caso il capitale) aumenta del p%, il suo valore diviene uguale a:

$$x\left(1 + \frac{p}{100}\right)$$

Il *coefficiente di incremento* (non il tasso di interesse!) che, applicato per tutti e due gli anni, lascia invariato il capitale finale è quello che conserva il prodotto dei due coefficienti $(1 + 0{,}02)$ e $(1 + 0{,}03)$, quindi è la loro media geometrica:

$$\sqrt{1{,}02 \cdot 1{,}03} \simeq 1{,}025$$

Poiché $1{,}025 = 1 + 0{,}025 = 1 + \frac{2{,}5}{100}$, deduciamo che il tasso di interesse che, applicato tutti e due gli anni, lascia invariato il capitale finale è circa il 2,5%.

Esercizi p. 27

3. Rapporti statistici

Il confronto di dati statistici

Per confrontare e interpretare correttamente dati quantitativi provenienti da analisi statistiche è spesso utile calcolare opportune *differenze* e *rapporti* tra i dati stessi.

ESEMPI Confronto di dati mediante differenze e rapporti

Supponiamo di sapere che il fatturato di un'azienda nel 2015 è stato di 4 milioni di euro e nel 2016 è stato di 4,5 milioni di euro.

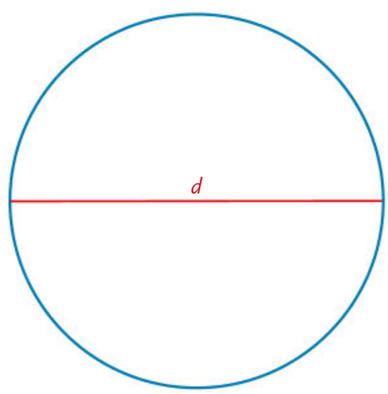
Per confrontare i fatturati dei due anni possiamo procedere in molti modi diversi; per esempio:

a. calcolandone la differenza: $4{,}5 - 4 = 0{,}5$, vale a dire: il fatturato dell'azienda è cresciuto di 0,5 milioni di euro dal 2015 al 2016;

b. calcolandone il rapporto: $\frac{4{,}5}{4} = 1{,}125$, vale a dire: per ogni milione di euro di fatturato realizzato nel 2015, l'azienda ha fatturato 1,125 milioni di euro nel 2016;

c. calcolandone la differenza relativa: $\frac{4{,}5 - 4}{4} = \frac{0{,}5}{4} = 0{,}125$, vale a dire: il fatturato nel 2016 è cresciuto del 12,5% rispetto al 2015.

Come messo in luce dagli ultimi due esempi, per il confronto di dati *quantitativi* è particolarmente utile e frequente l'utilizzo dei cosiddetti *rapporti* statistici. Un **rapporto statistico** è un rapporto tra due dati, di cui almeno uno deve essere di natura statistica (cioè riferirsi a un collettivo), tra i quali sussiste qualche legame logico.

Esempio	Controesempio
Il rapporto tra il numero dei nati in un dato anno e il numero complessivo delle unità della popolazione è un rapporto statistico. 	Il rapporto tra la lunghezza di una circonferenza e il suo diametro **non** è un rapporto statistico (nessuno dei due dati è di natura statistica). 

Esistono molti tipi di rapporti statistici; qui ci limiteremo a descrivere alcuni dei più comuni e importanti, precisamente: i rapporti di composizione, coesistenza, derivazione, densità e i numeri indice.

Rapporti di composizione, coesistenza, derivazione e densità

Tipo di rapporto	Definizione	Esempi	Osservazioni
di composizione	È il rapporto tra due dati, di cui quello al numeratore è *parte* di quello al denominatore.	– Il *tasso di occupazione*, definito come rapporto tra il numero di occupati e la totalità della popolazione (o la totalità della popolazione di età superiore o uguale a 15 anni). – Il *tasso di disoccupazione*, definito come rapporto tra il numero di persone in cerca di lavoro e il numero di persone in forza di lavoro (ossia il numero complessivo di persone occupate e in cerca di occupazione).	– Le frequenze relative, ossia i rapporti tra le frequenze assolute e il numero complessivo di unità di un collettivo, sono particolari rapporti di composizione. – Un rapporto di composizione è sempre compreso tra 0 e 1 ed è privo di unità di misura. – La somma di tutti i rapporti di composizione relativi a una stessa distribuzione è sempre uguale a 1 (o a 100% se i rapporti sono espressi in percentuale).
di derivazione	È il rapporto tra due dati, di cui quello al denominatore è la *causa* o *il presupposto logico* di quello al numeratore.	– L'*indice di sofferenza*, definito per una banca come rapporto tra fidi rientrati e fidi concessi. – Considerando la popolazione come la «causa» di molti fenomeni demografici quali nascite, morti, matrimoni, si considerano rapporti di derivazione, per esempio, l'*indice di natalità* (definito come rapporto tra il numero dei nati e la popolazione in un certo anno), l'*indice di mortalità* (definito analogamente), l'*indice di nuzialità* (rapporto tra il numero di matrimoni e la popolazione in un dato anno).	– Non sempre i rapporti di derivazione danno luogo a un numero compreso tra 0 e 1; per esempio, il rapporto tra la quantità prodotta di un dato bene e il numero di ore necessarie per produrre quella quantità può dare luogo a un rapporto maggiore di 1. – Non sempre un rapporto di derivazione è privo di unità di misura (ciò accade solo se i due termini che vengono posti a rapporto sono omogenei tra loro).
di coesistenza	È il rapporto tra le frequenze di due diverse modalità di un carattere su un medesimo collettivo.	– Il rapporto tra il numero dei nati maschi rispetto alle nate femmine di una popolazione in un dato anno (detto *rapporto di mascolinità alla nascita*). – Il rapporto tra la popolazione di 65 anni o più e la popolazione fino a 14 anni di età in un dato anno (detto *indice di vecchiaia*).	– Le due modalità di cui si mettono a rapporto le frequenze sono spesso due modalità «antitetiche» per le quali è interessante lo studio relativo. – Un rapporto di coesistenza può assumere valori maggiori di 1 ed è costituito da un numero puro, ossia privo di unità di misura.
di densità	È il rapporto tra un dato e una misura (l'area o il volume o altro) del campo in cui il dato viene osservato.	– La *densità di popolazione* di una località, definita come rapporto tra il numero di abitanti di quella località e l'area della sua superficie. – La *densità abitativa* di un complesso residenziale, definita come rapporto tra il numero di abitanti del complesso e il suo volume. – La *densità automobilistica*, definita come rapporto tra il numero di autovetture immatricolate e la popolazione residente (in questo caso la «misura» del campo in cui il dato viene osservato, ossia della popolazione, è la numerosità della popolazione stessa).	– Un rapporto di densità confronta due grandezze non omogenee, ed è perciò dotato di una unità di misura. – Può assumere valori maggiori di 1.

È importante fare alcune osservazioni.

1. I rapporti analizzati in tabella non sono mutuamente esclusivi, ossia una stessa tipologia di rapporto può rientrare in classi diverse.
   Per esempio, i rapporti di composizione possono anche considerarsi particolari rapporti di derivazione; non sempre però è vero il viceversa: per esempio, il rapporto tra il fatturato di un'azienda e il numero di dipendenti dell'azienda può considerarsi un rapporto di derivazione, ma non è un rapporto di composizione.
2. L'utilizzo dei rapporti statistici agevola i confronti perché permette di «depurare» i dati dai fattori che li influenzano, come mostriamo nell'esempio seguente.

ESEMPIO Utilizzo di un rapporto statistico

Abbiamo a disposizione i seguenti dati (fonte: *Italia in cifre*, Istat, 2011):

Area geografica	Occupati	Forze di lavoro
Nord	11 905	12 574
Centro	4832	5209
Sud	6288	7187
Italia	**23 025**	**24 970**

In quale area geografica la situazione, in termini di occupazione, è migliore? E in quale è peggiore?

In questo caso non sarebbe corretto un confronto basato soltanto sui dati *assoluti*, poiché il numero di occupati dipende chiaramente dalla numerosità della popolazione nell'area geografica considerata.
Per un confronto corretto dobbiamo «depurare» i dati da questo fattore, ossia mettere a *rapporto* il numero degli occupati di ciascuna area geografica con la corrispondente forza di lavoro. Costruiamo così i rapporti di composizione in quest'altra tabella:

Area geografica	Rapporti tra occupati e forze di lavoro
Nord	$\frac{11\,905}{12\,574} \simeq 0{,}95 = 95\%$
Centro	$\frac{4832}{5209} \simeq 0{,}93 = 93\%$
Sud	$\frac{6288}{7187} \simeq 0{,}87 = 87\%$
Italia	$\frac{23\,025}{24\,970} \simeq 0{,}92 = 92\%$

Rapporti di composizione

Dall'analisi dei rapporti (espressi sotto forma di percentuale) vediamo che:

- la situazione migliore si registra al Nord, con una percentuale di occupati del 95%;
- la situazione peggiore si ha invece nel Sud, con una percentuale di occupati che è solo dell'87%.

Osserva che, mentre nel Nord si registra anche il maggior numero di occupati in termini assoluti, il minore numero di occupati in termini assoluti non si ha nel Sud bensì nel Centro (che presenta però una situazione migliore in termini di occupazione rispetto al Sud, essendo il tasso di occupati del 93%).

Numeri indice

Nelle applicazioni economiche e sociali è molto frequente imbattersi nell'analisi di una **serie storica**, cioè di dati che derivano dall'osservazione di uno stesso fenomeno (quantitativo) a intervalli regolari di tempo.

ESEMPIO Serie storica

La seguente tabella, che riporta il numero di biglietti venduti al cinema in Italia a intervalli regolari di 3 anni dal 2006 al 2015, rappresenta una serie storica.

Anno	Biglietti venduti al cinema (in migliaia)
2006	100 911
2009	105 030
2012	104 980
2015	109 229

OSSERVA
Quando si rileva un fenomeno ripetutamente nel tempo, le unità statistiche sono gli istanti temporali di osservazione; nell'esempio qui a fianco, i quattro anni: 2006, 2009, 2012, 2015.

Poiché nell'analisi di una serie storica interessa studiare l'*evoluzione* di un fenomeno nel tempo, non avrebbe senso costruire una distribuzione di frequenze (perché si perderebbe proprio una delle informazioni fondamentali, ossia l'*ordine temporale* in cui sono state osservate le varie modalità); spesso risulta invece utile costruire particolari *rapporti* statistici, detti **numeri indice**, che consentono di quantificare la *variazione* del fenomeno tra due diversi istanti temporali.

DEFINIZIONE | Numero indice

Si chiama **numero indice** il rapporto tra due modalità di un carattere quantitativo rilevate in due differenti istanti temporali.

Se indichiamo con x_1, x_2, ..., x_t le varie modalità osservate, rispettivamente negli istanti temporali 1, 2, ..., t, un numero indice è quindi un rapporto del tipo $\frac{x_i}{x_j}$, con $i, j \in \{1, 2, ..., t\}$.

Esistono due tipologie di numeri indice, a **base fissa** e a **base mobile**.

- I numeri indice **a base fissa** si costruiscono scegliendo un istante temporale come base (in generale l'istante iniziale di rilevazione) e costruendo la sequenza dei numeri ottenuti rapportando ciascuna modalità osservata con la modalità corrispondente all'istante-base, diciamo x_1; si ottengono così esattamente t numeri indice a base fissa:

$$\frac{x_1}{x_1}, \frac{x_2}{x_1}, ..., \frac{x_t}{x_1}$$

- I numeri indice **a base mobile** si costruiscono rapportando ciascuna modalità x_i con la modalità x_{i-1} relativa all'istante temporale precedente. La base (ossia il denominatore) cambia in ogni rapporto, motivo per cui si parla di base *mobile*. I rapporti di questo tipo possono costruirsi solo a partire dal secondo istante di rilevazione (perché altrimenti non sarebbe definito il termine precedente); dunque si avranno esattamente $t - 1$ numeri indice a base mobile:

$$\frac{x_2}{x_1}, \frac{x_3}{x_2}, ..., \frac{x_t}{x_{t-1}}$$

MODI DI DIRE
Una *sequenza* dei numeri indice (a base fissa o mobile) viene anche chiamata **serie** di numeri indice.

ESEMPIO Numeri indice a base fissa e mobile

Calcoliamo, per la serie storica rappresentata nell'esempio precedente, i numeri indice a base fissa (scegliendo come anno base il 2006) e i numeri indice a base mobile.

Organizziamo il lavoro nella seguente tabella.

Anno	Biglietti venduti al cinema (migliaia)	Numeri indice a base fissa (anno base 2006)	Numeri indice a base mobile
2006	100 911	$1 = 100\%$	Non definito
2009	105 030	$\frac{105\,030}{100\,911} \simeq 1{,}041 = 104{,}1\%$	$\frac{105\,030}{100\,911} \simeq 1{,}041 = 104{,}1\%$
2012	104 980	$\frac{104\,980}{100\,911} \simeq 1{,}040 = 104\%$	$\frac{104\,980}{105\,030} \simeq 0{,}999 = 99{,}9\%$
2015	109 229	$\frac{109\,229}{100\,911} \simeq 1{,}082 = 108{,}2\%$	$\frac{109\,229}{104\,980} \simeq 1{,}040 = 104\%$

OSSERVA

Per determinare le *variazioni* percentuali basta calcolare le differenze tra i rapporti espressi in percentuale e il 100%; per esempio, relativamente ai numeri indici a base fissa e mobile corrispondenti al 2012:
$104\% - 100\% = +4\%$
$99{,}9\% - 100\% = -0{,}1\%$

Dall'analisi dei numeri indice ricavati possiamo trarre importanti conseguenze sulle *variazioni* percentuali dei biglietti venduti negli anni presi in considerazione.
In particolare, dall'analisi dei numeri *indice a base fissa* possiamo dedurre che il numero di biglietti venduti nel 2009, 2012, 2015 è cresciuto, rispetto all'anno 2006, rispettivamente circa del 4,1%, del 4% e dell'8,2%. Dall'analisi dei numeri *indice a base mobile* possiamo dedurre che nel 2009 il numero di biglietti venduti è cresciuto di circa il 4,1% rispetto al 2006, che nel 2012 il numero di biglietti venduti è diminuito di circa lo 0,1% rispetto al 2009 e infine che nel 2015 il numero di biglietti venduti è cresciuto di circa il 4% rispetto al 2012.

Esercizi p. 33

4. Indicatori di efficacia, efficienza e qualità

I rapporti statistici di cui abbiamo parlato nel precedente paragrafo, oltre a consentire confronti altrimenti difficoltosi, vengono spesso utilizzati in ambito economico-aziendale per la costruzione di numeri, detti **indicatori**, utili all'investigazione di fenomeni, quali per esempio la qualità di un prodotto o di un servizio, che per la loro complessità **non** possono essere rilevati in modo diretto.
I parametri per una valutazione quantitativa della *qualità* sono generalmente tre: l'**efficacia**, l'**efficienza** e la **soddisfazione del cliente** (*customer satisfaction*):

1. l'*efficacia* esprime il raggiungimento degli obiettivi dell'impresa (per esempio: soddisfare la domanda, raggiungere un determinato fatturato, «fidelizzare» la clientela, assicurare un servizio atteso);
2. l'*efficienza* esprime la capacità dell'azienda di utilizzare al meglio le risorse disponibili (per esempio: minimizzare i costi e le ore di lavorazione oppure garantire forniture in tempi rapidi);
3. la *soddisfazione del cliente* è il risultato dei giudizi espressi dagli utilizzatori di un dato prodotto o servizio.

Gli indicatori di questi tre parametri sono perlopiù definiti tramite frequenze assolute oppure tramite *differenze* o *rapporti statistici* oppure tramite indici sintetici di posizione o di variabilità, quali per esempio *medie*, *mediane* o *deviazioni standard*.

ESEMPIO Indicatori di efficacia ed efficienza

I titolari di uno studio dentistico si pongono l'obiettivo di incrementare il più possibile il profitto. Nella tabella sono riportati alcuni dati relativi all'anno 2017, circa il numero di pazienti, il profitto complessivo e il tempo totale di attesa dei pazienti. Definiamo dei possibili indicatori di efficacia ed efficienza e commentiamo i risultati raggiunti dallo studio dentistico in termini di efficacia ed efficienza nell'anno in esame.

Quadrimestre (anno 2017)	Numero di pazienti	Profitto complessivo (in euro)	Tempo complessivo di attesa dei pazienti (in ore)
Primo	625	62 000	104
Secondo	450	48 000	65
Terzo	712	85 000	115

Poiché l'obiettivo dichiarato è il raggiungimento del massimo profitto, come indicatore di *efficacia* possiamo considerare proprio il dato quadrimestrale del profitto complessivo (terza colonna). Tale indicatore è *massimo* nel terzo quadrimestre (in corrispondenza, tra l'altro, anche del massimo numero di pazienti). Come indicatore di efficienza possiamo considerare il tempo *medio* di attesa dei pazienti, calcolato come rapporto tra il tempo complessivo di attesa e il numero dei pazienti. Calcoliamolo, quadrimestre per quadrimestre, osservando che il servizio offerto dallo studio dentistico è tanto più efficiente quanto più l'indicatore fornisce un valore piccolo:

Primo quadrimestre: $\frac{104}{625} \simeq 0{,}17$

Secondo quadrimestre: $\frac{65}{450} \simeq 0{,}14$

Terzo quadrimestre: $\frac{115}{712} \simeq 0{,}16$

È interessante osservare che la *massima efficienza* del servizio corrisponde alla *peggiore efficacia*: il confronto degli indicatori potrebbe suggerire ai titolari dello studio alcuni correttivi di tipo gestionale; per esempio una diversa pianificazione degli appuntamenti o l'assunzione di assistenti sanitari più qualificati o l'allestimento di un'ulteriore sala.

Nel paragrafo di esercizi corrispondente a questo paragrafo ti proporremo alcuni problemi in cui dovrai utilizzare le tue conoscenze di statistica per la costruzione di opportuni indicatori di qualità. Per tale compito non ci sono ovviamente regole generali, in quanto l'opportunità di un indicatore dipende in modo essenziale dal particolare contesto cui si riferisce e dallo scopo che l'analisi si prefigge; in generale, però, un buon indicatore dovrebbe essere il più possibile semplice da costruire, ricavabile in modo oggettivo, facile da rappresentare (per mezzo di tabelle, istogrammi, diagrammi ecc.) e semplice da interpretare.

Esercizi p. 35

5. Distribuzione normale e introduzione all'inferenza

Distribuzione normale

Sappiamo che, per rappresentare tramite un istogramma una distribuzione di frequenze di un carattere quantitativo continuo, si è soliti suddividere l'intervallo delle possibili modalità in un numero finito n di intervalli (classi), ciascuno dei quali sarà la base di uno dei rettangoli dell'istogramma. Al fine di poter confrontare agevolmente istogrammi relativi a suddivisioni differenti, si conviene di solito di costruire istogrammi la cui area complessiva abbia sempre valore 1: un tale istogramma, se le classi hanno tutte le stessa ampiezza e se si assume sull'asse delle ascisse come unità di misura l'ampiezza di una classe, si può ottenere semplicemente costruendo rettangoli la cui altezza sia uguale alla frequenza relativa della classe. Se la popolazione considerata è molto numerosa, il numero n di classi può essere aumentato a piacere, diminuendone l'ampiezza.
Per esempio, consideriamo l'istogramma rappresentato in **Fig. 1**: al crescere di n si ottengono istogrammi via via più regolari e il cosiddetto poligono delle frequenze (colorato in rosso nelle **Figg. 1** e **2**) può essere approssimato da una curva continua, che tende generalmente ad assumere la forma «a campana» illustrata in **Fig. 3**.

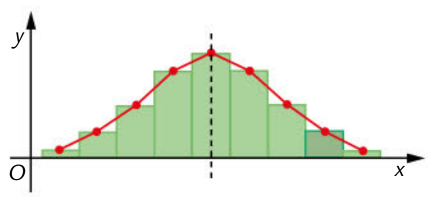


Figura 1

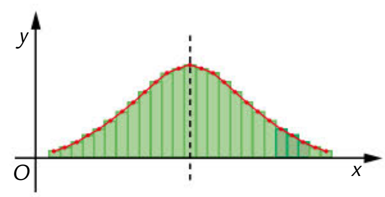


Figura 2

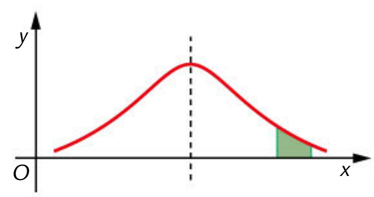


Figura 3

Questa curva limite, che studieremo più ampiamente nella prosecuzione del corso, ha equazione della forma:

$$y = Ae^{-B(x-C)^2}$$

dove i coefficienti A, B e C possono essere individuati, una volta che si conoscano la media μ e la deviazione standard σ della distribuzione; risulta infatti:

$$A = \frac{1}{\sigma\sqrt{2\pi}} \qquad B = \frac{1}{2\sigma^2} \qquad C = \mu$$

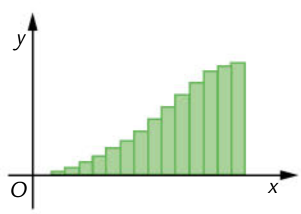


Figura 4

Una distribuzione di frequenze che individua una tale curva a campana è detta **normale** o **gaussiana**. Naturalmente non tutte le distribuzioni di frequenze sono normali: per esempio non lo è la distribuzione rappresentata in **Fig. 4**, chiaramente asimmetrica. In generale solo dall'analisi della distribuzione è possibile stabilire se essa è o meno gaussiana; tuttavia in molti casi è possibile prevedere a priori in base a delle considerazioni *teoriche* (su cui torneremo nel proseguimento del corso) che un certo insieme di dati sperimentali avrà una distribuzione normale. Per esempio, se si misura ripetutamente una medesima grandezza, a causa di piccoli errori casuali le misure ottenute generalmente non coincidono tra loro ma (in assenza di errori sistematici) tali misure tendono a distribuirsi secondo una distribuzione normale intorno alla «vera» misura della grandezza in esame, che è identificabile con la media dei risultati delle varie misurazioni effettuate.

Si dimostra che, se una distribuzione è normale, allora:

- nell'intervallo $[\mu - \sigma, \mu + \sigma]$ cade circa il 68,27% dei dati;
- nell'intervallo $[\mu - 2\sigma, \mu + 2\sigma]$ cade circa il 95,45% dei dati;
- nell'intervallo $[\mu - 3\sigma, \mu + 3\sigma]$ cade circa il 99,73% dei dati.

Questa proprietà è illustrata in **Fig. 5**.

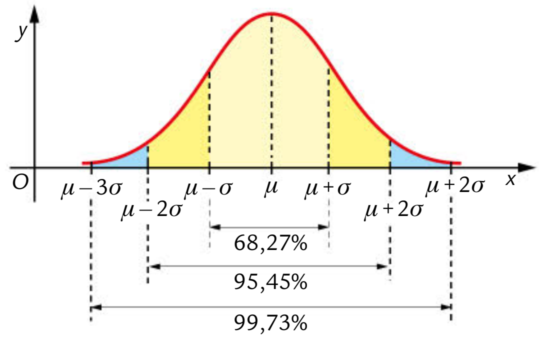


Figura 5 Intervalli tipici della normale.

Altri intervalli fondamentali per gli sviluppi che vedremo sono i seguenti:

- nell'intervallo $[\mu - 1{,}65\sigma,\ \mu + 1{,}65\sigma]$ cade circa il 90% dei dati;
- nell'intervallo $[\mu - 1{,}96\sigma,\ \mu + 1{,}96\sigma]$ cade circa il 95% dei dati;
- nell'intervallo $[\mu - 2{,}58\sigma,\ \mu + 2{,}58\sigma]$ cade circa il 99% dei dati.

ESEMPIO Intervalli tipici della normale

Si sono eseguite 20 misure della concentrazione di un farmaco nel sangue di diversi pazienti e si è ottenuta una concentrazione media di 5 mg/ml con una deviazione standard di 0,2 mg/ml. Supponendo che le varie misure siano distribuite normalmente, individuiamo un intervallo entro cui cadono il 95% delle misure effettuate.

In base a quanto detto poc'anzi, possiamo essere certi che circa il 95% delle misurazioni cadono nell'intervallo $[\mu - 1{,}96\sigma,\ \mu + 1{,}96\sigma]$; essendo in questo caso $\mu = 5$ e $\sigma = 0{,}2$, l'intervallo diventa:

$$[5 - 1{,}96 \cdot 0{,}2,\ 5 + 1{,}96 \cdot 0{,}2] \approx [4{,}6;\ \ 5{,}4]$$

Ciò significa che il 95% circa delle misure effettuate fornisce una concentrazione compresa tra 4,6 mg/ml e 5,4 mg/ml.

Introduzione all'inferenza: gli intervalli di confidenza

Consideriamo i seguenti problemi.

◆ PROBLEMI

1. Si vuole stimare la statura media degli abitanti di una certa regione. È emerso che la statura media di un campione di 100 abitanti della regione è di 174 cm, con deviazione standard di 8 cm. A partire da queste informazioni è possibile trarre informazioni sulla statura media degli abitanti dell'intera regione?
2. In base a un sondaggio effettuato su un campione di 500 elettori, è emerso che il 25% dei componenti del campione voteranno per il partito A. A partire da questa informazione è possibile trarre informazioni sulla percentuale di elettori dell'intera popolazione che voteranno per il partito A?

Questi problemi hanno una caratteristica in comune: si vuole stimare un parametro ignoto della *popolazione* (l'altezza media nel primo problema e la percentuale di elettori che voteranno il partito A nel secondo) a partire da dati raccolti soltanto su un *campione*. È intuitivo pensare che si possa ottenere una *stima* dei parametri che vogliamo calcolare sulla *popolazione* calcolando tali parametri sul *campione*; questa idea intuitiva è supportata da considerazioni teoriche che vedremo nel proseguimento del corso, in base alle quali la media μ di una popolazione è effettivamente stimata «bene» dalla media del campione (che indichiamo con $\bar{x}$) e la percentuale p di una popolazione che soddisfa una determinata caratteristica (nel secondo problema votare il partito A) è effettivamente stimata «bene» dalla corrispondente percentuale calcolata sul campione (che indicheremo con $\hat{p}$).

Tuttavia la stima *cambia* al variare del campione; dunque la stima è soggetta a una *incertezza*, che è espressa dal cosiddetto **errore standard**, che indicheremo con e_s e che è così definito:

Errore standard da cui è affetta la stima $\bar{x}$, calcolata su un campione di numerosità n, della media μ di una popolazione	**Errore standard da cui è affetta la stima $\hat{p}$, calcolata su un campione di numerosità n, della percentuale p di una popolazione soddisfacente una data proprietà**
$e_s = \dfrac{\sigma}{\sqrt{n-1}}$ σ è la deviazione standard calcolata sul campione	$e_s = \sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$

Proprio a causa della variabilità, al variare del campione, della stima della media o della percentuale non è opportuno esprimere tale stima con un solo valore, ma piuttosto fornire un intervallo, detto **intervallo di confidenza**, che contiene il vero valore del parametro da stimare con un prefissato grado di fiducia, detto **livello di confidenza**: per esempio, un livello di confidenza del 95% indica che è uguale al 95% la probabilità di estrarre un campione casuale il cui corrispondente intervallo di confidenza contiene effettivamente l'ignoto valore del parametro da stimare.
Gli intervalli di confidenza per la media e la proporzione, nel caso di campioni sufficientemente numerosi (deve essere $n > 30$), sono descritti negli schemi seguenti.

MODI DI DIRE

In alternativa alle espressioni «intervallo di confidenza» e «livello di confidenza» si utilizzano le espressioni «intervallo di fiducia» e «livello di fiducia».

Intervallo di confidenza per la media	**Intervallo di confidenza per la percentuale**
costante che dipende dal livello di confidenza $[\bar{x} - k e_s, \bar{x} + k e_s]$ stima della media calcolata sul campione errore standard cui è soggetta la stima della media	costante che dipende dal livello di confidenza $[\hat{p} - k e_s, \hat{p} + k e_s]$ stima della percentuale calcolata sul campione errore standard cui è soggetta la stima della percentuale

I valori di k da utilizzare, in corrispondenza dei più frequenti livelli di confidenza, sono riassunti nella tabella seguente:

Livello di confidenza	90%	95%	99%
k	1,645	1,96	2,58

ESEMPI Intervalli di confidenza

In riferimento ai due problemi proposti all'inizio del sottoparagrafo, determiniamo:

a. un intervallo di confidenza al livello del 95% per l'altezza media della popolazione;
b. un intervallo di confidenza al livello del 95% per la proporzione della popolazione che voterà per il partito A.

a. Per determinare l'intervallo di confidenza ci servono $\bar{x}$, k ed e_s; abbiamo che:

$$\underbrace{\bar{x} = 174}_{\text{altezza media sul campione}} \qquad \underbrace{k = 1{,}96}_{\text{valore di } k \text{ corrispondente al livello del 95\%}} \qquad e_s = \underbrace{\frac{\sigma}{\sqrt{n-1}} = \frac{8}{\sqrt{100-1}}}_{\text{la deviazione standard sul campione è di 8 cm e la numerosità del campione è } n = 100} = \frac{8}{\sqrt{99}}$$

Dunque l'intervallo di confidenza richiesto è:

$$\left[174 - 1{,}96 \cdot \frac{8}{\sqrt{99}};\quad 174 + 1{,}96 \cdot \frac{8}{\sqrt{99}}\right] \approx [172{,}4;\quad 175{,}6]$$

In conclusione, al livello di confidenza del 95%, l'altezza media della popolazione sarà compresa tra 172,4 cm e 175,6 cm.

b. Per determinare l'intervallo di confidenza ci servono $\hat{p}$, k ed e_s; abbiamo che:

$$\underbrace{\hat{p} = 25\% = 0{,}25}_{\text{percentuale valutata sul campione}} \qquad \underbrace{k = 1{,}96}_{\text{valore di } k \text{ corrispondente al livello del 95\%}} \quad e_s = \underbrace{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0{,}25(1-0{,}25)}{500}}}_{\text{la percentuale stimata sul campione è il 25\% e la numerosità del campione è } n = 500} = \frac{\sqrt{15}}{200}$$

Dunque l'intervallo di confidenza richiesto è:

$$\left[0{,}25 - 1{,}96 \cdot \frac{\sqrt{15}}{200};\quad 0{,}25 + 1{,}96 \cdot \frac{\sqrt{15}}{200}\right] \approx [0{,}21;\quad 0{,}29]$$

In conclusione, al livello di confidenza del 95%, la percentuale di elettori che voteranno per il partito A sarà compresa tra il 21% e il 29%.

ATTENZIONE!

Nell'eseguire i calcoli per determinare un intervallo di confidenza è bene approssimare l'estremo inferiore dell'intervallo per *difetto* e l'estremo superiore per *eccesso*, in modo da essere certi di ottenere un intervallo di confidenza approssimato che *contiene* quello esatto. D'ora in avanti, per approssimare gli estremi di un intervallo di confidenza procederemo in questo modo.

Gli intervalli di confidenza e la verifica di ipotesi

Gli intervalli di confidenza possono essere utilizzati anche per stabilire se una determinata *ipotesi* circa il valore della media o della percentuale di una data caratteristica della popolazione è supportata dai dati ricavati su un campione (e quindi può essere *accettata*) oppure se va *rifiutata*. Naturalmente l'accettazione o il rifiuto dell'ipotesi dipende dal livello di confidenza fissato; occorrerà quindi anzitutto fissare tale livello di confidenza, quindi determinare l'intervallo di confidenza corrispondente: se il valore ipotizzato per la media o la percentuale appartiene a tale intervallo accetteremo l'ipotesi, altrimenti la rifiuteremo.

ESEMPIO Verifica di una ipotesi tramite gli intervalli di confidenza

Una casa farmaceutica afferma che un farmaco di sua produzione impiega in media 15 minuti ad agire. In base alle analisi effettuate su un campione di 50 pazienti, il beneficio è stato ottenuto in media in 18 minuti, con una deviazione standard di 5 minuti. A un livello di confidenza del 95% possiamo accettare quanto dichiarato dalla casa farmaceutica?

L'intervallo di confidenza della media dei tempi di azione, al livello del 95%, è:

$$\left[\bar{x} - 1{,}96 \cdot \frac{\sigma}{\sqrt{n-1}};\ \bar{x} + 1{,}96 \cdot \frac{\sigma}{\sqrt{n-1}}\right] = \left[18 - 1{,}96 \cdot \frac{5}{\sqrt{49}};\ 18 + 1{,}96 \cdot \frac{5}{\sqrt{49}}\right]$$

ossia approssimativamente:

[16,6; 19,4]

Poiché il tempo medio dichiarato (15 minuti) **non** appartiene a tale intervallo, abbiamo motivo di ritenere che quanto dichiarato dalla casa farmaceutica non sia veritiero e quindi di *rifiutare* l'ipotesi che il tempo medio di azione sia di 15 minuti.

Esercizi p. 38

Statistica

Studio quantitativo di fenomeni collettivi, osservabili nella realtà sociale, in natura o in laboratorio.

Indici di posizione

Numeri che sintetizzano i dati di un'indagine statistica e che permettono di cogliere aspetti importanti del fenomeno in esame.

Moda

È la modalità che si presenta con la massima frequenza in un'indagine statistica.

ESEMPIO

La moda della sequenza di numeri 1, 2, 2, 3, 3, 4, 5, 5, 5 è il numero 5.

Mediana

Dati n numeri, ordinati in senso crescente o decrescente, è:
- il numero che occupa la posizione centrale, se n è dispari;
- la media aritmetica dei due numeri che occupano le posizioni centrali, se n è pari.

ESEMPI

- 0, 2, $\boxed{4}$, 5, 6 → **4 è la mediana**
- 1, 2, $\boxed{4, 5}$, 6, 8 → $\frac{4+5}{2} = 4{,}5$ **è la mediana**

Media aritmetica

Dati n numeri $x_1, x_2, \ldots, x_n$ è il numero indicato con il simbolo $\bar{x}$ e definito dalla formula:

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

ESEMPIO

Un alunno ha preso 5, 6, 8 e 4 nei compiti di matematica.

La media aritmetica dei suoi voti è

$$\frac{5+6+8+4}{4} = 5{,}75$$

Indici di variabilità

Numeri che permettono di misurare la **variabilità** del fenomeno in esame, cioè l'attitudine del fenomeno a manifestarsi sulle varie unità statistiche con modalità diverse e distanti tra loro.

Varianza

Dati n numeri $x_1, x_2, \ldots, x_n$ di media aritmetica $\bar{x}$, è la media aritmetica dei quadrati degli scarti da $\bar{x}$ e si indica σ^2:

$$\sigma^2 = \frac{(x_1 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2}{n}$$

In alternativa vale la formula:

$$\sigma^2 = \frac{x_1^2 + \ldots + x_n^2}{n} - \bar{x}^2$$

Deviazione standard (o scarto quadratico medio)

Dati n numeri $x_1, x_2, \ldots, x_n$, la loro **deviazione standard** (che si indica con σ) è la radice quadrata della loro varianza.

ESEMPIO

Un alunno ha preso 5, 6, 8 e 4 nei compiti di matematica. La media dei voti è 5,75, quindi la varianza è:

$$\sigma^2 = \frac{5^2 + 6^2 + 8^2 + 4^2}{4} - 5{,}75^2 = 2{,}1875$$

e la deviazione standard vale quindi: $\sigma = \sqrt{2{,}1875} \simeq 1{,}48$.

Coefficiente di variazione

Il **coefficiente di variazione** C_V di un insieme di dati la cui media aritmetica è $\bar{x}$, con $\bar{x} \neq 0$, e la cui deviazione standard è σ, è il rapporto tra la deviazione standard e la media aritmetica stessa:

$$C_V = \frac{\sigma}{\bar{x}}$$

Rapporti statistici

Rapporti che si utilizzano per confrontare e interpretare dati quantitativi.

Di composizione

Rapporto tra due dati, di cui quello al numeratore è parte di quello al denominatore.

ESEMPIO

Il rapporto tra il numero degli occupati e la totalità della popolazione.

Di derivazione

Rapporto tra due dati, di cui quello al numeratore è causa o presupposto logico di quello al denominatore.

ESEMPIO

Il rapporto tra la qualità prodotta di un dato bene e il tempo necessario a produrla.

Di coesistenza

Rapporto tra le frequenza di due modalità di un carattere su un medesimo collettivo.

ESEMPIO

Il rapporto tra il numero dei nati maschi e il numero dei nati femmine di una popolazione in un dato anno.

Di densità

Rapporto tra un dato e una misura (per esempio l'area) del campo in cui il dato viene osservato.

ESEMPIO

Il rapporto tra il numero di abitanti di una località e l'area della sua superficie.

Numeri indice

Rapporto tra due modalità di un carattere quantitativo rilevate in due differenti istanti temporali. Esistono due tipi di numeri indice: a **base fissa** e a **base mobile**.

Intervallo di confidenza

Intervallo che contiene un parametro incognito da sistemare (per esempio una media o una percentuale), con un prefissato grado di fiducia, detto **livello di confidenza**.

Per la media

$$[\overline{x} - ke_s, \overline{x} + ke_s]$$

$\overline{x}$ è la media calcolata sul campione

$$\underbrace{e_s}_{\text{errore standard}} = \frac{\sigma}{\sqrt{n-1}}$$

σ è la deviazione standard calcolata sul campione n è la numerosità del campione

$$k = \begin{cases} 1{,}645 & \text{se il livello di confidenza è 90\%} \\ 1{,}96 & \text{se il livello di confidenza è 95\%} \\ 2{,}58 & \text{se il livello di confidenza è 99\%} \end{cases}$$

Per la percentuale

$$[\hat{p} - ke_s, \hat{p} + ke_s]$$

$\hat{p}$ è la percentuale calcolata sul campione

$$\underbrace{e_s}_{\text{errore standard}} = \frac{p(1-p)}{n}$$

n è la numerosità del campione

$$k = \begin{cases} 1{,}645 & \text{se il livello di confidenza è 90\%} \\ 1{,}96 & \text{se il livello di confidenza è 95\%} \\ 2{,}58 & \text{se il livello di confidenza è 99\%} \end{cases}$$

1. Richiami di statistica descrittiva

Teoria p. 2

Il linguaggio della statistica

1 Si osservano tutte le automobili che transitano dal casello autostradale di Firenze Sud in un dato periodo di tempo e si registra la marca dell'automobile. Individua il collettivo, le unità statistiche, il carattere e alcune possibili modalità di questa indagine statistica. Precisa se il carattere è qualitativo o quantitativo.

2 Nello stabilimento Fiat di Mirafiori si rilevano gli stipendi mensili dei dipendenti. Individua il collettivo, le unità statistiche, il carattere e alcune possibili modalità di questa indagine statistica. Precisa se il carattere è qualitativo o quantitativo.

3 Indica alcune delle modalità che possono assumere i seguenti caratteri:

a. voto finale all'esame di Stato;
b. colore dei capelli;
c. genere di un romanzo;
d. numero di giorni di malattia di un lavoratore in un mese.

Stabilisci inoltre quali caratteri sono quantitativi e quali sono qualitativi.

4 Stabilisci quali fra i seguenti caratteri quantitativi sono continui e quali discreti:

a. peso delle foglie dello stesso tipo in un giardino in un pomeriggio;
b. numero di automobili per famiglia;
c. famiglie per condominio;
d. passeggeri per giorno all'aeroporto di Malpensa;
e. lunghezza di un capello.

Le distribuzioni di frequenze

5 **Vero o falso?**
A un concorso fotografico è stato assegnato un voto da 1 a 5 a ciascuna foto partecipante. La tabella che rappresenta la distribuzione delle frequenze dei voti è la seguente.

Voto	1	2	3	4	5
Frequenza assoluta	6	2	8	3	1

Stabilisci quali affermazioni sono vere e quali sono false:

a. le foto complessivamente in concorso sono 20 V F
b. la frequenza relativa del voto 3 è 0,4 V F
c. la frequenza assoluta del voto 2 è maggiore della frequenza assoluta del voto 4 V F
d. la frequenza cumulata del voto 4 è 11 V F
e. la frequenza cumulata del voto 5 è 20 V F

[3 affermazioni vere e 2 false]

6 Completa la seguente tabella (arrotonda le frequenze relative a meno di un decimo).

Modalità	2	1,5	1,8	5	**Totale**
Frequenza assoluta	23	18	.....	15	**80**
Frequenza relativa	.....	.....	.....	.....	**1**

7 Completa la seguente tabella (arrotonda le frequenze relative a meno di un decimo).

Modalità	1	2	3	4	**Totale**
Frequenza assoluta	.....	20	.....	.....	.....
Frequenza relativa	0,2	0,25	.....	0,5	1

8 **Videolezione** Si sono rilevati gli stipendi mensili (in euro) degli impiegati di una piccola azienda e si sono ottenuti i seguenti dati grezzi:

1200	1100	1000	1400	1200	1500	2000	1800	2200	3200	2000	1200	1000

Costruisci la tabella che rappresenta la distribuzione delle frequenze assolute e relative degli stipendi rilevati.

9 Uno studente registra il tempo che impiega per svolgere quindici esercizi di matematica, simili a quelli che deve affrontare nel compito in classe il giorno successivo. I tempi di svolgimento degli esercizi, espressi in minuti, sono i seguenti:

4	5	6	12	10	8	3	4	7	14	18	15	19	13	8

Supponi di suddividere i dati nelle quattro classi:

$0 \leq t < 5 \qquad 5 \leq t < 10 \qquad 10 \leq t < 15 \qquad 15 \leq t < 20$

(t indica il tempo, espresso in minuti, impiegato per risolvere un esercizio) e completa la seguente tabella, che rappresenta la distribuzione delle frequenze assolute, relative e percentuali delle classi.

Classe	Frequenza assoluta	Frequenza relativa	Frequenza percentuale
$0 \leq t < 5$	.....	.....	.....
$5 \leq t < 10$	.....	.....	.....
$10 \leq t < 15$	.....	.....	.....
$15 \leq t < 20$	.....	.....	.....

10 In un grande magazzino è stata compiuta un'indagine sulle taglie di pantaloni da uomo venduti durante una settimana e si sono ottenuti i dati riassunti nella seguente tabella:

Taglia	44	46	48	50	52	54	56
Frequenza	2	7	6	12	10	2	1

Completa la tabella con tre righe, in cui riporterai le frequenze relative, le frequenze percentuali e le frequenze percentuali cumulate, quindi rispondi alle seguenti domande.

a. Qual è la percentuale di pantaloni venduti di taglia 48?
b. Qual è la percentuale di pantaloni venduti di taglia inferiore o uguale alla 50?
c. Qual è la percentuale di pantaloni venduti di taglie superiori alla 50?
d. Qual è la percentuale di pantaloni venduti di taglie comprese tra la 48 e la 52, incluse la 48 e la 52?

[**a.** 15%; **b.** 67,5%; **c.** 32,5%; **d.** 70%]

Le rappresentazioni grafiche

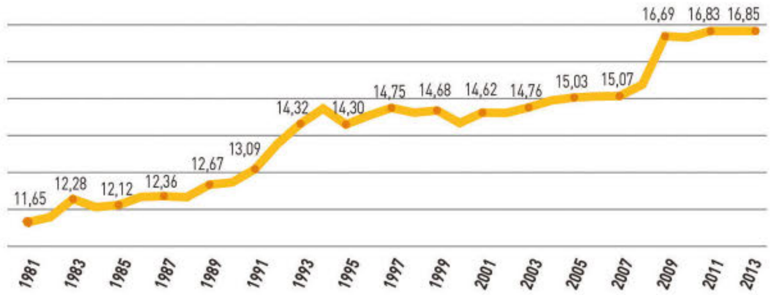
11 La seguente tabella si riferisce alla distribuzione delle frequenze dei giocattoli venduti in un negozio in una settimana.

Giorno	Lunedì	Martedì	Mercoledì	Giovedì	Venerdì	Sabato
Frequenza assoluta	25	32	50	15	30	68

Costruisci il diagramma cartesiano corrispondente.

12 Gli studenti iscritti all'istituto tecnico di Mattia sono così ripartiti: 256 all'indirizzo meccanico, 232 all'indirizzo elettronico, 184 all'indirizzo informatico. Rappresenta la distribuzione degli studenti tramite un diagramma circolare.

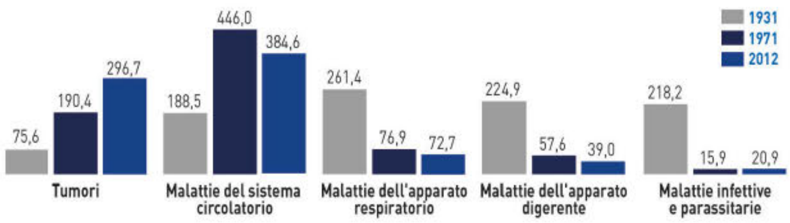
13 Il grafico mostra l'incidenza percentuale sul PIL della spesa pensionistica in Italia negli anni 1981-2013.
(*Fonte*: Italia in cifre, ISTAT)



Individua, relativamente alla sequenza degli anni indicati in tabella:

a. In quale anno l'incidenza sul PIL è risultata massima;
b. In quale anno l'incidenza sul PIL è risultata minima;
c. i due anni successivi tra cui si è registrato il massimo aumento nell'incidenza sul PIL;
d. i due anni successivi tra cui si è registrato il minimo aumento nell'incidenza sul PIL;
e. i due anni successivi tra cui si è registrata la massima diminuzione nell'incidenza sul PIL;
f. i due anni successivi tra cui si è registrata la minima diminuzione nell'incidenza sul PIL.

14 Nel diagramma sono confrontate alcune cause di mortalità nei tre anni 1931, 1971, 2012.
(*Fonte*: Italia in cifre, ISTAT)



a. Qual era la più frequente causa di mortalità, fra quelle indicate, nel 1931? E nel 1971? E nel 2012?
b. Di quanto è aumentata in percentuale la mortalità per tumori dal 1931 al 2012?
c. Di quanto è diminuita in percentuale la mortalità per malattie del sistema circolatorio dal 1971 al 2012?

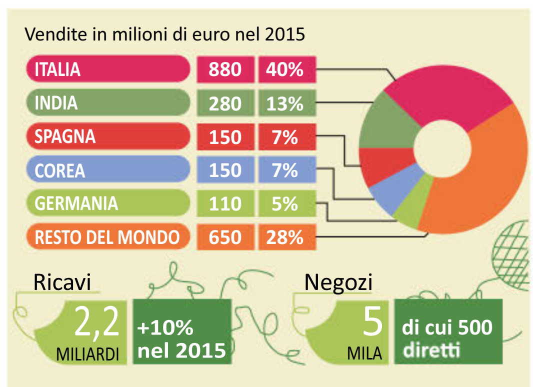
[**b.** Circa 292,5%; **c.** circa 13,8%]

15 **Dal giornale**

Il grafico in figura è comparso in un articolo riguardante le vendite del gruppo Benetton nel mondo nel 2015. L'aumento percentuale dei ricavi è calcolato rispetto al 2014.
(*Fonte*: *Corriere della sera*, 17 febbraio 2016)

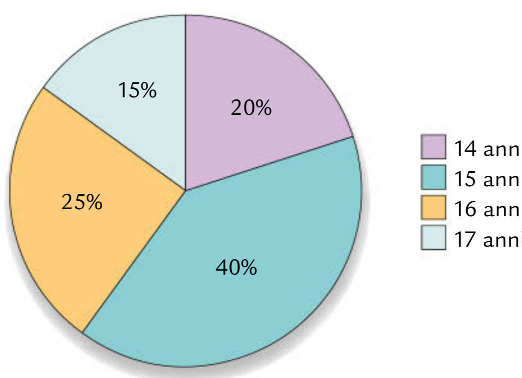
a. Calcolando le percentuali indicate e approssimando i risultati per difetto alle decine di milioni di euro, i dati riportati relativi alle vendite risultano tutti corretti, tranne uno: trova l'errore.
b. Il diagramma circolare rappresenta correttamente i dati delle vendite riportati o contiene errori?
c. Quale è stato il ricavo del gruppo nel 2014?

[**c.** 2 miliardi]



16 Il seguente diagramma a torta rappresenta la suddivisione per età dei 240 allievi di una scuola.

a. Quanti allievi hanno 17 anni?
b. Qual è la frequenza relativa degli allievi che hanno 16 anni?
c. Quanti allievi hanno meno di 16 anni?
d. Quanti allievi hanno più di 15 anni?



[**a.** 36; **b.** 0,25; **c.** 144; **d.** 96]

17 Si è rilevato il numero di automobili possedute da un campione di 20 famiglie italiane e si sono ottenuti i dati grezzi riportati nella seguente tabella.

Famiglia	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Numero di automobili possedute	0	1	2	0	1	1	2	2	4	2	0	3	1	2	0	2	3	2	2	1

a. Costruisci la distribuzione delle frequenze assolute e relative del numero di automobili possedute dalle famiglie.
b. Rappresenta la distribuzione delle frequenze assolute tramite il diagramma che ritieni opportuno.

18 Nella tabella sottostante sono riportate le aree delle superfici degli oceani. Rappresenta questi dati con il grafico che ritieni più opportuno.

Oceano	Superficie (in milioni di km^2)
Pacifico	183,4
Atlantico	106,7
Indiano	73,8

2. Richiami sugli indici di posizione e di variabilità

Teoria p. 5

Media, moda e mediana

19 **Vero o falso?**

a. la media aritmetica può essere calcolata qualsiasi sia il tipo di carattere V F
b. la modalità che ha frequenza maggiore si chiama moda e può non essere unica V F
c. la media aritmetica di un insieme di numeri è sempre un numero positivo o nullo V F
d. per distribuzioni suddivise in classi non è possibile individuare il valore esatto della media, ma è possibile determinarne un valore approssimato V F
e. la classe mediana è la classe che ha frequenza relativa più elevata V F
f. media, moda e mediana non possono mai assumere lo stesso valore V F

[Solo 2 affermazioni sono vere]

20 Per tre giorni consecutivi in una certa località è stata rilevata la temperatura a intervalli di tempo regolari (alle ore 3, 9, 15, 21). Le temperature rilevate sono riportate nella tabella seguente. Per ciascuna giornata completa la tabella calcolando la temperatura media e mediana (la prima riga è compilata come esempio).

Giorno	Temperature rilevate (°C)	Media	Mediana
1	20 25 28 18	La media aritmetica delle temperature è: $\dfrac{20\,°C + 25\,°C + 28\,°C + 18\,°C}{4} =$ $= 22{,}75\,°C$	Le temperature disposte in ordine *crescente* sono: $18\,°C, \underbrace{20\,°C, 25\,°C}_{\text{valori centrali}}, 28\,°C$ La temperatura mediana è: $\dfrac{20\,°C + 25\,°C}{2} = 22{,}5\,°C$
2	20, 27, 30, 18	....	....
3	19, 24, 25, 17	....	....

Trova la media aritmetica dei seguenti gruppi di numeri.

21 $-2, 2, -4, 4, 5, -5$ [0]

22 $-3, -2, 4, 5, 6$ [2]

23 $-\frac{3}{5}, \frac{7}{5}, \frac{1}{5}, \frac{1}{2}$ $\left[\frac{3}{8}\right]$

24 $1, 5, 4, 4, 2, 6$ $\left[\frac{11}{3}\right]$

25 $11, 15, 14, 14, 12, 16$ $\left[\frac{41}{3}\right]$

26 $a, \quad \frac{2}{3}a, \quad -2+3a, \quad \frac{4}{3}, \quad -\frac{a}{3}$ $\left[\frac{13a-2}{15}\right]$

27 $\sqrt{2}, \quad 2\sqrt{2}-1, \quad \sqrt{32}+2, \quad -3, \quad 2(\sqrt{2}+1)$ $\left[\frac{9\sqrt{2}}{5}\right]$

Trova la mediana dei seguenti gruppi di numeri.

28 $-2, 2, -4, 4, 5, -5$ [0]

29 $-3, -2, 4, 5, 6$ [4]

30 $\frac{19}{15}, -\frac{3}{5}, \frac{7}{5}, \frac{1}{5}, \frac{1}{2}, \frac{5}{4}, \frac{13}{10}$ $\left[\frac{5}{4}\right]$

31 $0,\bar{3}; 0,3; 0,5; 0,\bar{5}$ $\left[\frac{5}{12}\right]$

32 **E se?** Sia $a > 0$. Determina la mediana dei seguenti numeri:

$$\frac{3}{2}a, \quad \frac{5}{3}a, \quad (1-\sqrt{2})a, \quad (2-\sqrt{2})a, \quad \frac{9}{7}a$$

▶ Cambierebbe la risposta, se fosse $a < 0$? $\left[\frac{9}{7}a;\text{ no}\right]$

Trova la moda dei seguenti gruppi di numeri.

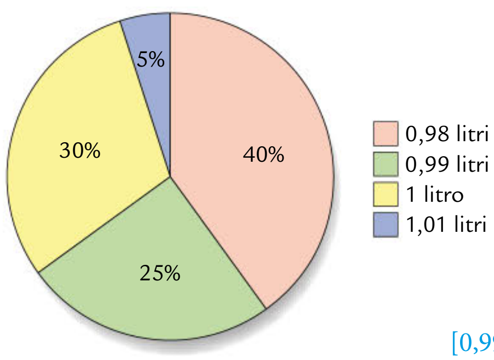
33 $2, 3, 4, 5, 5, 4, 1, 5$ [5]

34 $10, 15, 14, 14, 12, 15, 16$ [14 e 15]

35 $-1, 0, 0, 1, 3, 0, -1, 0$ [0]

36 $2, 4, 3, 6, 7, 1$ [Tutti i numeri dati]

37 Si è rilevata, su un campione di bottiglie da 1 litro di acqua minerale, la quantità di acqua effettivamente contenuta in esse. I risultati ottenuti sono riassunti nel diagramma a torta a lato. Calcola la quantità di acqua media contenuta in una bottiglia.



[0,99 litri]

38 **Videolezione** Un ristorante propone 10 diversi primi piatti, i cui prezzi presentano la distribuzione di frequenze a lato.

Prezzo (euro)	10	12	16	18
Frequenza	4	3	2	1

a. Qual è il prezzo medio di un primo?

b. Se il proprietario del ristorante decidesse di raddoppiare il prezzo di ciascun piatto, quale diventerebbe il prezzo medio di un primo?

c. Se il proprietario del ristorante decidesse di aumentare del 10% il prezzo di ciascun piatto, quale diventerebbe il prezzo medio di un primo? [**a.** 12 euro e 60 centesimi; **b.** 25 euro e 20 centesimi; **c.** 13 euro e 86 centesimi]

Interpretazione di grafici

39 Il grafico rappresenta la distribuzione di frequenza del numero di euro che lunedì scorso ciascuno dei 22 alunni di una classe aveva in tasca. Qual è il valore della mediana della distribuzione?

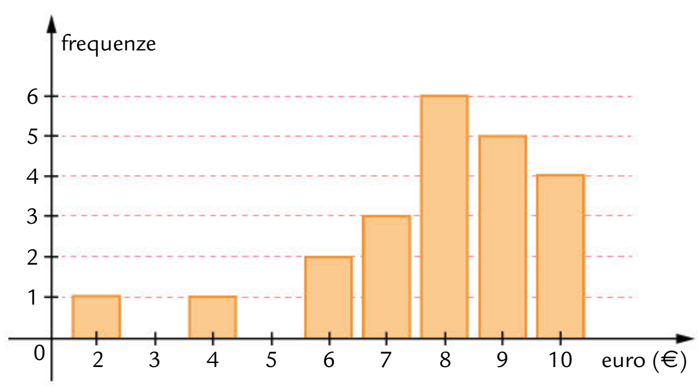
A 7

B 6

C 8

D 11,5

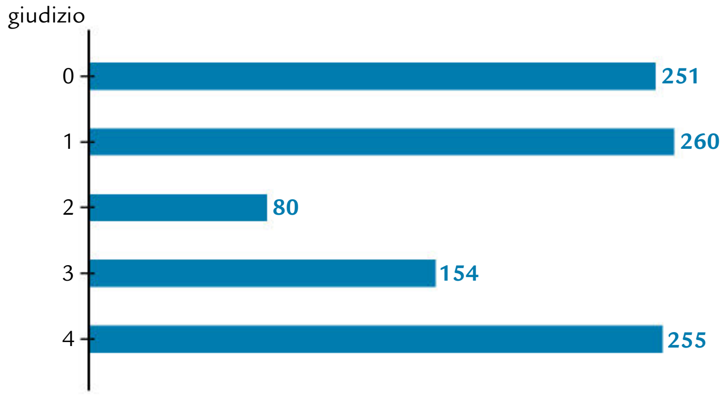
(Olimpiadi della statistica 2016)



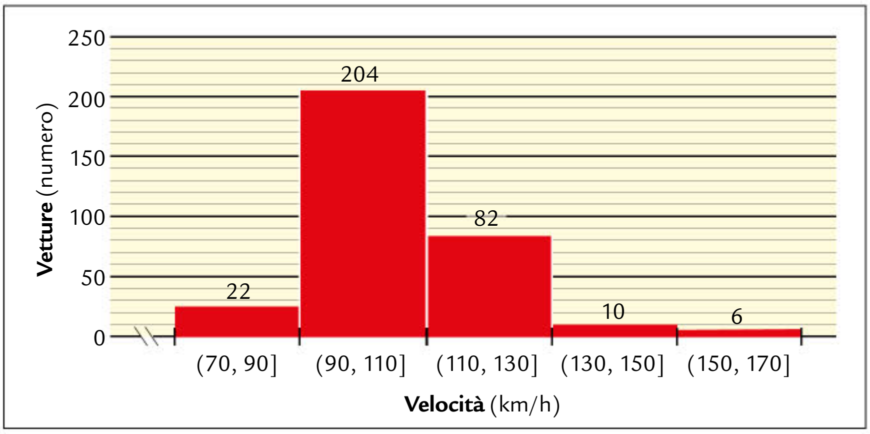
40 Ai 1000 abitanti di un piccolo comune viene chiesto di esprimere un giudizio su un nuovo servizio comunale, usando una scala da 0 a 4 (0 = pessimo, 4 = ottimo). Le risposte ottenute sono riassunte nel grafico.
Delle seguenti terne di valori, qual è la terna corretta relativa a moda, mediana e media aritmetica della distribuzione rappresentata nel suddetto grafico?

A 1; 1; 2,1 C 1; 2; 1,9
B 1; 2; 2,1 D 1; 1; 1,9

(Olimpiadi della statistica 2017)



41 Viene effettuato un controllo sulla velocità delle auto in un tratto autostradale in cui, a causa di alcuni lavori in corso, il limite consentito è di 110 km/h. Le frequenze delle velocità rilevate (raggruppate in classi) sono rappresentate nel seguente istogramma.



a. Qual è la velocità media degli automobilisti che sono stati oggetti del controllo (arrotonda il risultato a meno di una unità)?
b. Per uno sforamento del limite di velocità superiore ai 40 km/h è previsto il ritiro e la sospensione della patente. Qual è la percentuale di automobilisti che incorre in questa sanzione (arrotonda il risultato a meno di un centesimo)? [**a.** 106 km/h; **b.** 1,85%]

42 In una classe formata da 16 ragazze e 9 ragazzi, l'altezza media delle ragazze è 160 cm e l'altezza media dei ragazzi è 170 cm. Qual è l'altezza media degli studenti della classe? [163,6 cm]

43 I dipendenti di un'azienda sono per il 40% donne e per il 60% uomini. Il salario medio mensile delle donne è di 1400 euro, quello medio degli uomini è di 1600 euro. Qual è il salario medio mensile dei dipendenti dell'azienda? [1520 euro]

44 In un'azienda ci sono 10 dirigenti con un'anzianità media di 18,5 anni e 50 impiegati con un'anzianità media di 6,5 anni. Qual è l'anzianità media di tutto il personale dell'azienda? [8,5 anni]

45 I voti nell'ultimo compito di matematica in una classe sono stati i seguenti:

5	6	6	7	7,5	6,5	8	4	7	7	10	3	4,5	5,5	8	8	7	9	4	6	6,5	5

a. Rappresenta tramite una tabella la distribuzione delle frequenze dei voti del compito.
b. Determina la media, la mediana e la moda della distribuzione.
c. Nel giorno dello svolgimento del compito in classe c'era un assente, che ha recuperato il compito in classe la settimana successiva. Considerando anche il voto dell'assente, la media dei voti nel compito in classe è stata di 6,5. Quale voto ha avuto l'alunno assente nel compito?

$\left[\textbf{b.}\ \text{Media} = \frac{281}{44} \simeq 6{,}4;\ \text{mediana} = 6{,}5;\ \text{moda} = 7;\ \textbf{c.}\ 9\right]$

46 Si è lanciato 20 volte un dado. La seguente tabella riporta i risultati ottenuti.

Numero della faccia	1	2	3	4	5	6
Frequenza	6	2	4	3	4	1

Determina:

a. il valore medio della distribuzione; **b.** il valore mediano; **c.** la moda. [**a.** 3; **b.** 3; **c.** 1]

47 La ripartizione degli stipendi in un'azienda è rappresentata nella tabella qui sotto. Determina:

a. lo stipendio medio;
b. lo stipendio mediano;
c. la classe modale.

Stipendio S in euro	Frequenza
$1000 \leq S < 1400$	6
$1400 \leq S < 1800$	10
$1800 \leq S < 2200$	4

[**a.** 1560 euro; **b.** 1600 euro; **c.** $1400 \leq S < 1800$]

Varianza, deviazione standard e coefficiente di variazione

48 **Vero o falso?**

a. la varianza è la media dei quadrati degli scarti dei valori osservati dalla loro media aritmetica V F
b. la varianza è il quadrato della deviazione standard V F
c. la deviazione standard di n numeri è sempre minore della loro varianza V F
d. se $x_1, x_2, ..., x_n$ sono le n modalità osservate di un carattere quantitativo, la loro deviazione standard ha la stessa unità di misura delle modalità V F

[3 affermazioni vere e 1 falsa]

49 Se la deviazione standard di n numeri è nulla, che cosa si può dire degli n numeri? La deviazione standard di n numeri può essere uguale a 1?

50 Sono state rilevate le età di tre gruppi di ragazzi (indicati con A, B, C). Per ciascun gruppo, completa la tabella seguente, calcolando media, varianza e deviazione standard delle età del gruppo (la prima riga è compilata come esempio).

Gruppo	Età (in anni)	Media	Varianza	Deviazione standard	Coefficiente di variazione
A	12, 15, 21	$\frac{12+15+21}{3} = \underbrace{16}_{\text{età media}}$	$\underbrace{\frac{12^2+15^2+21^2}{3}}_{\text{media dei quadrati delle età}} - \underbrace{16^2}_{\text{quadrato dell'età media}} = \underbrace{14}_{\text{varianza}}$	$\underbrace{\sqrt{14}}_{\text{varianza}} \simeq \underbrace{3{,}7}_{\text{deviazione standard}}$	$\frac{\overbrace{\sqrt{14}}^{\text{deviazione standard}}}{\underbrace{16}_{\text{media}}} \simeq \underbrace{0{,}23}_{\text{coeff. di variazione}}$
B	9, 12, 18				
C	12, 15, 20, 21				

Trova la varianza, la deviazione standard e il coefficiente di variazione (se esiste) dei seguenti gruppi di numeri. Arrotonda i risultati alla seconda cifra decimale.

51 $-2, 2, -4, 4, 5, -5$ [$\sigma^2 = 15$, $\sigma \simeq 3{,}87$, $\nexists\, C_V$]

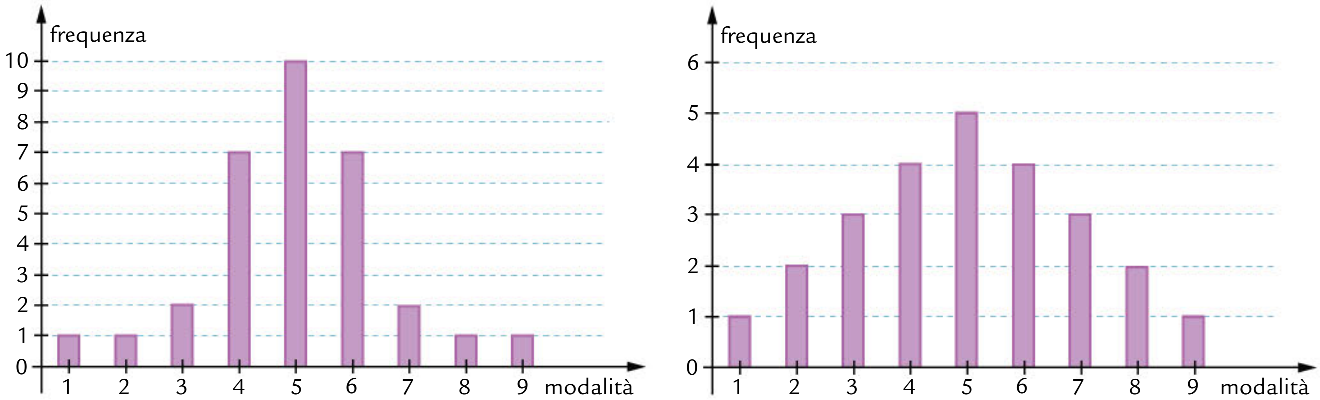
52 $-3, -2, 4, 5, 6$ [$\sigma^2 = 14$, $\sigma \simeq 3{,}74$, $C_V \simeq 1{,}87$]

53 $-\frac{3}{5}, \frac{7}{5}, \frac{1}{5}, \frac{1}{2}$ [$\sigma^2 = 0{,}51$, $\sigma \simeq 0{,}72$, $C_V \simeq 1{,}92$]

54 $1, 5, 4, 4, 2, 6$ [$\sigma^2 \simeq 2{,}89$, $\sigma \simeq 1{,}70$, $C_V \simeq 0{,}46$]

55 $11, 15, 14, 14, 12, 16$ [$\sigma^2 \simeq 2{,}89$, $\sigma \simeq 1{,}70$, $C_V \simeq 0{,}12$]

56 **Interpretazione di grafici** Fra le due distribuzioni di frequenze rappresentate in figura, qual è quella la cui deviazione standard risulta maggiore?



57 Determina per quale valore di a la varianza dei numeri a; $2a$; $3a$; $4a$, dove $a > 0$, è uguale a 45. [$a = 6$]

58 Sono dati i numeri: a, $2a$, $3a$, essendo $a > 0$. Determina a in modo che la varianza di questi numeri sia 24. [$a = 6$]

59 **Videolezione** Andrea ha ottenuto, in quattro compiti in classe, i seguenti voti: 5, 7, 6, 8.

a. Calcola la media e la deviazione standard dei voti di Andrea.

b. Se il professore di Andrea decidesse di aumentare i voti di tutti i compiti in classe di 1 punto, quali diventerebbero la media e la deviazione standard dei voti di Andrea? [**a.** Media $= 6{,}5$, $s = \frac{\sqrt{5}}{2} \simeq 1{,}2$; **b.** media $= 7{,}5$; s invariata]

60 Una piccola azienda effettua un'indagine circa il numero dei figli dei propri dipendenti. I risultati sono riassunti nella seguente tabella. Determina la media, la varianza e la deviazione standard del numero di figli dei dipendenti di quell'azienda.

Numero di figli	0	1	2	3	4
Numero di dipendenti	10	18	16	4	2

[Media $= 1{,}4$; $V = 1{,}04$; $s \simeq 1{,}02$]

61 **Zoom sull'enunciato** Alberto, Barbara, Carlo ricevono dai rispettivi padri «paghette» mensili di 50, 100, 150 euro. I padri percepiscono stipendi mensili di 1000, 1500, 2000 euro. Risulta maggiore la *variabilità* degli stipendi dei genitori o delle paghette dei figli?

[I coefficienti di variazione delle paghette e degli stipendi valgono all'incirca rispettivamente 0,41 e 0,27]

Guida all'interpretazione del testo

La *variabilità* di un insieme di dati si studia tramite la *deviazione standard*; in questo caso tuttavia è richiesto un *confronto* tra la variabilità di *due* insiemi di dati (stipendi e paghette) quindi occorre confrontare i *coefficienti di variazione*.

Perché non sarebbe stato significativo il confronto tra le deviazioni standard?

62 In una classe la statura media delle femmine è 160 cm e la deviazione standard è 3,9 cm, mentre la statura media dei maschi è 178 cm e la varianza è 16 cm^2. È vero che la statura dei maschi è più variabile di quella delle femmine?

[No; confronta i coefficienti di variazione]

Media armonica e media geometrica

63 ESERCIZIO SVOLTO

Determiniamo la media armonica e la media geometrica dei numeri: 3, 6, 12.

- In base alla definizione, la media armonica dei tre numeri dati è:

$$m_{\text{armonica}} = \frac{3}{\frac{1}{3}+\frac{1}{6}+\frac{1}{12}} = \frac{3}{\frac{7}{12}} = \frac{36}{7}$$

- In base alla definizione, la media geometrica dei tre numeri dati è:

$$m_{\text{geometrica}} = \sqrt[3]{3 \cdot 6 \cdot 12} = \sqrt[3]{216} = 6$$

Trova la media armonica dei seguenti gruppi di numeri.

64 $-2, 2, -4, 4, 5, -5$ [La media armonica non è definita]

65 $-\frac{3}{5}, \frac{7}{5}, \frac{1}{5}, \frac{1}{2}$ $\left[\frac{84}{127}\right]$

66 $2 \cdot 10^{-4}, \quad 4 \cdot 10^{-4}$ $\left[\frac{8}{3} \cdot 10^{-4}\right]$

67 $2^{-3}, \quad 2^3$ $\left[\frac{16}{65}\right]$

68 $a, 2a, 3a$ con $a \neq 0$ $\left[\frac{18}{11}a\right]$

Trova la media geometrica dei seguenti gruppi di numeri.

69 2, 9, 12 [6]

70 3, 4, 8, 9, 9 [6]

71 $2 \cdot 10^{-4}, \quad 4 \cdot 10^{-4}$ $[2\sqrt{2} \cdot 10^{-4}]$

72 $2 \cdot 10^{-4}, \quad 4 \cdot 10^{4}$ $[2\sqrt{2}]$

73 $2a^3, \quad 8a$ con $a > 0$ $[4a^2]$

Problemi su medie di vario tipo

74 Una macchina in 4 anni dimezza il suo valore commerciale. Qual è il tasso medio percentuale di deprezzamento annuo? [Circa 15,91%]

75 Uno studente, in vista di un esame, ha programmato di studiare 10 pagine di testo al giorno, per 60 giorni. Fatti i primi quindici giorni di studio, però, lo studente, a causa di imprevisti, deve sospendere lo studio per 6 giorni. Quale media giornaliera egli dovrà imporsi per completare la preparazione alla stessa data già stabilita all'inizio? [Circa 11,54 pagine/giorno]

76 Un ragazzo ha programmato di fare un risparmio medio giornaliero di un euro, fino ad arrivare ad acquistare un nuovo casco da 110 euro. Dopo aver effettuato la raccolta per venti giorni, la interrompe per una settimana, durante la quale utilizza cinque euro di quelli già accantonati. Quale media giornaliera di risparmio gli permetterà di arrivare all'acquisto alla stessa data che aveva previsto? [Circa 1,14 euro]

77 Il tasso di crescita dell'inflazione in un certo Paese è stato, in due anni successivi, del 4% e del 6%. Qual è il tasso medio annuo di aumento dell'inflazione in quel Paese? [Circa 4,995%]

78 Per arrivare sulla spiaggia di una rinomata località di mare della Liguria bisogna servirsi di una lunga scalinata. Se si scende di 2 gradini al secondo e, successivamente, si risale di 2 gradini ogni 3 secondi, per salire e scendere si impiega, complessivamente, un certo tempo t. Quanti gradini al secondo si devono percorrere sia in discesa sia in salita senza che vari il tempo totale t? [Un gradino al secondo]

79 Un prodotto finanziario ha dato negli ultimi cinque anni i seguenti rendimenti: 6%, 4%, 6%, 8%, 6%. Qual è stato il rendimento medio annuo nei cinque anni? [Circa 5,99%]

80 Una lega di metalli è costituita al 20% da ferro e all'80% da nichel. La densità del ferro è 7,8 g/cm^3, quella del nichel è 8,9 g/cm^3; con quale media è possibile calcolare la densità della lega? Qual è la densità della lega? [Media armonica ponderata delle singole densità; densità $\simeq$ 8,66 g/cm^3]

3. Rapporti statistici

Teoria p. 12

Esercizi introduttivi

81 Stabilisci il tipo di ciascuno dei seguenti rapporti.

a. Numero di autovetture in una città su numero di abitanti.
b. Numero di alunni di una classe promossi senza debito su numero di alunni con debito.
c. Numero di studenti di una scuola di età inferiore ai 15 anni su numero di studenti della scuola.
d. Spesa totale sostenuta da un'impresa per la produzione di un bene in un dato mese su numero di unità del bene prodotte.

82 **Vero o falso?**

a. un rapporto di composizione è sempre compreso tra 0 e 1 V F
b. un rapporto di coesistenza è sempre compreso tra 0 e 1 V F
c. un rapporto di densità è privo di unità di misura V F
d. un rapporto di composizione è privo di unità di misura V F
e. un rapporto di derivazione è sempre anche un rapporto di composizione V F

[2 affermazioni vere e 3 false]

83 **Vero o falso?**

a. il numero di morti in Italia osservato ogni mese dell'anno costituisce una serie storica V F
b. in una serie di numeri indice a base fissa, il numero indice relativo all'anno base è sempre uguale a 1 V F
c. una serie storica riporta il prezzo di un bene il primo gennaio di ogni anno, per tutti gli anni dal 2005 al 2015; se il numero indice a base fissa relativo al prezzo dell'anno 2010 rispetto al prezzo dell'anno base 2005 è uguale a 1,34 significa che il prezzo nel 2010, rispetto a quello del 2005, è aumentato del 134% V F
d. data una serie di numeri indici a base mobile, dividendo ogni numero indice con il numero indice precedente si ottiene una serie di numeri indice a base fissa V F
e. una serie storica riporta il prezzo di un bene il primo gennaio di ogni anno, per tutti gli anni dal 2005 al 2015; il numero indice a base fissa relativo al prezzo dell'anno 2015, rispetto al prezzo dell'anno base 2005, è sempre inferiore al numero indice a base fissa relativo al prezzo dell'anno 2015, rispetto al prezzo dell'anno base 2016 V F

[2 affermazioni vere e 3 false]

Problemi

84 Utilizzando i dati in tabella, calcola il tasso percentuale di occupazione e disoccupazione maschile, femminile e complessivo per l'Italia nel 2010. Ricorda che per tasso di occupazione si intende il rapporto tra gli occupati e la popolazione residente mentre per tasso di disoccupazione si intende il rapporto tra le persone in cerca di occupazione e le forze di lavoro. Arrotonda i risultati con una cifra decimale. Che tipo di rapporti hai costruito?

Forze di lavoro in Italia nell'anno 2010 (migliaia)				
Sesso	Occupati	In cerca di occupazione	Forze di lavoro (totale tra occupati e in cerca di occupazione)	Popolazione residente
Maschi	13 634	1114	14 748	29 181
Femmine	9238	989	10 227	30 871

[Tassi di occupazione: 46,7% (maschi), 29,9% (femmine), 38,1% (complessivo); tassi di disoccupazione: 7,6% (maschi), 9,7%(femmine), 8,4% (complessivo)]

85 In riferimento ai dati in tabella, determina i tassi di natalità e mortalità su mille abitanti negli anni 2008, 2009 e 2010. Fornisci i risultati arrotondati con una cifra decimale. Di che tipo di rapporti si tratta?

Anno	Popolazione (al 1° gennaio)	Nascite	Morti
2008	59 619 290	576 659	585 126
2009	60 045 068	568 857	591 663
2010	60 340 328	561 944	587 488

[2008: 9,7‰ (natalità) e 9,8‰ (mortalità); 2009: 9,5‰ (natalità) e 9,9‰ (mortalità); 2010: 9,3‰ (natalità) e 9,7‰ (mortalità)]

86 Nella seguente tabella è riportata la distribuzione percentuale delle età degli italiani nel 2012, per area geografica.

a. Calcola, per ogni area geografica, l'*indice di dipendenza*, ossia il rapporto percentuale tra la popolazione avente età inferiore o uguale ai 14 anni oppure superiore o uguale ai 65 anni e la popolazione avente età compresa tra 15 e 64 anni.

b. Calcola, per ogni area geografica, l'indice di dipendenza degli anziani, ossia il rapporto percentuale tra la popolazione con 65 o più anni e la popolazione con età compresa fra 15 e 64 anni.

	0-14 anni	15-64 anni	65 anni e oltre
Nord	13,7	64,6	21,7
Centro	13,4	64,8	21,8
Sud	14,9	66,7	18,4
Italia	14,1	65,3	20,6

c. I rapporti calcolati ai punti **a.** e **b.** sono rapporti di composizione? In caso negativo, di che tipo di rapporti si tratta? Le percentuali riportate in tabella possono considerarsi rapporti di composizione?

[**a.** Circa 54,80%, 54,32%, 49,93% (arrotondando alla seconda cifra decimale); **b.** circa 33,59%, 33,64%, 27,59% (arrotondando alla seconda cifra decimale)]

87 In un dato Paese le forze di lavoro, per titolo di studio, sono distribuite come indicato in tabella.

Titolo di studio	Forze lavoro nel 2010	Forze lavoro nel 2015
Licenza elementare	2152	2280
Licenza media	8244	9300
Diploma	10 052	12 160
Laurea di primo o secondo livello o dottorato	4625	5082

Costruisci degli opportuni rapporti statistici che ti permettano di confrontare le variazioni nella composizione delle forze di lavoro tra il 2010 e il 2015, rispondendo così alle seguenti domande.

a. La quota di forze lavoro con licenza elementare è aumentata o diminuita dal 2010 al 2015?
b. La quota di forze lavoro con licenza media è aumentata o diminuita dal 2010 al 2015?
c. La quota di forze lavoro con diploma è aumentata o diminuita dal 2010 al 2015?
d. La quota di forze lavoro con laurea è aumentata o diminuita dal 2010 al 2015?

Che tipo di rapporti hai costruito per rispondere alle domande precedenti?

[**a.** È diminuita, passando circa dall'8,58% al 7,91%; **b.** è diminuita, passando circa dal 32,88% al 32,27%; **c.** è aumentata, passando circa dal 40,1% al 42,19%; **d.** è diminuita, passando circa dal 18,45% al 17,63%]

88 Nella seguente tabella è riportato il numero di procedimenti civili di primo grado per ufficio giudiziario, negli anni 2005-2009. Definisci degli opportuni rapporti statistici che permettano di eseguire confronti e rispondere alle seguenti domande.

Anni	Uffici giudiziari		
	Giudice di pace	Tribunale	Corte di appello
2005	475 309	953 825	8256
2006	458 438	899 046	4632
2007	449 126	925 674	3007
2008	471 460	914 145	3083
2009	497 907	880 465	3006

a. La quota di procedimenti civili di primo grado svolti dai giudici di pace è aumentata o diminuita dal 2005 al 2006?
b. La quota di procedimenti civili di primo grado svolti nei tribunali è aumentata o diminuita dal 2006 al 2007?
c. La quota di procedimenti civili di primo grado svolti in corte di appello è aumentata o diminuita dal 2005 al 2006?

Che tipo di rapporti hai costruito per rispondere alle domande precedenti?

[**a.** È aumentata passando circa dal 33,1% al 33,7%; **b.** è aumentata passando circa dal 66% al 67,2%; **c.** è diminuita passando circa dallo 0,6% allo 0,3%]

89 Nella seguente tabella sono riportati i prezzi in euro di una unità di un dato bene dal 2013 al 2017.

Anno	2013	2014	2015	2016	2017
Prezzo in euro	11,80	12,30	13	12,60	13,50

a. Calcola la serie dei numeri indici a base fissa dei prezzi, assumendo come anno base il 2013. Arrotonda i risultati alla terza cifra decimale.
b. Deduci dai numeri indice determinati al punto precedente la variazione percentuale del prezzo tra il 2013 e il 2015 e la variazione percentuale tra il 2013 e il 2017.
c. Calcola la serie dei numeri indice a base mobile. Arrotonda i risultati alla terza cifra decimale.
d. Deduci dai numeri indice determinati al punto precedente, la variazione percentuale del prezzo tra il 2015 e il 2016 e la variazione percentuale di prezzo tra il 2016 e il 2017.

[**a.** 1; 1,042, 1,102, 1,068, 1,144; **b.** +10,2% circa, +14,4% circa;
c. 1,042, 1,057, 0,969, 1,071; **d.** −3,1% circa, +7,1% circa]

90 Considera la serie storica rappresentata in tabella, relativa al numero di nascite in Italia nel periodo 2005-2010.

Anno	Nascite
2005	554 022
2006	560 010
2007	563 933
2008	576 659
2009	568 857
2010	561 944

a. Calcola la serie dei numeri indice a base fissa, assumendo come anno base il 2005. Arrotonda i risultati alla terza cifra decimale.
b. Deduci dai numeri indice determinati al punto precedente la variazione percentuale del numero di nati tra il 2005 e il 2007 e la variazione percentuale tra il 2005 e il 2008.
c. Calcola la serie dei numeri indice a base mobile. Arrotonda i risultati alla terza cifra decimale.
d. Deduci dai numeri indice determinati al punto precedente la variazione percentuale del numero di nati tra il 2007 e il 2008 e la variazione percentuale del numero di nati tra il 2008 e il 2009.

[**a.** 1, 1,011, 1,018, 1,041, 1,027, 1,014; **b.** +1,8% e +4,1%; **c.** 1,011, 1,007, 1,023, 0,986, 0,988; **d.** +2,3%, −1,4%]

91 Nella seguente tabella è rappresentata una serie di numeri indice a base mobile. Determina la corrispondente serie dei numeri indice a base fissa, avente come anno base il 2010.

Anno	2010	2011	2012	2013	2014	2015
Numero indice a base mobile	–	1,4	1,2	1,6	1,5	1,8

[1, 1,4, 1,68, 2,688, 4,032, 7,2576]

4. Indicatori di efficacia, efficienza e qualità

Teoria p. 16

92 **Vero o falso?**
a. un ufficio si prefigge l'obiettivo di evadere almeno l'80% degli ordini che arrivano in giornata. Il rapporto tra il numero di ordini evasi in una giornata e il numero di ordini pervenuti è un indicatore di efficacia V F
b. in riferimento all'ufficio di cui al punto precedente, la valutazione del tempo impiegato per evadere 100 ordini è un indicatore di efficienza V F
c. un ufficio postale ha cinque sportelli e si prefigge di servire 500 utenti alla settimana. Il numero di utenti effettivamente serviti dall'ufficio in una settimana è un indicatore di efficienza V F
d. in riferimento all'ufficio postale di cui al punto precedente, il rapporto tra il numero di utenti serviti settimanalmente da uno sportello e il numero complessivo di utenti serviti è un indicatore di efficacia V F

[2 affermazioni vere e 2 false]

93 Un'azienda promuove i suoi prodotti anche per via telefonica. L'obiettivo dell'azienda è di ottenere il maggior numero possibile di ordini, così da incrementare i profitti. Alcuni dati relativi alla promozione telefonica nell'anno 2011 e al corrispondente tempo di evasione degli ordini sono riportati nella seguente tabella, suddivisi per trimestre.

Trimestri (anno 2011)	Numero di clienti contattati tramite una telefonata	Numero di clienti che hanno effettivamente eseguito un ordine	Costo complessivo delle telefonate per contattare i clienti (in euro)	Tempo complessivo necessario all'evasione degli ordini (in ore)
1° trimestre	12 000	3250	3200	525
2° trimestre	15 000	4215	3420	410
3° trimestre	9500	3312	1850	350
4° trimestre	16 000	5680	2120	364

Considera i seguenti tre indicatori I_1, I_2, I_3 e rispondi ai quesiti **a-b-c-d**.

I_1 = rapporto tra clienti che hanno eseguito un ordine e clienti contattati

I_2 = costo medio di una telefonata

I_3 = tempo medio per l'evasione di un ordine

a. Stabilisci se si tratta di indicatori di efficacia o efficienza.

b. Calcola, per ciascuno dei trimestri presi in considerazione, i tre indicatori I_1, I_2, I_3; esprimi I_1 in percentuale (arrotondando il risultato con una cifra decimale), I_2 in euro (arrotondando il costo ai centesimi) e I_3 in minuti (arrotondando a un numero intero). Supponi che ogni potenziale cliente sia stato contattato tramite una sola telefonata e che ogni cliente che ha immesso un ordine ne abbia effettuato uno solo.

c. Alla luce degli indicatori determinati al punto **b.**, commenta i risultati raggiunti dall'azienda in termini di efficacia ed efficienza nei quattro trimestri presi in considerazione.

d. Come indicatori di efficacia e di efficienza annuali, che indichiamo con $\overline{I_1}$, $\overline{I_2}$, $\overline{I_3}$, l'azienda decide di assumere la media dei corrispondenti indicatori trimestrali I_1, I_2, I_3. Calcola gli indicatori per l'anno 2011.

[**b.** 1° trimestre: $I_1 \simeq 27{,}1\%$, $I_2 \simeq 27$ centesimi, $I_3 \simeq 10$ minuti; 2° trimestre: $I_1 = 28{,}1\%$, $I_2 \simeq 23$ centesimi, $I_3 \simeq 6$ minuti; 3° trimestre: $I_1 \simeq 34{,}9\%$, $I_2 \simeq 19$ centesimi, $I_3 \simeq 6$ minuti; 4° trimestre: $I_1 = 35{,}5\%$, $I_2 \simeq 13$ centesimi, $I_3 \simeq 4$ minuti; **d.** $\overline{I_1} = 31{,}4\%$, $\overline{I_2} = 20{,}5$ centesimi, $\overline{I_3} = 6{,}5$ minuti]

94 Un'azienda produce un determinato bene in serie e il suo obiettivo è la vendita del maggior numero possibile di unità del bene. Alcuni dati relativi al processo produttivo negli anni 2010-2012 sono riportati in tabella.

Anno	Unità vendute	Unità prodotte	Spesa complessiva (in euro)	Ore di produzione complessive
2010	12 150	12 200	150 000	2250
2011	11 085	12 650	144 000	2150
2012	10 052	11 500	135 000	1740

Considera i seguenti quattro indicatori:

I_1 = unità invendute

I_2 = rapporto tra unità vendute e unità prodotte

I_3 = costo medio per unità

I_4 = tempo medio di produzione per unità

a. Stabilisci se si tratta di indicatori di efficacia o efficienza.

b. Calcola, per ciascuno dei tre anni presi in considerazione, i quattro indicatori I_1, I_2, I_3, I_4; esprimi I_2 in percentuale (arrotondando il risultato con una cifra decimale), I_3 in euro (arrotondando il costo ai centesimi) e I_4 in minuti (arrotondando a un numero intero).

c. Alla luce degli indicatori determinati al punto **b.**, commenta i risultati raggiunti dall'azienda in termini di efficacia ed efficienza nel triennio preso in considerazione.

[**b.** 2010: $I_1 = 50$, $I_2 \simeq 99{,}6\%$, $I_3 \simeq 12{,}30$ euro, $I_4 \simeq 11$ minuti; 2011: $I_1 = 1565$, $I_2 \simeq 87{,}6\%$, $I_3 \simeq 11{,}38$ euro, $I_4 \simeq 10$ minuti; 2012: $I_1 = 1448$, $I_2 \simeq 87{,}4\%$, $I_3 \simeq 11{,}74$ euro, $I_4 \simeq 9$ minuti]

95 Una società di consulenza informatica ha come obiettivo la soddisfazione del maggior numero possibile di clienti e il raggiungimento del massimo profitto. Alcuni dati relativi agli anni 2010-2012 sono riportati in tabella.

Anno	Numero di clienti	Numero di clienti soddisfatti	Costo complessivo per la gestione dei clienti (euro)	Fatturato annuo (euro)	Numero annuo complessivo di ore di lavoro per la gestione dei clienti
2010	125	115	182 000	285 000	4600
2011	180	150	265 000	360 000	4400
2012	165	158	226 000	320 000	4150

Considera i seguenti quattro indicatori:

I_1 = rapporto tra il numero di clienti insoddisfatti e il numero di clienti soddisfatti

I_2 = costo medio per la gestione di un cliente

I_3 = profitto medio per cliente

I_4 = tempo medio di gestione di un cliente

a. Stabilisci se si tratta di indicatori di efficacia o efficienza.

b. Calcola, per ciascuno dei tre anni presi in considerazione, i quattro indicatori I_1, I_2, I_3, I_4; esprimi I_1 in percentuale (arrotondando il risultato a una cifra decimale), I_2 e I_3 in euro (arrotondando ai centesimi) e I_4 in ore (arrotondando a un numero intero).

c. Alla luce degli indicatori determinati al punto **b.**, commenta i risultati raggiunti dall'azienda in termini di efficacia ed efficienza nel triennio preso in considerazione.

[**b.** 2010: $I_1 \simeq 8{,}7\%$, $I_2 = 1456$ euro, $I_3 = 824$, $I_4 \simeq 37$ ore; 2011: $I_1 = 20\%$, $I_2 \simeq 1472{,}22$ euro; $I_3 \simeq 527{,}78$ euro, $I_4 \simeq 24$ ore; 2012: $I_1 \simeq 4{,}4\%$, $I_2 = 1369{,}70$ euro, $I_3 = 569{,}7$ euro, $I_4 \simeq 25$ ore]

96 Un'azienda di trasporti gestisce il collegamento tra due città con gli autobus. L'orario degli autobus prevede un tempo di percorrenza di 1 h e 16 min e l'obiettivo dell'azienda è di rispettarlo; nelle ultime otto corse i due autisti che gestiscono la tratta hanno impiegato i tempi riportati in tabella.
L'azienda vuole confrontare la bontà del servizio offerto da parte di due autisti, utilizzando come indicatori di efficacia il ritardo medio accumulato dai due autisti nelle corse prese in considerazione e la variabilità nei tempi di percorrenza. Stabilisci quale dei due autisti è stato mediamente più puntuale e quale dei due ha percorso la tratta con tempi più variabili.

Autista A	Autista B
1 h 32 min	1 h 20 min
1 h 25 min	1 h 16 min
1 h 15 min	1 h 30 min
1 h 22 min	1 h 24 min
1 h 16 min	1 h 20 min
1 h 14 min	1 h 28 min
1 h 12 min	1 h 22 min
1 h 20 min	1 h 12 min

[Ritardo medio di A = 3 min e 30 s; ritardo medio di B = 5 min e 30 s; deviazione standard dei tempi di A ≃ 6 min e 15 s; deviazione standard dei tempi di B ≃ 5 min e 33 s; per confrontare la variabilità nei tempi di percorrenza dei due autisti devi calcolare i coefficienti di variazione, da cui risultata che l'autista che ha percorso la tratta con tempi più variabili è stato A]

97 La direzione di un call center in cui lavorano 10 operatori effettua un monitoraggio del servizio, valutando come indicatori di efficienza il numero medio e il numero mediano di telefonate gestite dagli operatori in una settimana, nonché il relativo coefficiente di variazione. La direzione ritiene che il livello di efficienza sia troppo basso in ciascuno dei seguenti casi: se il numero medio di telefonate gestite dagli operatori è inferiore a 150, oppure se il numero mediano di telefonate è inferiore a 200 oppure se il coefficiente di variazione è superiore al 30%. Se si verifica almeno una di queste tre situazioni, sono previste attività supplementari di formazione per gli operatori. Nell'ultima settimana il numero di telefonate gestite dagli operatori è risultato quello in tabella:

Operatore	1	2	3	4	5	6	7	8	9	10
Numero di chiamate	232	140	186	202	150	190	161	244	210	135

a. Determina il numero medio e il numero mediano di telefonate gestite.
b. Determina il coefficiente di variazione.
c. Stabilisci se l'azienda, in base ai parametri prefissati, deve provvedere a organizzare ulteriori attività di formazione per gli operatori. [**a.** Numero medio = 185, numero mediano = 188; **b.** circa 19,5%]

5. Distribuzione normale e introduzione all'inferenza

Teoria p. 18

Distribuzione normale

Test

98 La statura media di una popolazione è di 172 cm con una deviazione standard di 18 cm. Sapendo che la distribuzione delle stature è normale, in quale dei seguenti intervalli cade all'incirca il 95,45% delle stature della popolazione?
A 154 cm, 190 cm
B 136 cm, 208 cm
C 118 cm, 226 cm
D Nessuno dei precedenti

99 La statura media di una popolazione è di 172 cm con una deviazione standard di 18 cm. Sapendo che la distribuzione delle stature è normale, in quale dei seguenti intervalli cade all'incirca il 99,73% delle stature della popolazione?
A 154 cm, 190 cm
B 136 cm, 208 cm
C 118 cm, 226 cm
D Nessuno dei precedenti

100 Una partita di 2000 lampadine ha un tempo di vita medio di 12 000 ore. Se il tempo di vita delle lampadine è distribuito normalmente, con deviazione standard uguale a 500 ore, quante lampadine all'incirca hanno un tempo di vita compreso tra 11 500 ore e 12 500 ore?
A 1365
B 1909
C 1995
D Non è possibile stabilirlo con i dati a disposizione

101 Durante una fase di sperimentazione di un farmaco, si è misurata la concentrazione raggiunta dal farmaco nel sangue dopo 1 ora dall'assunzione su un campione di 150 individui. Si è ottenuta una distribuzione normale delle concentrazioni misurate, con concentrazione media di 4 μg/ml e una deviazione standard di 0,15 μg/ml. In quanti individui del campione la concentrazione del farmaco dopo 1 ora dalla somministrazione è risultata compresa tra 3,7 μg/ml e 4,3 μg/ml? [Circa 143]

102 Ai fini di un controllo di qualità, si misura il diametro di 250 pezzi meccanici prodotti. Le misure risultano distribuite normalmente, con una media di 5,4 cm e una deviazione standard di 0,2 cm. Quanti dei pezzi meccanici esaminati presentano un diametro compreso tra 5,2 cm e 5,6 cm? [Circa 171]

Intervalli di confidenza

103 **Caccia all'errore.** Pietro afferma che l'ampiezza dell'intervallo di confidenza è indipendente dalla media del campione e direttamente proporzionale alla deviazione standard; Rosa aggiunge che l'ampiezza dell'intervallo di confidenza è inversamente proporzionale alla numerosità del campione. Uno dei due è in errore. Chi e perché?

104 Una azienda produce dadi per bulloni, di diametro medio μ incognito. Il diametro medio di un lotto di 100 dadi è risultato di 2,5 cm, con deviazione standard di 0,1 cm. Determina un intervallo di confidenza per μ, al livello del 95%.
[2,48 cm$\leq \mu \leq$2,52 cm]

105 **E se?** In un campione di 500 persone di una data popolazione si è osservata una statura media di 175 cm, con deviazione standard di 5 cm. Determinare l'intervallo di confidenza al 90% della statura media h della popolazione.

▸ Come cambia l'intervallo di confidenza se si vuole un livello di confidenza del 99% anziché del 90%?
[174, 6 cm $\leq h \leq$ 175,4 cm; 174,4 cm $\leq h \leq$ 175,5 cm]

106 Un call center ha calcolato il tempo medio della durata delle telefonate su un campione di 150 telefonate. Il tempo medio delle telefonate nel campione è risultato di 6 min con deviazione standard di 2 min. Determina un intervallo di confidenza al 95% per la durata media μ delle telefonate che arrivano al call center. [5,6 min $\leq \mu \leq$ 6,4 min]

107 In un seggio elettorale sono state scrutinate 250 schede e si è ottenuto che in 60 di esse è espresso il voto per un dato partito. A un livello di confidenza del 95%, quale è la percentuale p di voti ottenuta dal partito? [18,7% $\leq p \leq$ 29,3%]

108 Ai fini di sperimentare un nuovo farmaco, 200 individui vengono sottoposti a una terapia con tale farmaco e 80 di essi guariscono dalla patologia per cui il farmaco è stato formulato. A un livello di confidenza del 95%, qual è la percentuale p di malati che il farmaco guarisce? [33,2% $\leq p \leq$ 46,8%]

109 In un campione di 150 articoli prodotti da una azienda, 10 sono risultati difettosi. A un livello di confidenza del 99% qual è la percentuale p di articoli difettosi? [1,4% $\leq p \leq$ 12%]

110 **E se?** Un'azienda che produce lampadine dichiara che la durata media di un certo tipo di lampadine è di 785 ore. Una verifica effettuata su un campione di 150 lampadine ha dato come esito una durata media di 780 ore, con una deviazione standard di 10 ore. A un livello di confidenza del 95%, può ritenersi veritiero quanto dichiarato dall'azienda?

▸ Cambierebbe la risposta a un livello di confidenza del 90%?
[No; no]

111 La durata media di una batteria è dichiarata di 600 ore. Una verifica effettuata su un campione di 100 batterie ha dato come esito una durata media di 590 ore, con una deviazione standard di 20 ore. A un livello di confidenza del 95%, può ritenersi veritiero quanto dichiarato dall'azienda? [Sì]

112 Una casa farmaceutica afferma che un antidolorifico di sua produzione impiega in media 22 minuti per agire. In base alle analisi effettuate su un campione di 50 pazienti, il beneficio è stato ottenuto in media in 23 minuti con una deviazione standard di 8 minuti. A un livello di confidenza del 99%, può ritenersi veritiero quanto dichiarato dall'azienda? [Sì]

Esercizi di riepilogo

Esercizi interattivi

113 **Vero o falso?**
Stabilisci se le seguenti proposizioni sono corrette: se lo sono, giustificale, altrimenti mostra che sono false tramite un controesempio.

a. se due sequenze di numeri hanno la stessa media, allora hanno anche la stessa mediana [V] [F]
b. se due sequenze di numeri hanno la stessa mediana, allora hanno anche la stessa media [V] [F]
c. esistono sequenze di numeri per cui la moda, la media e la mediana coincidono [V] [F]
d. la moda di una sequenza di numeri può non essere unica [V] [F]
e. sia la varianza sia la deviazione standard presentano la stessa unità di misura dei dati [V] [F]

Test

114 Un impiegato ha percepito per i primi 3 mesi dell'anno uno stipendio mensile di 1000 euro. Nei 9 mesi successivi lo stipendio mensile è aumentato di 400 euro. Qual è lo stipendio medio nell'anno di quell'impiegato?

[A] 1250 euro [C] 1350 euro
[B] 1300 euro [D] 1400 euro

115 La media aritmetica di 11 numeri è 4850. Se ciascuno degli undici numeri viene diminuito di 10, quanto diventa la loro media aritmetica?

[A] 4830 [B] 4840 [C] 4850
[D] I dati forniti non sono sufficienti per determinarla

116 Mario, Luigi e Giacomo pesano complessivamente 210 kg. Sapendo che Mario e Luigi pesano rispettivamente 3 kg in meno e 4 kg in più della media aritmetica fra i pesi di tutti e tre, quanto pesa Giacomo?

[A] 68 kg [B] 69 kg [C] 70 kg [D] 71 kg

117 I voti in matematica di 10 studenti in un compito in classe sono stati: 9, 8, 7, 6, 5, 7, 8, 8, 5, 4. Il voto minimo della metà degli studenti più bravi è:

[A] la media dei voti
[B] la moda dei voti
[C] la mediana dei voti
[D] nessuna delle altre risposte

118 Stabilisci quale delle seguenti affermazioni è vera.

[A] La media di un insieme di dati non può mai essere uguale a zero
[B] La mediana di un insieme di dati può essere uguale alla media aritmetica
[C] La deviazione standard di un insieme di dati non può mai essere uguale a zero
[D] La moda di un insieme di dati non può essere uguale alla mediana

119 La media dei voti ottenuti in un compito in classe è stata 6 e la mediana 5,5. Il professore decide di alzare tutti i voti di mezzo punto. Allora:

[A] la media resta invariata e la mediana aumenta di 0,5
[B] la media aumenta di 0,5 e la mediana resta invariata
[C] sia la media sia la mediana restano invariate
[D] sia la media sia la mediana aumentano di 0,5

120 Stabilisci quale delle seguenti affermazioni è *falsa*.
La *deviazione standard* di un insieme di dati:

[A] è sempre un numero non negativo
[B] può essere uguale a zero
[C] non cambia se a tutti i dati si aggiunge uno stesso numero
[D] non cambia se tutti i dati vengono moltiplicati per uno stesso numero

121 Al fine di svolgere un'indagine sull'utilizzo di Internet da parte dei propri studenti, una scuola media rileva il tempo di connessione giornaliero (in minuti) degli studenti delle sezioni A e B dell'ultimo anno. Sapendo che per le due sezioni il numero di studenti, la media e la varianza del tempo di connessione sono rispettivamente:

$n_A = 30$, $\mu_A = 20$, $\sigma_A^2 = 49$ e $n_B = 20$, $\mu_B = 30$, $\sigma_B^2 = 36$

stabilisci quale delle seguenti affermazioni è corretta.

[A] La media su entrambe le classi del tempo di connessione è 25 minuti
[B] La variabilità del tempo di connessione è la medesima per le due sezioni
[C] La media su entrambe le classi del tempo di connessione è di 24 minuti
[D] La variabilità del tempo di connessione è maggiore nella sezione B
[E] Nessuna delle altre affermazioni

(Olimpiadi della Statistica 2013)

122 In una classe si è rilevato il mezzo di trasporto con cui gli allievi si recano a scuola. I dati grezzi ottenuti sono:

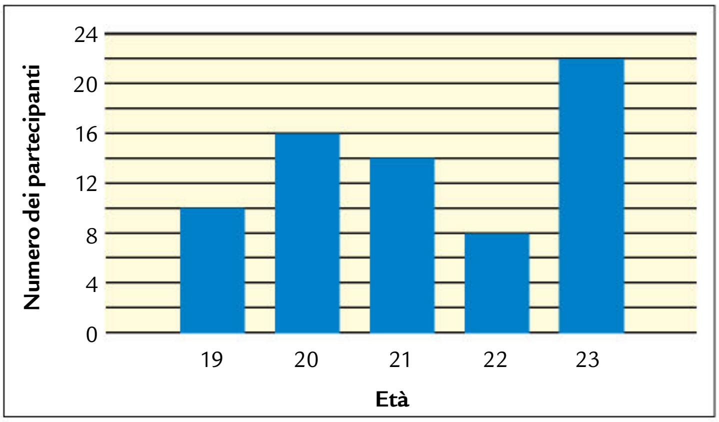
a piedi	in treno	in autobus	in auto	in auto	in treno	in autobus	a piedi	a piedi
in auto	in auto	a piedi	in treno	in autobus	in auto	a piedi	in auto	in autobus

a. Precisa il carattere studiato, le relative modalità, il collettivo e le unità statistiche.
b. Determina la tabella che rappresenta la distribuzione delle frequenze assolute, relative e percentuali.

123 Il diagramma a barre rappresenta la distribuzione di frequenze delle età dei giocatori in un torneo.

a. Qual è il numero complessivo dei partecipanti al torneo?
b. Qual è l'età media dei partecipanti?
c. Qual è l'età mediana?
d. Quale età rappresenta la moda della distribuzione?
e. Qual è la percentuale dei partecipanti aventi più di 20 anni?

[**a.** 70; **b.** $\frac{743}{35} \simeq 21{,}2$; **c.** 21; **d.** 23; **e.** circa il 62,86%]



124 Sono dati i numeri: a, $a + 1$, $a + 2$, $a + 4$, essendo $a \in \mathbf{R}$. Determina a in modo che la media di questa sequenza di numeri sia uguale al triplo della mediana. $\left[a = -\frac{11}{8}\right]$

125 **Giustificare e argomentare** Considera i numeri seguenti, dipendenti dal parametro reale k:

$$k + \frac{1}{2}, \qquad -2k - 1, \qquad 3k + 1, \qquad 2k + \frac{3}{2}$$

a. Calcola la loro media aritmetica, verificando che risulta indipendente da k.
b. Spiega perché, senza ulteriori informazioni, non è possibile calcolare la loro mediana.
c. Calcola la loro mediana nel caso in cui sia $k > \frac{1}{2}$.
d. La formula che esprime la mediana per $k > \frac{1}{2}$ non è più valida per $k = 0$. Verificalo e spiega perché ciò accade.

126 Si svolge un'indagine sull'importo della «paghetta» che i ragazzi di 16 anni ricevono dai genitori. Il 75% degli intervistati riceve una media di 100 euro al mese; il 15% una media mensile di 80 euro e il restante 10% una media di 160 euro. Qual è la paghetta media mensile dei ragazzi nel campione esaminato? [103 euro]

127 Si è rilevato su un campione di persone il numero di riviste acquistate in una data settimana. La tabella seguente rappresenta la distribuzione delle frequenze ottenuta:

Numero di riviste	0	1	2	3	4
Frequenza assoluta	12	13	15	6	4

Determina il numero medio e il numero mediano di riviste acquistate. [Media = 1,54; mediana = 1,5]

128 La seguente tabella riporta (per classi di età) il numero di persone seguite dalla ASL di Pavia, nel 2002, per tossicodipendenza (*fonte*: annuario statistico regionale).

Classi di età	[0, 14]	[15, 19]	[20, 24]	[25, 29]	[30, 34]	[35, 39]	Oltre 39
Frequenza	4	25	133	268	450	421	426

Trova l'età media, l'età mediana e la classe modale della distribuzione rappresentata nella tabella, chiudendo l'ultima classe a 50. [Media $\simeq$ 34,5; mediana = 32; classe modale = 30-34]

129 Un grande magazzino ha rilevato le taglie di pantaloni vendute in una settimana. La seguente tabella riporta la distribuzione delle frequenze relative dei dati raccolti.

Taglia	44	46	48	50	52	54
Frequenza relativa	0,08	0,15	0,20	.....	0,15	0,09

a. Qual è il carattere oggetto di questa indagine statistica? Quali modalità può assumere il carattere?
b. Qual è la frequenza relativa della taglia 50? Quale taglia rappresenta la moda della distribuzione?

Supposto che nella settimana in cui sono stati rilevati i dati siano stati venduti complessivamente 200 pantaloni, rispondi ai seguenti ulteriori quesiti:
c. Determina quanti pantaloni della taglia 50 sono stati venduti.
d. Calcola la taglia media e la taglia mediana.
e. Calcola la deviazione standard delle taglie dei pantaloni. Esprimi il risultato arrotondato alla seconda cifra decimale.
f. Calcola il coefficiente di variazione delle taglie dei pantaloni.

[**b.** 0,33, taglia modale = 50; **c.** 66; **d.** taglia media = 49,18, taglia mediana = 50; **e.** 2,73; **f.** 0,056]

130 La prova d'esame di un concorso cui hanno partecipato 50 candidati è costituita da 8 esercizi. La seguente tabella riporta i dati relativi a quanti esercizi ha svolto ciascun candidato.

Numero di esercizi svolti	1	2	3	4	5	6	7	8
Frequenza	2	1	3	20	12	3	5	4

a. Qual è il carattere oggetto di questa indagine statistica? Quali modalità può assumere il carattere?
b. Costruisci la tabella delle frequenze percentuali e rappresentala con il grafico che ritieni più opportuno.
c. Completa la tabella delle frequenze percentuali, calcolando anche le frequenze percentuali cumulate.
d. Quanti sono, in percentuale, i candidati che hanno svolto un numero di esercizi minore o uguale a 4?
e. Quanti sono, in percentuale, i candidati che hanno svolto più di 6 esercizi?
f. Calcola il numero medio e il numero mediano di esercizi svolti da un candidato.

[**d.** 52%; **e.** 18%; **f.** numero medio = 4,76, numero mediano = 4]

131 Si è rilevato il tasso di colesterolo nel sangue di 100 persone e si sono ottenuti i seguenti risultati:

Tasso colesterolo (in mg/dl)	[80, 120)	[120, 160)	[160, 200)	[200, 240)	[240, 280)	[280, 320)	[320, 360)
Frequenza	2	18	40	24	12	3	1

a. Rappresenta la distribuzione delle frequenze con un istogramma.
b. Determina la percentuale del campione avente colesterolo inferiore a 240 e la percentuale del campione avente colesterolo maggiore o uguale a 200.
c. Determina la mediana e la media.
d. Determina la varianza e la deviazione standard.
e. Determina il coefficiente di variazione.

[**b.** 84%; **c.** mediana = 180, media = 195,6; **d.** $V = 2012{,}64$, $s \simeq 44{,}86$; **e.** 0,23 (circa)]

132 Cinque tuffatori A, B, C, D, E eseguono una serie di quattro tuffi ciascuno, con diverso grado di difficoltà, ottenendo dalla giuria i seguenti punteggi:

	A	B	C	D	E
1° tuffo	5,50	8	6	7	7
2° tuffo	6	7,90	5,25	7,25	6
3° tuffo	6,50	6,75	7	7,25	5,50
4° tuffo	7	7,50	8,50	7	7,25

Dato che i tuffi hanno diversi gradi di difficoltà, a ognuno di essi è stato assegnato un peso diverso: il primo tuffo ha peso 1, il secondo ha peso 2, il terzo ha peso 3 e il quarto ha peso 2,5. Calcola la media aritmetica ponderata dei punteggi di ciascun tuffatore, stilando la classifica finale.

[Nell'ordine, i punteggi medi sono: $B \simeq 7{,}39$; $D \simeq 7{,}15$; $C \simeq 6{,}91$; $A \simeq 6{,}41$; $E \simeq 6{,}31$]

133 In una scuola con diversi indirizzi operano tre commissioni d'esame. I candidati della prima commissione sono i $\frac{3}{4}$ dei candidati della seconda e hanno avuto un voto medio di 72; i candidati della seconda commissione, che sono i $\frac{5}{3}$ dei candidati della terza, hanno avuto un voto medio di 76, e gli allievi della terza commissione hanno avuto un voto medio di 70. Qual è il voto medio degli alunni di tutta la scuola? [Circa 73,19]

134 In una classe di 25 alunni ci sono due alunni assenti nel giorno della correzione e riconsegna del compito in classe di matematica. L'insegnante non dice il voto degli assenti, ma dice che sia il voto medio dei 23 presenti, sia il voto medio di tutti e 25 gli alunni è 6,5 e che, dei due assenti, uno ha preso tre punti in più dell'altro. Quali sono i voti dei due assenti? [5 e 8]

135 Fra i partecipanti a un viaggio, l'età media delle donne è inferiore di 3 a quella degli uomini. Sapendo che gli uomini sono $\frac{1}{3}$ delle donne e che la media delle età di tutti i partecipanti è 43 anni, determina l'età media delle donne e l'età media degli uomini. [Età media degli uomini = 45,25; età media delle donne = 42,25]

136 Un gruppo assicurativo ha 10 agenti. In un mese otto agenti hanno stipulato 6 polizze ognuno, il nono agente ha stipulato 5 polizze e il decimo agente ha stipulato una polizza in più del numero medio di polizze che un agente del gruppo ha stipulato in quel mese. Qual è il numero complessivo di polizze stipulate in quel mese da tutti gli agenti?

137 Considerando i dati in tabella, calcola, per l'anno scolastico 2008/2009:

a. il tasso percentuale di iscritti alle scuole non statali rispetto al numero complessivo di iscritti;

b. il tasso percentuale di iscritti di sesso femminile rispetto al numero complessivo di iscritti;

c. il tasso percentuale di ripetenti rispetto al numero complessivo di iscritti e il tasso di ripetenti maschi rispetto al numero complessivo di ripetenti;

d. la densità degli studenti per scuola e per classe.

Arrotonda i risultati dei punti **a.**, **b.**, **c.** con una cifra decimale e i risultati del punto **d.** a un numero intero. Di che tipo di rapporti si tratta?

Anno scolastico	Scuole	Classi	Totale numero iscritti	di cui maschi	Numero studenti iscritti in scuole statali	Numero studenti ripetenti	Numero femmine ripetenti su 100 femmine iscritte
2008/09	6809	130 784	2 723 562	1 389 562	2 568 319	209 714	5,8

[**a.** 5,7%; **b.** 49%; **c.** 7,7%, 63,1%; **d.** 400 studenti/scuola, 21 studenti/classe]

138 Nella seguente tabella è riportata, *in parte*, la serie dei numeri indice a base mobile (espressi in percentuale) dei prezzi in euro di una rivista mensile negli anni 2005-2010.

Anno	2005	2006	2007	2008	2009	2010
Prezzo in euro	4	4,2	4,3	.....	.....	.....
Numero indice a base mobile	–	.....	.....	110%	115%	120%

a. Completa la serie dei prezzi della rivista negli anni mancanti, arrotondando il risultato ai centesimi.

b. Completa la serie dei numeri indice a base mobile, sempre esprimendoli in percentuale e arrotondandoli a un numero intero.

c. Determina la serie dei numeri indice a base fissa, assumendo come anno base il 2005 ed esprimendoli in percentuale.

[**a.** 4,73 euro, 5,44 euro, 6,53 euro; **b.** 105%, 102%; **c.** 105%, 107,5%, 118,25%, 136%, 163,25%]

139 **Inventa tu.** Un'azienda produce delle macchine per caffè espresso. L'obiettivo è di avere il maggior numero possibile di clienti, e la loro soddisfazione.
Alcuni dati relativi al numero di macchine per caffè vendute, alla soddisfazione dei clienti e ai costi di produzione negli anni 2015-2017 sono riportati nella seguente tabella.

Anno	Numero di macchine per caffè vendute	Numero di clienti soddisfatti (ogni cliente ha acquistato esattamente una macchina per caffè)	Costo complessivo per la produzione della macchine per caffè vendute (euro)
2015	15 500	9620	620 000
2016	13 200	10 150	594 000
2017	16 400	12 670	607 200

Definisci degli indicatori di efficacia ed efficienza, calcolane il valore e, alla luce di essi, commenta i risultati ottenuti dall'azienda in termini di efficacia ed efficienza nel triennio preso in esame.

Esercizi *più*

140 **Il paradosso dei voti.** In due prime classi di una stessa scuola è stato proposto lo stesso compito in classe di matematica. Le due classi sono composte entrambe da 24 alunni, di cui 8 ragazzi e 16 ragazze nella sezione A, 16 ragazzi e 8 ragazze nella sezione B. I voti riportati dagli allievi sono riportati nelle seguenti tabelle.

<table>
<tr><th colspan="17">Classe 1A</th></tr>
<tr><td>Ragazzi</td><td colspan="2">3</td><td colspan="2">4</td><td colspan="2">4</td><td colspan="2">5</td><td colspan="2">3</td><td colspan="2">3</td><td colspan="2">6</td><td colspan="2">4</td></tr>
<tr><td>Ragazze</td><td>7</td><td>8</td><td>6</td><td>9</td><td>7,5</td><td>7</td><td>7</td><td>8</td><td>8</td><td>6</td><td>6</td><td>7,5</td><td>9</td><td>9</td><td>7</td><td>8</td></tr>
</table>

<table>
<tr><th colspan="17">Classe 1B</th></tr>
<tr><td>Ragazzi</td><td>4</td><td>5</td><td>5</td><td>6</td><td>4</td><td>4</td><td>7</td><td>5</td><td>5</td><td>6</td><td>4</td><td>5</td><td>5</td><td>7</td><td>4</td><td>5</td></tr>
<tr><td>Ragazze</td><td colspan="2">8</td><td colspan="2">9</td><td colspan="2">7</td><td colspan="2">10</td><td colspan="2">8,5</td><td colspan="2">8</td><td colspan="2">9</td><td colspan="2">8,5</td></tr>
</table>

a. L'insegnante della sezione A afferma: «I miei alunni hanno, in media, ottenuto voti migliori degli alunni della sezione B».
b. L'insegnante della sezione B afferma: «Sia il gruppo dei ragazzi, sia il gruppo delle ragazze della mia classe hanno, in media, ottenuto voti migliori rispetto ai voti del gruppo dei ragazzi e del gruppo delle ragazze della sezione A».

Come si spiega questo apparente paradosso? Avrebbe potuto verificarsi un paradosso analogo se nelle due classi (sempre supposte ugualmente numerose) il numero dei ragazzi della sezione A fosse stato uguale al numero dei ragazzi della sezione B?

Dalle gare

141 Mariolina invita i visitatori del suo blog a dare i loro suggerimenti e a esprimere con un voto il gradimento del sito. La media attuale dei voti dei visitatori è visibile sul sito. Giuliano, cui piace molto il sito della sua amica, decide di dare come voto la media visibile sul sito più 1. In seguito al voto di Giuliano, la media dei voti del sito aumenta di 0,02 punti. Quanti visitatori del sito hanno votato prima di Giuliano sul blog di Mariolina?

(Mathématiques sans frontières, febbraio 2008) [49]

142 Michele si prepara all'ultimo compito in classe di matematica dell'anno; lo affronta con tranquillità, sapendo che se prenderà 10 avrà la media del 9, mentre prendendo 5 la media diverrà 8. Quanti compiti ha già fatto quest'anno Michele?

A 2 B 3 C 4 D 5 E I dati non sono sufficienti per dare la risposta

(Giochi di Archimede 2004) [C]

143 Andrea, Beatrice, Chiara, Davide, Enea e Federico sono molto amici. La loro età media è di 16 anni. Se a loro si uniscono tre amici di Enea, l'età media dell'intero gruppo diventa di 18 anni. Qual è l'età media dei tre amici di Enea?

A 18 B 19 C 22 D 21 E 20

(Giochi di Archimede, 2015)

144 I seguenti dati rappresentano il grasso totale contenuto negli hamburger e nel pollo preparati da un campione di catene di fast food:

Hamburger

19	31	34	35	39	39	43

Pollo

7	9	15	16	16	18	22	25	27	33	39

Quale delle seguenti affermazioni è corretta?

A La media aritmetica e la mediana del grasso contenuto negli hamburger sono entrambe maggiori rispettivamente della media aritmetica e della mediana del grasso contenuto nel pollo

B La mediana del grasso contenuto negli hamburger è maggiore di quella del pollo, mentre la media aritmetica del grasso contenuto negli hamburger è minore di quella del pollo

C La mediana del grasso contenuto negli hamburger è minore di quella del pollo, mentre la media aritmetica del grasso contenuto negli hamburger è maggiore di quella del pollo

D La media aritmetica e la mediana del grasso contenuto negli hamburger sono entrambe minori rispettivamente della media aritmetica e della mediana del grasso contenuto nel pollo

(Olimpiadi della statistica 2013)

145 Sull'isola *Kenoncè* tutti gli studenti che quest'anno hanno terminato la scuola superiore hanno partecipato al test per l'ammissione all'università. Considerando l'intera popolazione scolastica, il punteggio medio al test è stato 11. Se però ci si restringe agli studenti che hanno partecipato alle gare di matematica, che sono il 25% del totale, il punteggio medio al test è stato molto più alto: ben 18 punti. Allora, indicando con p il punteggio medio degli studenti che NON hanno partecipato alle gare di matematica, si ha:

A $8 \leq p < 9$ C $9 \leq p < 10$ E $6 \leq p < 7$

B $10 \leq p < 11$ D $7 \leq p < 8$ F $p < 6$

(Gara Nazionale Classi prime, 2017)

Prova di autoverifica

ESERCIZI

Richiami e complementi di statistica in una variabile

1 Specifica il tipo di rapporto statistico: di composizione, di derivazione, di coesistenza, di densità.

a. Il rapporto tra l'incasso di un esercizio commerciale e il numero degli scontrini emessi.
b. Il rapporto tra il numero dei telefoni cellulari in Italia e la popolazione italiana.
c. Il rapporto tra il numero di studenti maggiorenni e quello di studenti minorenni nella tua scuola.
d. Il rapporto tra il numero di laureati di sesso femminile e il numero complessivo di laureati.
e. Il rapporto tra la quantità di vino prodotto in Emilia-Romagna e la superficie della regione coltivata a viti.
f. Il rapporto tra il numero di autovetture di produzione italiana e il numero di autovetture di produzione straniera circolanti in Italia nell'anno 2016.

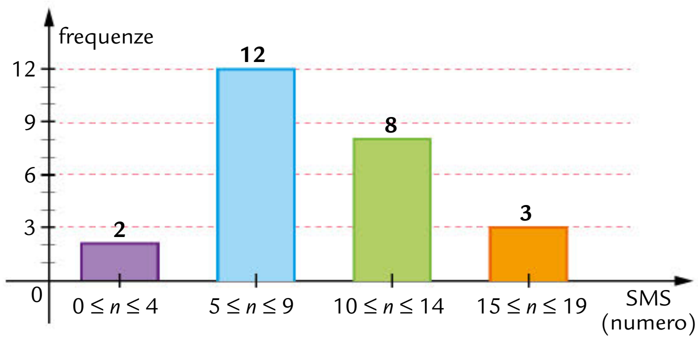
2 La tabella riporta la serie dei numeri indice a base fissa e a base mobile (espressi in percentuale) del numero degli studenti nell'istituto di Francesca. Completa la tabella, tenendo presente che l'anno di riferimento per gli indici a base fissa è il primo della rilevazione, cioè il 2012.

Anno	2012	2013	2014	2015	2016	2017
Numero studenti	800			840		
Indice a base fissa		105%				130%
Indice a base mobile			95%		110%	

3 Si è registrato il numero n di SMS ricevuti in un giorno dagli studenti di una classe e si sono ottenuti i dati riportati nel diagramma in figura.

Determina:

a. il numero medio di SMS ricevuti
b. il numero mediano di SMS ricevuti;
c. la deviazione standard del numero di SMS ricevuti.



4 La media dei voti di un compito in classe è stata 6,5, la mediana 6 e la deviazione standard 1,5.

a. Se il professore decide di aumentare tutti i voti di 1 punto, come variano la media, la mediana e la deviazione standard dei voti del compito?
b. Se il professore decide di aumentare tutti i voti del 10%, come variano la media, la mediana e la deviazione standard dei voti del compito?

5 Si vuole determinare il peso medio degli adulti di sesso maschile. Nel campione esaminato, costituito da 200 persone, si è osservata una media di 82 kg e una deviazione standard di 5 kg.

a. A un livello di confidenza del 95%, in quale intervallo è compreso il peso medio dell'intera popolazione adulta di sesso maschile?
b. A parità di media, deviazione standard e livello di confidenza, come cambierebbe la risposta data in precedenza se il campione fosse costituito da 500 persone?

6 Ai fini di valutare il gradimento di un prodotto alimentare appena posto sul mercato, sono state interpellate 100 persone, delle quali 76 hanno formulato un giudizio pienamente positivo.

a. A un livello di confidenza del 90%, in quale intervallo è compresa la percentuale dei consumatori soddisfatti?
b. E a un livello di confidenza del 99%?

Valutazione

Esercizio	1	2	3	4	5	6	**Totale**
Punteggio massimo	$0,25 \cdot 6 = 1,5$	2	$0,75 + 0,75 + 1 = 2,5$	$0,5 \cdot 2 = 1$	$0,75 \cdot 2 = 1,5$	$0,75 \cdot 2 = 1,5$	10
Punteggio ottenuto							

Tempo indicativo: 1 h

Risposte p. 218

Statistica bivariata, correlazione e regressione

Unità 2

Tema E

1. Tabelle a doppia entrata

Nell'**Unità 1** abbiamo appreso le nozioni fondamentali relative alla statistica **univariata**, ossia a quella parte della statistica che si occupa dell'analisi dei dati provenienti dalla rilevazione di *un solo* carattere su una data popolazione. In questa Unità vogliamo invece introdurre le nozioni di base relative alla statistica **bivariata**: vogliamo cioè vedere come si estendono le nozioni della statistica univariata quando vengono rilevati congiuntamente *due* caratteri, diciamo X e Y. L'obiettivo ulteriore che ci porremo, in questo nuovo contesto, sarà quello di scoprire e mettere in luce eventuali *relazioni* tra X e Y.

- **Con GeoGebra**
- **Esercizi interattivi**

Distribuzioni congiunte e marginali

Supponiamo dunque che due caratteri X e Y siano osservati (insieme) su ciascuna delle n unità che compongono la popolazione in esame.
Il risultato della rilevazione è un insieme di coppie ordinate (x, y) che possono essere rappresentate in una tabella come quella qui sotto, detta **tabella dei dati grezzi**:

Unità statistiche	Modalità di X rilevata	Modalità di Y rilevata
1	x_1	y_1
2	x_2	y_2
.....	.....	.....
n	x_n	y_n

ESEMPIO Tabella dei dati grezzi

Su una popolazione formata da cinque amici si sono rilevati due caratteri: l'età (X) e la città di nascita (Y). I dati grezzi possono essere organizzati nella tabella sottostante.

Nome	Età	Città di nascita
Alberto	30	Milano
Maria	35	Torino
Giovanni	32	Milano
Paola	30	Milano
Alessandro	32	Roma

Nell'ambito della statistica univariata abbiamo visto che per rendere i dati grezzi meglio leggibili è utile costruire la tabella che ne rappresenta la distribuzione di frequenze. Similmente si procede nell'ambito della statistica bivariata, costruendo una **tabella a doppia entrata**, che riporti le frequenze con cui si manifestano le varie *coppie* di modalità osservate.
Le frequenze assolute di queste coppie sono dette **frequenze congiunte** e la tabella così costruita viene detta **distribuzione doppia di frequenze**.

ALTRE NOTAZIONI
La frequenza congiunta della coppia (x_i, y_j) viene spesso indicata anche con il simbolo f_{ij}.

ESEMPIO Distribuzione doppia di frequenze

In riferimento all'esempio precedente, il carattere X (*età*) manifesta tre modalità distinte: 30, 32, 35; così pure il carattere Y (*città di nascita*) manifesta tre modalità distinte: Milano, Torino, Roma. I dati grezzi possono essere organizzati nella seguente tabella a doppia entrata, dove nella prima colonna sono riportate le modalità distinte di X, nella prima riga le modalità distinte di Y e all'incrocio tra una riga e una colonna è scritta la frequenza congiunta della coppia corrispondente.

Per esempio, la coppia (30, Milano) si presenta due volte, quindi abbiamo posto un 2 all'incrocio della riga e della colonna corrispondenti; invece la coppia (30, Torino) non si presenta mai, quindi abbiamo posto uno 0 all'incrocio della riga e della colonna corrispondenti. Le frequenze congiunte delle coppie (30, Milano) e (30, Torino) sono dunque rispettivamente 2 e 0.

X \ Y	Milano	Torino	Roma
30	2	0	0
32	1	0	1
35	0	1	0

La tabella a doppia entrata che rappresenta le frequenze congiunte di X e Y si completa di solito con un'ultima riga, dove vengono riportate le *somme* delle frequenze di ciascuna colonna, e un'ultima colonna, dove vengono riportate le *somme* delle frequenze di ciascuna riga. L'ultima riga e l'ultima colonna rappresentano le cosiddette **distribuzioni marginali** dei due caratteri, cioè le distribuzioni di X e Y che si avrebbero se ciascuno di essi fosse stato rilevato *singolarmente*.
All'incrocio dell'ultima riga e dell'ultima colonna si pone il numero complessivo di unità della popolazione.

ESEMPIO Distribuzioni marginali

In riferimento alla tabella dell'esempio precedente, ne risultano le distribuzioni marginali messe in evidenza sull'ultima riga e sull'ultima colonna:

OSSERVA
Il numero complessivo di unità del collettivo è uguale sia alla somma delle frequenze marginali di X, sia alla somma delle frequenze marginali di Y.

X \ Y	Milano	Torino	Roma	Totale
30	2	0	0	2
32	1	0	1	2
35	0	1	0	1
Totale	3	1	1	5

Distribuzione marginale di X

Distribuzione marginale di Y

Numero complessivo

Distribuzioni condizionate

Facciamo ancora riferimento alla tabella dell'ultimo esempio. Se per esempio fissiamo l'attenzione sulla riga corrispondente alla prima modalità di X, $x_1 = 30$, leggiamo come si distribuisce il carattere Y *tra le unità della popolazione che manifestano la modalità* x_1 *di* X. Per questo motivo si dice che tale riga rappresenta **la distribuzione condizionata** di Y rispetto alla modalità x_1 di X (*vedi* la tabella a pagina seguente).

Distribuzione condizionata di Y rispetto alla modalità $x_1 = 30$ di X

X \ Y	Milano	Torino	Roma	Totale
30	2	0	0	2
32	1	0	1	2
35	0	1	0	1
Totale	3	1	1	5

IN SIMBOLI
La distribuzione di X condizionata a una modalità y_j di Y si indica con il simbolo $X|y_j$.
Analogamente va interpretato il simbolo $Y|x_i$.

Più in generale, fissare l'attenzione su una sola riga (o colonna) della tabella (escludendo quelle delle modalità e dei totali) significa restringersi alla sottopopolazione che presenta una data modalità di X (o di Y): ciascuna di queste righe o colonne, singolarmente presa, rappresenta perciò una particolare **distribuzione condizionata**. È anche possibile costruire le **distribuzioni condizionate relative**, ponendo a rapporto le frequenze congiunte che appartengono alla distribuzione condizionata considerata con i corrispondenti totali di riga o di colonna. Per esempio, in riferimento alla tabella sopra, la distribuzione condizionata di X rispetto a y_1 e la corrispondente distribuzione condizionata relativa sono rappresentate qui sotto.

X	Distribuzione condizionata $X|y_1$	Distribuzione condizionata relativa
30	2	2/3
32	1	1/3
35	0	0/3
Totale	**3**	**1**

Generalizziamo

Più in generale, supponiamo che il carattere X abbia manifestato le modalità distinte $x_1, x_2, ..., x_k$ e che il carattere Y abbia manifestato le modalità distinte $y_1, y_2, ..., y_h$. Indicando con $f(x_i, y_j)$ la frequenza congiunta della coppia (x_i, y_j), con $f(x_1), ..., f(x_k)$ le frequenze (dette **marginali**) della distribuzione marginale di X, con $f(y_1), ..., f(y_k)$ le frequenze della distribuzione marginale di Y, i dati raccolti nella rilevazione congiunta di X e Y possono essere organizzati in una tabella a doppia entrata come la **Tab. 1**.

ALTRE NOTAZIONI
La frequenza marginale $f(x_i)$ viene indicata anche con il simbolo $f_{i.}$ e la frequenza marginale $f(y_j)$ con il simbolo $f_{.j}$. Il punto posto prima o dopo l'indice indica il fenomeno che è stato trascurato, per ricordare che si sta lavorando con frequenze marginali.

Tabella 1

Frequenza *assoluta* con cui la modalità x_i si presenta *congiuntamente* alla modalità y_j

X \ Y	y_1	....	y_j	.....	y_h	Totale
x_1	....	....	....	....	....	$f(x_1)$
....	....	....	....	....	....	....
x_i	....	....	$f(x_i, y_j)$	....	....	....
....	....	....	....	....	....	....
x_k	....	....	....	....	....	$f(x_k)$
Totale	$f(y_1)$	....	....	....	$f(y_h)$	n

Somme per *riga*: forniscono la distribuzione **marginale** di X

Somme per *colonna*: forniscono la distribuzione **marginale** di Y

Numero complessivo di unità del collettivo

L'organizzazione dei dati in una tabella a doppia entrata permette dunque di riassumere molti tipi di informazioni:

1. il comportamento *congiunto* di X e di Y: è rappresentato nelle caselle della tabella non appartenenti ai bordi, dove sono riportate le frequenze *congiunte* di X e Y;
2. il comportamento di X e Y, considerati *singolarmente*: è rappresentato sull'ultima colonna e sull'ultima riga della tabella, dove sono riportate le frequenze *marginali* di X e Y;
3. il comportamento di un carattere (X o Y) condizionatamente a una modalità dell'altro: è rappresentato sulle righe e sulle colonne *interne* alla tabella, considerate singolarmente.

Esercizi p. 66

2. Dipendenza e indipendenza statistica

Come già anticipato, lo studio statistico di due caratteri X e Y, rilevati congiuntamente su una data popolazione, si pone tra i vari obiettivi anche quello di stabilire se sussiste qualche *relazione di dipendenza* tra X e Y.
I metodi che esporremo in questo paragrafo per valutare l'eventuale dipendenza tra due caratteri possono essere applicati sia a caratteri quantitativi sia a caratteri qualitativi, perché faranno riferimento solo alle *frequenze*; tuttavia, nella pratica si utilizzano prevalentemente per caratteri di tipo *qualitativo*, poiché per caratteri quantitativi esistono strumenti statistici più adeguati (che presenteremo nel **Paragrafo 3**).

Dipendenza e indipendenza

Dati due caratteri X e Y, per stabilire se X dipende o meno da Y viene naturale l'idea di confrontare le distribuzioni di X *condizionate* alle modalità di Y con la distribuzione *marginale* di X (che esprime il comportamento di X considerato singolarmente). Se c'è indipendenza, c'è da aspettarsi che il condizionamento di X alle modalità di Y non abbia alcun effetto, ossia che le distribuzioni condizionate si mantengano uguali a quella marginale. Occorre però prestare attenzione a un aspetto: le frequenze *marginali* si riferiscono all'*intera popolazione*, mentre le frequenze *condizionate* si riferiscono soltanto alla *sottopopolazione* che presenta la modalità rispetto cui stiamo condizionando. Non sarebbe perciò corretto eseguire il confronto tra le frequenze assolute: il confronto deve essere fatto tra frequenze *relative*. Queste considerazioni portano alla seguente definizione.

DEFINIZIONE | Indipendenza

Il carattere X si dice **indipendente** da Y se le distribuzioni **condizionate relative** di X rispetto alle modalità di Y sono uguali alla distribuzione **marginale relativa** di X.

Ogni qualvolta due caratteri **non** sono tra loro indipendenti, si dirà che esiste una **dipendenza** tra di essi; in particolare, se i due caratteri sono qualitativi, la dipendenza si chiama **connessione**.

OSSERVA
La condizione di indipendenza [1] richiede in pratica che *ogni* frequenza congiunta sia uguale al prodotto delle corrispondenti frequenze marginali, diviso per *n*.

TEOREMA 1 | Indipendenza tra due caratteri

Due caratteri X e Y, di cui sono state osservate rispettivamente le modalità distinte $x_1, ..., x_k$ e $y_1, .., y_h$, su una popolazione costituita da n unità, sono **indipendenti** se e solo se risulta:

$$f(x_i, y_j) = \frac{f(x_i)f(y_j)}{n} \text{ per ogni } i = 1, ..., k \text{ e per ogni } j = 1, ..., h \qquad [1]$$

DIMOSTRAZIONE

Facciamo riferimento per maggiore chiarezza alla generica tabella che rappresenta la distribuzione doppia di frequenze di X e Y.

$X \backslash Y$	y_1	y_2	....	y_j	.....	y_h	**Totale**
x_1	$f(x_1, y_1)$	$f(x_1, y_2)$	....	$f(x_1, y_j)$	....	$f(x_1, y_h)$	$f(x_1)$
....	....	....	....	....	....	....	....
x_i	....	....	....	....	....	....	....
....	....	....	....	....	....	....	....
x_k	....	....	....	....	....	....	$f(x_k)$
Totale	$f(y_1)$	$f(y_2)$	....	$f(y_j)$	....	$f(y_h)$	n

Distribuzione di Y condizionata a x_1

Distribuzione marginale di Y

Supponiamo che Y sia indipendente da X, cioè che le distribuzioni condizionate relative di Y siano uguali alla distribuzione marginale relativa di Y (il caso simmetrico in cui X è indipendente da Y si può trattare in modo del tutto analogo). Fissiamo l'attenzione sulla seconda riga della tabella, che rappresenta la distribuzione di Y condizionata alla modalità x_1 di X. La corrispondente distribuzione condizionata *relativa* si ottiene ponendo a rapporto le frequenze con il corrispondente totale di riga:

$$\frac{f(x_1, y_1)}{f(x_1)}, \frac{f(x_1, y_2)}{f(x_1)}, \ldots, \frac{f(x_1, y_h)}{f(x_1)}$$

Analogamente, la distribuzione marginale relativa di Y è data da:

$$\frac{f(y_1)}{n}, \frac{f(y_2)}{n}, \ldots, \frac{f(y_h)}{n}$$

Affinché queste due distribuzioni di frequenze relative siano uguali dovrà essere:

$$\frac{f(x_1, y_1)}{f(x_1)} = \frac{f(y_1)}{n}, \frac{f(x_1, y_2)}{f(x_1)} = \frac{f(y_2)}{n}, \ldots, \frac{f(x_1, y_h)}{f(x_1)} = \frac{f(y_h)}{n}$$

ossia:

$$f(x_1, y_1) = \frac{f(x_1)f(y_1)}{n} \quad f(x_1, y_2) = \frac{f(x_1)f(y_2)}{n}, \quad \ldots, \quad f(x_1, y_h) = \frac{f(x_1)f(y_h)}{n}$$

Altrettante analoghe uguaglianze scaturiscono condizionando Y rispetto alle modalità $x_2, \ldots, x_k$.

In definitiva deve essere:

$$f(x_i, y_j) = \frac{f(x_i)f(y_j)}{n} \quad \text{per ogni } i = 1, \ldots, k \text{ e per ogni } j = 1, \ldots, h$$

È importante fare alcune osservazioni.

a. La formula [1] è simmetrica rispetto a x_i e y_j: da ciò segue che X è indipendente da Y se e solo se Y è indipendente da X.

b. Le frequenze congiunte che realizzano la condizione di *indipendenza* statistica, ossia le frequenze che si ottengono tramite la formula [1], vengono chiamate **frequenze teoriche di indipendenza**, per distinguerle da quelle effettivamente osservate: per non confonderle con queste ultime, indicheremo le frequenze teoriche di indipendenza con un apice, con il simbolo $f'(x_i, y_j)$.

c. A ogni tabella che rappresenta una distribuzione doppia di frequenze osservate è possibile affiancare la tabella *teorica* di indipendenza, che si costruisce mantenendo fisse le distribuzioni marginali e sostituendo le frequenze congiunte osservate con quelle teoriche di indipendenza. L'indipendenza tra i due caratteri in esame sarà verificata se e solo se la tabella teorica di indipendenza coincide con la tabella delle frequenze osservate.

ESEMPIO Indipendenza tra due caratteri

Consideriamo la tabella qui sotto a sinistra (che si riferisce a due caratteri qualitativi osservati su un collettivo di 50 unità) e costruiamo la tabella teorica di indipendenza (a destra).

X \ Y	y_1	y_2	Totale
x_1	15	10	25
x_2	7	8	15
x_3	8	2	10
Totale	30	20	50

Tabella osservata

$\Longrightarrow$

X \ Y	y_1	y_2	Totale
x_1	$\frac{30 \cdot 25}{50} = 15$	$\frac{20 \cdot 25}{50} = 10$	25
x_2	$\frac{30 \cdot 15}{50} = 9$	$\frac{20 \cdot 15}{50} = 6$	15
x_3	$\frac{30 \cdot 10}{50} = 6$	$\frac{20 \cdot 10}{50} = 4$	10
Totale	30	20	50

Tabella teorica di indipendenza

Poiché le due tabelle non coincidono, concludiamo che i due caratteri X e Y **non** sono statisticamente indipendenti.

È importante osservare che, mentre le frequenze congiunte osservate sono costituite da numeri *interi*, le frequenze teoriche di indipendenza in generale **non** lo sono (dal momento che sono ottenute come *rapporto* tra il prodotto delle corrispondenti frequenze marginali e il numero complessivo di unità della popolazione). La situazione di perfetta indipendenza statistica può quindi realizzarsi solo nel caso fortuito in cui tutte le frequenze teoriche di indipendenza sono numeri interi. Anche ammettendo che si verifichi questa ipotesi, è facile capire che è comunque molto raro che la tabella delle frequenze osservate coincida esattamente con la tabella teorica di indipendenza. Quest'ultima va quindi interpretata come una situazione *ideale* dalla quale è importante capire *quanto* i dati reali si trovano *distanti*.

La misura del grado di dipendenza

Per misurare il grado di dipendenza di due caratteri X e Y dobbiamo quindi confrontare la tabella delle frequenze osservate con quella teorica di indipendenza: il grado di dipendenza sarà *tanto più elevato* quanto più la tabella delle frequenze osservate è *lontana* dalla tabella delle frequenze teoriche di indipendenza. Gli indici statistici che misurano tale «lontananza» si basano sulle *differenze* tra le frequenze osservate e quelle teoriche; tali differenze sono dette **contingenze** e possono essere così definite:

$$\underbrace{c(x_i, y_j)}_{\text{contingenza della coppia } (x_i, y_j)} = \underbrace{f(x_i, y_j)}_{\text{frequenza congiunta della coppia } (x_i, y_j)} - \underbrace{f'(x_i, y_j)}_{\text{frequenza teorica di indipendenza della coppia } (x_i, y_j)}$$

Si può dimostrare che la somma di tutte le contingenze è sempre nulla. Pertanto, per sintetizzare in un unico indice tutte le differenze, non è possibile basarsi semplicemente sulla somma delle contingenze. L'indice sintetico più noto, dovuto al matematico e statistico Karl Pearson (1857-1936), si basa (per ragioni analoghe a quelle viste quando abbiamo introdotto la varianza) sui *quadrati* delle *contingenze* e viene indicato con la lettera greca χ (*chi*) elevata al quadrato.

RIFLETTI
Il significato della [2] è il seguente. Per calcolare l'indice chi-quadrato di una distribuzione doppia di frequenze occorre:
1. per ogni frequenza congiunta, calcolare la contingenza, elevarla al quadrato e dividerla per la corrispondente frequenza teorica di indipendenza;
2. sommare tutti i risultati ottenuti.
La doppia sommatoria nella formula [2] significa che la somma deve includere gli addendi provenienti da *tutte* le frequenze congiunte della distribuzione doppia.

DEFINIZIONE | Indice chi-quadrato

Dati due caratteri X e Y, siano $x_1, ..., x_k$ e $y_1, ..., y_h$ le differenti modalità con cui si manifestano, rispettivamente, X e Y; l'**indice chi-quadrato** (o **chi-quadro**) è così definito:

$$\chi^2 = \sum_{i=1}^{k} \sum_{j=1}^{h} \frac{c^2(x_i, y_j)}{f'(x_i, y_j)} \qquad [2]$$

L'indice χ^2 soddisfa le proprietà che ragionevolmente ci aspettiamo da un indice che misuri la dipendenza di due caratteri; infatti:

- è uguale a 0 se e solo se X e Y sono indipendenti (infatti vale 0 se e solo se tutte le contingenze sono nulle);
- cresce al crescere delle contingenze.

L'indice χ^2 può essere calcolato più rapidamente tramite la formula seguente, che si dimostra essere equivalente alla [2] e ha il vantaggio di consentire di evitare il calcolo delle contingenze e delle frequenze teoriche di indipendenza.

DEFINIZIONE | Formula abbreviata dell'indice chi-quadrato

Dati due caratteri X e Y, siano $x_1, ..., x_k$ e $y_1, ..., y_h$ le differenti modalità con cui si manifestano, rispettivamente, X e Y su una popolazione di n unità; l'**indice di connessione chi-quadrato** può essere calcolato tramite la formula:

$$\chi^2 = n\left(\sum_{i=1}^{k}\sum_{j=1}^{h}\frac{f^2(x_i, y_j)}{f(x_i)f(y_j)} - 1\right) \qquad [3]$$

RIFLETTI

In pratica, per calcolare l'indice chi-quadrato di una distribuzione doppia di frequenze in base alla [3] si procede così:
1. per ogni frequenza congiunta, la si eleva al quadrato e si divide il quadrato per il prodotto delle corrispondenti frequenze marginali;
2. si sommano tutti i risultati ottenuti, quindi si sottrae dalla somma 1 e si moltiplica il risultato per n.
La doppia sommatoria nella formula [3] significa che la somma deve includere gli addendi provenienti da *tutte* le frequenze congiunte della distribuzione doppia.

ESEMPIO Indice chi-quadrato

In riferimento alla tabella dell'esempio precedente, calcoliamo l'indice di connessione χ^2.

X \ Y	y_1	y_2	Totale
x_1	15	10	25
x_2	7	8	15
x_3	8	2	10
Totale	30	20	50

$$\Longrightarrow \chi^2 = \underbrace{50}_{n}\left(\frac{15^2}{30\cdot 25}+\frac{10^2}{20\cdot 25}+\frac{7^2}{30\cdot 15}+\frac{8^2}{20\cdot 15}+\frac{8^2}{30\cdot 10}+\frac{2^2}{20\cdot 10}-1\right) =$$

$$= \frac{25}{9} \simeq 2{,}8$$

Si pone ora il problema di *interpretare* l'indice chi-quadrato. Il valore trovato nell'esempio precedente è tanto o poco? È indice di una connessione forte o debole? Per rispondere a queste domande occorre *normalizzare* l'indice, cioè trasformarlo in un numero compreso tra 0 e 1, in modo che sia facilmente interpretabile. Questo obiettivo si ottiene dividendo l'indice per il suo valore massimo.
Si dimostra che il valore massimo che può assumere l'indice χ^2 (cioè il valore che assumerebbe nel caso di perfetta connessione) è uguale al prodotto tra n (il numero complessivo di unità del collettivo) e il minimo tra $k-1$ e $h-1$ (essendo k e h rispettivamente il numero di modalità differenti manifestate da X e Y).
Si giunge così alla formula seguente:

$$\text{Indice } \chi^2 \text{ normalizzato} = \frac{\chi^2}{n\cdot\min\{k-1,\, h-1\}}$$

ESEMPIO Indice chi-quadrato normalizzato

In riferimento all'ultimo esempio svolto, normalizziamo l'indice di connessione χ^2.

$$\text{Indice } \chi^2 \text{ normalizzato} = \frac{\chi^2}{n\cdot\min\{k-1,\, h-1\}} = \qquad n=50,\ k=3,\ h=2$$

$$= \frac{\frac{25}{9}}{50\cdot\min\{3-1,\, 2-1\}} = \frac{\frac{25}{9}}{50\cdot\underbrace{\min\{2,\, 1\}}_{1}} = \frac{25}{50\cdot 9} = \frac{1}{18} \simeq 0{,}056$$

Pertanto la connessione tra i due caratteri X e Y è circa il 5,6% della massima possibile: un grado di connessione molto basso.

OSSERVA

Poiché stiamo considerando statistiche bivariate, i caratteri X e Y manifestano almeno due modalità distinte ciascuno, quindi $k \geq 2$ e $h \geq 2$: ciò garantisce che $k-1 \neq 0$ e $h-1 \neq 0$, dunque la formula per l'indice χ^2 normalizzato è ben definita.

Esercizi p. 70

3. Correlazione e regressione

Come abbiamo anticipato nel paragrafo precedente, per indagare sull'esistenza di possibili relazioni tra due caratteri *quantitativi* X e Y è possibile introdurre strumenti più adeguati di quelli appena visti, perché è possibile lavorare anche sulle *modalità* di X e Y (mentre per caratteri qualitativi è necessario limitarsi alle *frequenze*). In questo paragrafo introduciamo questi nuovi strumenti.

Correlazione

Premettiamo che la dipendenza tra due caratteri di tipo *quantitativo* viene chiamata **correlazione** (per distinguerla dalla dipendenza tra due caratteri *qualitativi* che, come abbiamo visto, viene chiamata invece **connessione**).
Un indice statistico molto utilizzato per valutare la *correlazione* tra *due* caratteri quantitativi è la cosiddetta *covarianza*, così definita.

DEFINIZIONE | **Covarianza**

Siano X e Y due variabili statistiche di medie $\overline{x}$ e $\overline{y}$, rilevate congiuntamente su un collettivo di n unità. Siano $x_1, x_2, ..., x_n$ i valori osservati di X e $y_1, y_2, ..., y_n$ i corrispondenti valori osservati di Y. Si chiama **covarianza** di X e Y, e si indica con il simbolo σ_{XY}, il numero così definito:

$$\sigma_{XY} = \frac{\sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})}{n} \qquad [4]$$

Il significato della covarianza appare chiaro se interpretiamo geometricamente la formula [4]. Immaginiamo di avere rappresentato in un piano cartesiano i punti di coordinate (x_i, y_i), con $i = 1, ..., n$: si ottiene una «nuvola» di punti. Definiamo **baricentro** di questa nuvola il punto di coordinate $(\overline{x}, \overline{y})$ e tracciamo le parallele agli assi cartesiani passanti per tale punto. Queste rette dividono il piano in quattro angoli retti, che numeriamo in senso antiorario come indicato in **Fig. 1** a partire da quello in alto a destra. A seconda che il punto (x_i, y_i) sia interno all'angolo **I**, **II**, **III** o **IV**, gli scarti $(x_i - \overline{x})$ e $(y_i - \overline{y})$ hanno il segno illustrato in **Fig. 1** e di conseguenza il prodotto $(x_i - \overline{x})$ $(y_i - \overline{y})$ ha il segno illustrato in **Fig. 2**.

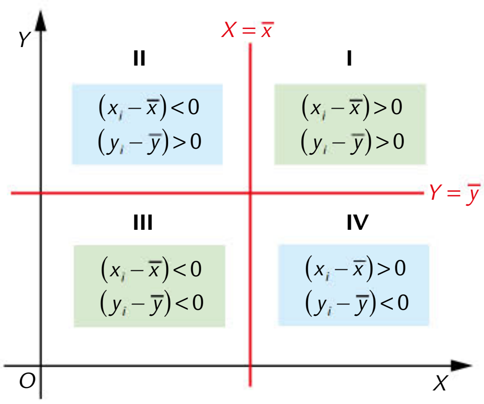


Figura 1 Il segno degli scarti dei valori osservati dalle rispettive medie.

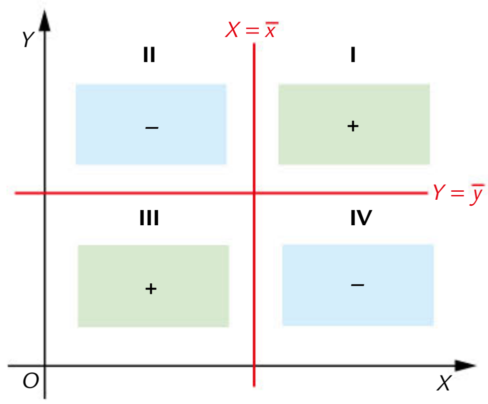


Figura 2 I punti interni ai quattro angoli retti numerati contribuiscono al calcolo della covarianza secondo i segni indicati.

Tenendo conto di queste ultime considerazioni, possiamo giungere alle seguenti riflessioni.

a. Se la covarianza è *positiva*, cioè se i prodotti $(x_i - \overline{x})(y_i - \overline{y})$ di segno positivo prevalgono su quelli di segno negativo, allora i punti di coordinate (x_i, y_i) cadono perlopiù internamente ai due angoli retti **I** e **III**; la nuvola di punti deve perciò avere la forma in **Fig. 3**: tale forma è indicativa di una relazione di *tipo lineare crescente* tra le variabili X e Y.

b. Se la covarianza è *negativa*, cioè se prevalgono prodotti $(x_i - \overline{x})(y_i - \overline{y})$ di segno negativo, allora i punti di coordinate (x_i, y_i) cadono perlopiù internamente ai due angoli retti **II** e **IV**; la nuvola di punti deve perciò avere la forma in **Fig. 4**: tale forma è indicativa di una relazione di *tipo lineare decrescente* tra le variabili X e Y.

c. Se la covarianza è *nulla*, i punti sono sparpagliati senza alcuna regolarità oppure sono disposti secondo relazioni diverse e lontane da quella lineare (per esempio ciò accade nel caso di una relazione quadratica).

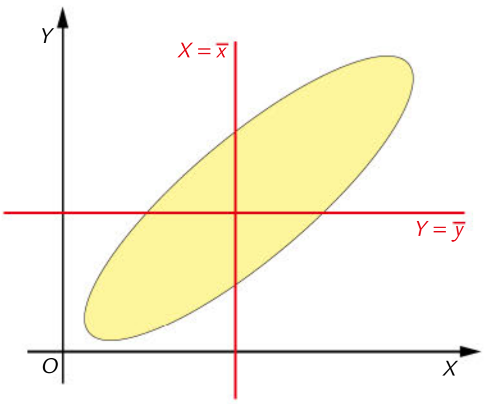


Figura 3 **Nuvola di punti che genera una correlazione positiva.**

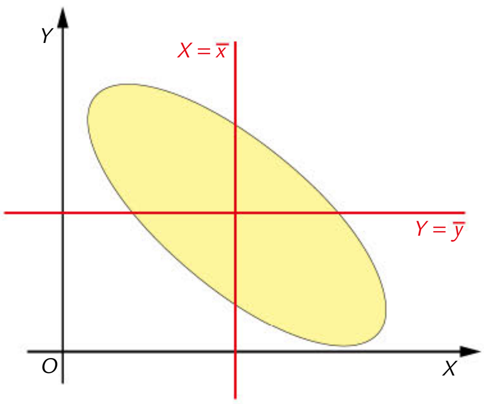


Figura 4 **Nuvola di punti che genera una correlazione negativa.**

La covarianza si può calcolare più rapidamente rispetto alla definizione mediante la formula seguente, che si può dimostrare essere equivalente alla [4].

DEFINIZIONE | Formula «abbreviata» per il calcolo della covarianza

Siano X e Y due variabili statistiche di medie $\overline{x}$ e $\overline{y}$, rilevate congiuntamente su un collettivo di n unità. Siano $x_1, x_2, ..., x_n$ i valori osservati di X e $y_1, y_2, ..., y_n$ i corrispondenti valori osservati di Y. La **covarianza** di X e Y è espressa dalla formula:

$$\sigma_{XY} = \frac{\sum_{i=1}^{n} x_i y_i}{n} - \overline{x} \cdot \overline{y}$$

Una volta appurata una correlazione tra due variabili statistiche X e Y, si pone il problema di stabilire se essa è *forte* o *debole*. Questo obiettivo si raggiunge, similmente a quanto già visto per l'indice chi-quadrato, costruendo un indice *relativo*, ottenuto ponendo la covarianza a rapporto con il suo valore *massimo*. A quest'ultimo proposito vale il seguente teorema.

TEOREMA 2 | Minimo e massimo della covarianza

La **covarianza** di due variabili X e Y può assumere valori appartenenti al seguente intervallo:

$$-\sigma_X \sigma_Y \leq \sigma_{XY} \leq \sigma_X \sigma_Y$$

dove σ_X e σ_Y sono le deviazioni standard di X e Y.

Tenendo conto del **Teorema 2**, siamo in grado di introdurre l'indice relativo cercato.

DEFINIZIONE | Coefficiente di correlazione lineare

Si chiama **coefficiente di correlazione lineare** di due variabili X e Y, e si indica con il simbolo r, il numero così definito:

$$r = \frac{\sigma_{XY}}{\sigma_X \cdot \sigma_Y}$$

ALTRE NOTAZIONI

1. Il coefficiente di correlazione lineare è stato proposto da **Karl Pearson** (1857-1936) e **Auguste Bravais** (1811-1863); per questo è anche noto come *indice di Bravais-Pearson*.

2. Il coefficiente di correlazione lineare tra due variabili X e Y viene talvolta indicato con la lettera ρ.

È importante fare alcune osservazioni.

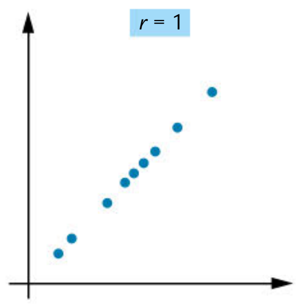
a. Per come è stato definito, risulta sempre:

$$-1 \leq r \leq 1$$

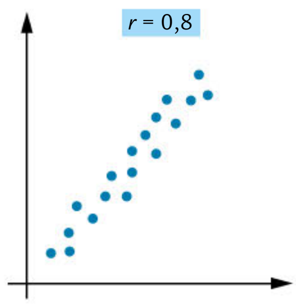
b. Il segno del coefficiente di correlazione lineare è lo stesso della covarianza e dà informazioni analoghe: un coefficiente $r > 0$ indica una relazione lineare *crescente*, mentre un coefficiente $r < 0$ indica una relazione lineare *decrescente*.

c. Si può dimostrare che l'indice di correlazione r è uguale a ± 1 se e solo se tra Y e X sussiste una *perfetta* relazione lineare. Tanto più r è vicino a ± 1, quanto più il modello lineare interpreta bene la relazione che sussiste tra Y e X; tanto più r è vicino a 0, quanto più il legame tra Y e X (se c'è) è distante da quello lineare, come illustrato nelle figure della seguente tabella.

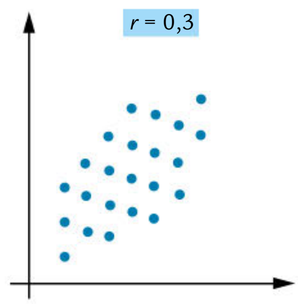
d. In base a come è definito il coefficiente di correlazione lineare r, tra X e Y sussiste una perfetta relazione lineare se e solo se il modulo della covarianza dei due caratteri è uguale al prodotto delle rispettive deviazioni standard, cioè se $|\sigma_{XY}| = \sigma_X \sigma_Y$.



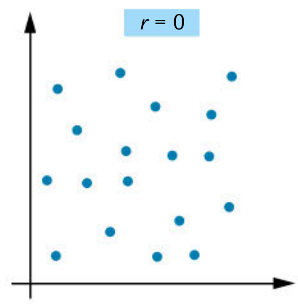
a. Tra X e Y sussiste una perfetta correlazione lineare positiva, ovvero i punti sono allineati su una retta con pendenza positiva.



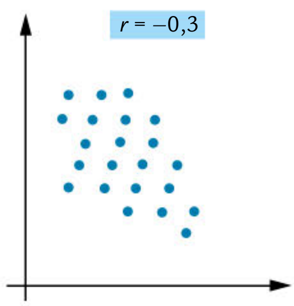
b. Tra X e Y sussiste una forte correlazione lineare positiva.



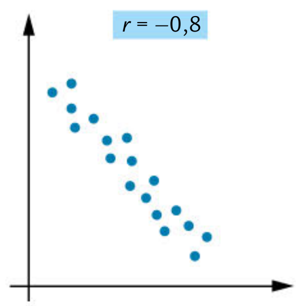
c. Tra X e Y sussiste una debole correlazione lineare positiva.



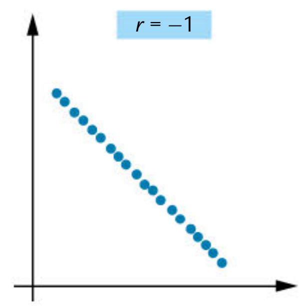
d. Tra X e Y non sussiste alcun legame di correlazione lineare.



e. Tra X e Y sussiste una debole correlazione lineare negativa.



f. Tra X e Y sussiste una forte correlazione lineare negativa.



g. Tra X e Y sussiste una perfetta correlazione lineare negativa, ovvero i punti sono allineati su una retta con pendenza negativa.

ESEMPIO Calcolo del coefficiente di correlazione lineare

In 4 supermercati di una nota catena sono stati rilevati la superficie di esposizione, in migliaia di metri quadrati (X), e il fatturato settimanale, in migliaia di euro (Y). Sono stati ottenuti i dati riassunti nella seguente tabella:

x_i	0,2	0,5	0,8	1
y_i	50	120	150	200

Determiniamo il coefficiente di correlazione lineare di X e Y.

Per determinare il coefficiente di correlazione lineare dobbiamo calcolare σ_X, σ_Y (le deviazioni standard di X e Y) e σ_{XY} (la covarianza di X e Y).
Al fine di agevolare i calcoli, è utile organizzare il lavoro in una tabella come quella qui di seguito.

Tabella 2

x_i	y_i	$x_i\, y_i$	x_i^2	y_i^2
0,2	50	10	0,04	2500
0,5	120	60	0,25	14 400
0,8	150	120	0,64	22 500
1,0	200	200	1,00	40 000
$\sum x_i = 2{,}5$	$\sum y_i = 520$	$\sum x_i y_i = 390$	$\sum x_i^2 = 1{,}93$	$\sum y_i^2 = 79\,400$

ATTENZIONE!
Talvolta, come nell'esempio qui a fianco, per brevità si omettono gli indici di sommatoria, sottintendendo che la somma va estesa a tutti gli indici *i*.

Possiamo ora comodamente determinare tutti gli elementi che ci servono:

$$\overline{x} = \frac{\sum x_i}{4} = \frac{2{,}5}{4} = 0{,}625 \qquad \overline{y} = \frac{\sum y_i}{4} = \frac{520}{4} = 130$$

$$\sigma_X^2 = \frac{\sum x_i^2}{4} - \overline{x}^2 = \frac{1{,}93}{4} - (0{,}625)^2 = 0{,}091875$$

$$\sigma_Y^2 = \frac{\sum y_i^2}{4} - \overline{y}^2 = \frac{79\,400}{4} - 130^2 = 2950$$

$$\sigma_{XY} = \frac{\sum x_i y_i}{4} - \overline{x} \cdot \overline{y} = \frac{390}{4} - 0{,}625 \cdot 130 = 16{,}25$$

Concludiamo che il coefficiente di correlazione lineare è uguale a:

$$r = \frac{\sigma_{XY}}{\sigma_X \cdot \sigma_Y} = \frac{16{,}25}{\sqrt{0{,}091875} \cdot \sqrt{2950}} \simeq 0{,}987$$

Regressione lineare

Dopo avere scoperto l'esistenza di una relazione lineare tra le due variabili X e Y, in base all'analisi di un diagramma cartesiano o al calcolo del coefficiente di correlazione (che deve essere vicino a ± 1), ci proponiamo di determinare la funzione lineare che interpreta *meglio* tale legame, nel senso che ora precisiamo.

Con GeoGebra Visualizzazione dinamica della retta di regressione

1. Consideriamo una generica funzione lineare di equazione $y = mx + q$ e, per ogni punto $P_i(x_i, y_i)$ appartenente alla nuvola che rappresenta i dati, consideriamo il corrispondente punto $Q_i(x_i, y_i')$ di ascissa x_i appartenente alla retta che costituisce il grafico della funzione $y = mx + q$. Costruiamo quindi i vari segmenti P_1Q_1, ..., P_iQ_i, ..., P_nQ_n (*vedi* la **Fig. 5**, che visualizza il caso in cui $n = 5$).

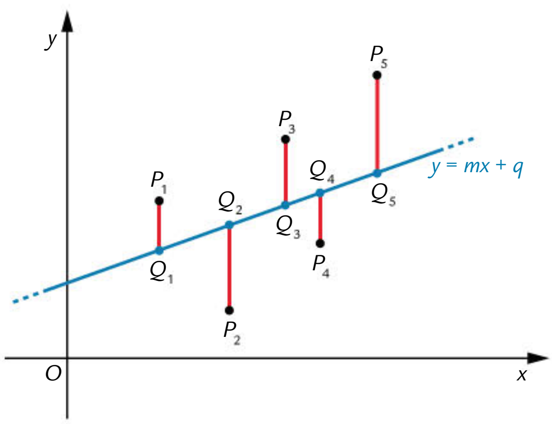


Figura 5

2. Calcoliamo le lunghezze dei vari segmenti P_1Q_1, ..., P_iQ_i, ..., P_nQ_n; in generale sarà: $\overline{P_iQ_i} = |y_i - y_i'|$
3. Eleviamo al quadrato tali lunghezze e le sommiamo:

$$\sum_{i=1}^{n}(y_i - y_i')^2 \qquad [5]$$

4. Questa somma esprime, tramite un *unico* numero, una misura della distanza complessiva fra i dati y_i osservati e i valori teorici y_i' calcolati sul grafico della retta. Scegliamo, come funzione lineare che *meglio* approssima (in linguaggio tecnico si dice **interpola**) i dati, quella per cui la somma [5] risulta *minima*. La retta che costituisce il grafico di questa funzione si chiama **retta di regressione** e la sua equazione può essere determinata in base al teorema seguente, che ci limitiamo a enunciare.

MODI DI DIRE

La retta di regressione è chiamata anche **retta dei minimi quadrati**.

TEOREMA 3 | Retta di regressione

Date due variabili X e Y, di valori medi rispettivamente $\overline{x}$ e $\overline{y}$, la **retta di regressione** che esprime Y in funzione di X è la retta che passa per il punto di coordinate $(\overline{x}, \overline{y})$ e che ha come coefficiente angolare m il cosiddetto **coefficiente di regressione**, così definito:

$$m = \frac{\sigma_{XY}}{\sigma_X^2} \qquad [6]$$

OSSERVA

Il punto di coordinate $(\overline{x}, \overline{y})$ è quello che abbiamo definito **baricentro** della nuvola di punti che rappresenta i dati osservati.

Ne segue che l'equazione della retta di regressione è:

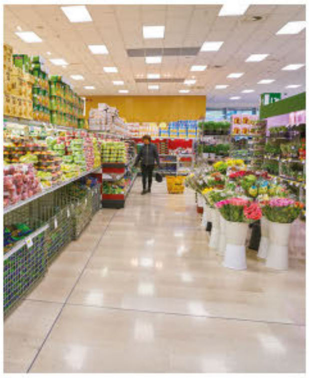
$$y - \overline{y} = m(x - \overline{x}) \qquad \text{dove } m = \frac{\sigma_{XY}}{\sigma_X^2}$$

Scrivendo il rapporto $\frac{\sigma_{XY}}{\sigma_X^2}$ tramite le formule abbreviate per il calcolo della covarianza e della deviazione standard, cioè:

$$\sigma_{XY} = \frac{\sum_{i=1}^{n} x_i y_i}{n} - \overline{x}\,\overline{y} = \frac{\sum_{i=1}^{n} x_i y_i - n\,\overline{x}\,\overline{y}}{n} \quad \text{e} \quad \sigma_X^2 = \frac{\sum_{i=1}^{n} x_i^2}{n} - \overline{x}^2 = \frac{\sum_{i=1}^{n} x_i^2 - n\,\overline{x}^2}{n}$$

la formula che fornisce il coefficiente angolare della retta di regressione di Y in funzione di X si può anche scrivere nella seguente forma equivalente:

$$m = \frac{\sum_{i=1}^{n} x_i y_i - n\,\overline{x}\,\overline{y}}{\sum_{i=1}^{n} x_i^2 - n\,\overline{x}^2} \qquad [7]$$



ESEMPIO Calcolo della retta di regressione

In 4 supermercati di una nota catena sono stati rilevati la superficie di esposizione, in migliaia di metri quadrati (X) e il fatturato settimanale, in migliaia di euro (Y). Sono stati ottenuti i seguenti dati:

x_i	0,2	0,5	0,8	1
y_i	50	120	150	200

Scriviamo l'equazione della retta di regressione che esprime Y in funzione di X.

- Nell'esempio precedente abbiamo visto che il *coefficiente di correlazione lineare* tra X e Y è circa 0,987. Essendo questo coefficiente prossimo a 1, la retta di regressione è certamente un modello che interpreta molto bene il legame tra X e Y.
- Abbiamo già calcolato nell'esempio precedente i valori che servono a scrivere l'equazione della retta di regressione:

  $\overline{x} = 0{,}625 \qquad \overline{y} = 130$

  $\sigma_X^2 = 0{,}091875 \qquad \sigma_{XY} = 16{,}25$
- Il coefficiente angolare della retta di regressione è uguale a:

  $$m = \frac{\sigma_{XY}}{\sigma_X^2} = \frac{16{,}25}{0{,}091875} \simeq 176{,}87$$ **Formula [6]**

  In alternativa, possiamo calcolare il coefficiente angolare tramite la formula [7]. Ricorda che in **Tab. 2** abbiamo calcolato che:

  $\sum x_i y_i = 390 \quad \text{e} \quad \sum x_i^2 = 1{,}93$

  Pertanto, essendo $n = 4$:

  $$m = \frac{\sum x_i y_i - n\overline{x} \cdot \overline{y}}{\sum x_i^2 - n\overline{x}^2} = \frac{390 - 4 \cdot 0{,}625 \cdot 130}{1{,}93 - 4 \cdot (0{,}625)^2} \simeq 176{,}87$$

  L'equazione della retta di regressione è perciò:

  $$\underbrace{y - 130}_{y - \overline{y}} = 176{,}87\underbrace{(x - 0{,}625)}_{x - \overline{x}} \qquad [8]$$

Se ne ricava che l'equazione esplicita della [8] è approssimativamente:

$y = 176{,}87x + 19{,}46$

È importante infine fare alcune osservazioni.

1. La retta di regressione può essere utilizzata sia per determinare (in linguaggio tecnico si dice **estrapolare**) dati mancanti all'interno dei dati osservati, sia per fare delle «previsioni» sull'andamento tendenziale (il cosiddetto **trend**) di un fenomeno nel futuro. Quest'ultimo impiego è particolarmente frequente nelle applicazioni economiche e sociali, nel caso in cui i dati osservati costituiscano una *serie storica* e la variabile indipendente sia il *tempo*.
2. Il metodo dei minimi quadrati garantisce che la distanza complessiva tra i dati osservati e quelli teorici sulla retta di regressione sia la *minima* possibile; il fatto che tale distanza sia minima non implica però che sia necessariamente *piccola*. È necessario perciò definire un *indice* che misuri la bontà di adattamento della retta di regressione ai dati e che sia compreso tra 0 e 1, in modo da essere facilmente interpretabile. Tale indice, di cui per semplicità omettiamo la definizione operativa, si chiama **coefficiente di determinazione** e si dimostra essere uguale al *quadrato* del *coefficiente di correlazione lineare*:

   coefficiente di determinazione $= r^2$

RICORDA
Si ha una serie storica quando un fenomeno viene rilevato a intervalli regolari di tempo.

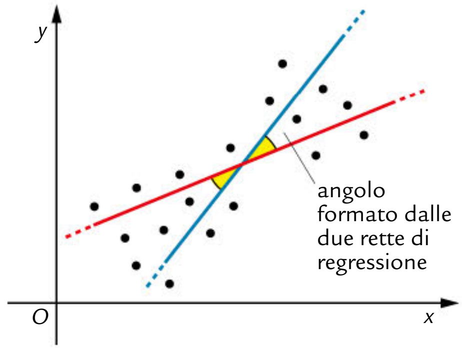
Come è naturale (in base a quanto osservato a proposito del coefficiente di correlazione lineare r), quanto più r è prossimo a ± 1, quindi quanto più il coefficiente di determinazione è prossimo a 1, tanto più la retta di regressione presenta un *buon* accostamento ai dati; quanto più r è prossimo a zero (per valori positivi o negativi) e, quindi, il coefficiente di determinazione è vicino a zero, tanto più la retta di regressione presenta un cattivo accostamento ai dati e anche eventuali estrapolazioni risultano inaffidabili.

PER SAPERNE DI PIÙ **La retta di regressione di *X* in funzione di *Y***

Dati due caratteri quantitativi X e Y, abbiamo visto come scrivere l'equazione della retta di regressione che esprime i valori di Y in funzione dei valori di X. È anche possibile determinare la retta di regressione che esprime i valori di X in funzione dei valori di Y. Si può dimostrare che l'equazione di tale retta è:

$$x - \bar{x} = m(y - \bar{y}) \qquad \text{dove} \qquad m = \frac{\sigma_{XY}}{\sigma_Y^2}$$

L'angolo formato dalle due rette di regressione che esprimono Y in funzione di X e X in funzione di Y è tanto più piccolo quanto più il coefficiente di determinazione è vicino a 1: quando il coefficiente di determinazione è 1, l'angolo è nullo e le due rette coincidono.



Regressione non lineare

A volte può accadere che un modello *lineare* **non** sia adatto a interpretare la relazione tra due variabili; per esempio, il legame tra due variabili potrebbe essere di tipo polinomiale di grado superiore al primo, esponenziale o logaritmico. Anche in questi casi si può costruire un modello di regressione, detto **non lineare**; in generale però la complessità dal punto di vista matematico cresce notevolmente. Ci limitiamo perciò a trattare due casi che si possono facilmente *ricondurre* al modello *lineare*, quelli in cui si cerca una funzione di regressione di tipo **potenza** o di tipo **esponenziale**, cioè una funzione di regressione la cui equazione sia rispettivamente in una delle seguenti forme:

$$y = ax^b \qquad y = ab^x$$

1. Funzioni potenza

Supponiamo che tra le due grandezze x e y sussista una relazione del tipo $y = ax^b$, con $a > 0$ e $x > 0$. Passando ai logaritmi otteniamo:

$$\ln y = \ln(ax^b) \underbrace{\quad\Rightarrow\quad}_{\substack{\text{proprietà del}\\ \text{logaritmo di}\\ \text{un prodotto}}} \ln y = \ln a + \ln x^b \underbrace{\quad\Rightarrow\quad}_{\substack{\text{proprietà del}\\ \text{logaritmo di}\\ \text{una potenza}}} \ln y = \ln a + b\ln x$$

Se ora poniamo $y' = \ln y$ e $x' = \ln x$, l'ultima relazione scritta si può scrivere nella forma:

$$y' = bx' + \ln a \qquad [9]$$

Scopriamo così che sussiste una relazione tipo *lineare* tra le due variabili x' e y', ossia tra il logaritmo di x e il logaritmo di y. Questa osservazione suggerisce come si può ricondurre al caso *lineare* il problema di trovare una funzione di regressione di tipo *potenza*.

a. Si determinano i logaritmi dei valori rilevati delle due variabili x e y;

b. si determina l'equazione della retta di regressione che esprime il legame tra le due variabili $x' = \ln x$ e $y' = \ln y$; sia tale equazione:

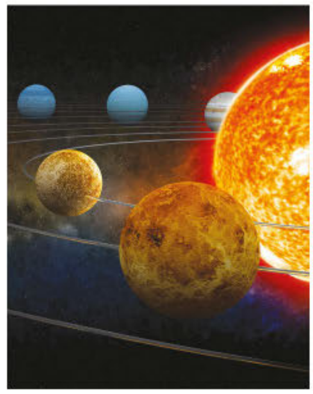
$$y' = mx' + q \qquad [10]$$

c. si ricavano i valori di a e di b, cioè dei parametri che definiscono l'equazione $y = ax^b$ della funzione potenza di regressione, confrontando le equazioni [9] e [10], da cui si deduce che $b = m$ e $\ln a = q$, quindi:

$$a = e^q \quad \text{e} \quad b = m \qquad [11]$$

ESEMPIO Regressione tramite una funzione potenza

Nella seguente tabella sono riportati, per i pianeti visibili, la distanza media x del pianeta dal Sole (in milioni di km) e il tempo di rivoluzione y intorno al Sole.



Pianeta	Mercurio	Venere	Terra	Marte	Giove	Saturno
Distanza media dal Sole (x)	58	108	150	228	778	1423
Tempo di rivoluzione (y), in giorni	88	225	365	687	4333	10 760

È noto che la funzione che esprime il tempo di rivoluzione in funzione della distanza media dal Sole è una funzione potenza. Cerchiamo una funzione di regressione della forma $y = ax^b$ che interpola i dati.

a. Determiniamo i logaritmi dei dati osservati.

Pianeta	Mercurio	Venere	Terra	Marte	Giove	Saturno
Logaritmi delle distanze ($x' = \ln x$)	4,06	4,68	5,01	5,43	6,66	7,26
Logaritmi dei tempi di rivoluzione ($y' = \ln y$)	4,48	5,42	5,90	6,53	8,37	9,28

b. Determiniamo ora l'equazione della retta di regressione che esprime y' in funzione di x'. Procedendo come abbiamo visto in precedenza o avvalendoci delle funzioni predefinite di un software di calcolo troviamo che tale retta ha equazione:

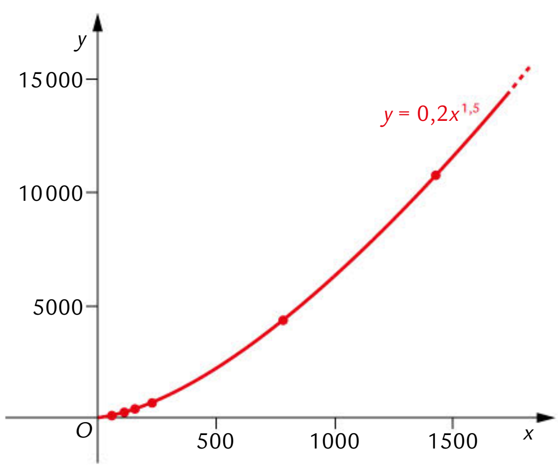
$$y' = 1{,}5x' - 1{,}6 \qquad [12]$$

c. In base alle formule [11]:

$$a = e^{-1{,}6} \simeq 0{,}2 \qquad b = 1{,}5$$

La funzione potenza che interpola i dati (vedi figura qui sotto) è dunque:

$$y = 0{,}2x^{1{,}5} \qquad [13]$$



OSSERVA

Dalla [13] segue, elevando i due membri al quadrato, che

$$y^2 = 0{,}04x^3 \Rightarrow \frac{y^2}{x^3} = 0{,}04$$

Ritroviamo così la ben nota terza legge di Keplero, la quale afferma che il rapporto tra il quadrato del periodo di rivoluzione di un pianeta e il cubo della distanza media del pianeta dal Sole è costante per tutti i pianeti.

2. Funzioni esponenziali

Supponiamo ora che tra le due grandezze x e y sussista una relazione del tipo $y = ab^x$, con $a > 0$ e $b > 0$.

Passando ai logaritmi otteniamo:

$$\ln y = \ln\ (ab^x) \quad \underbrace{\Rightarrow}_{\substack{\text{proprietà del}\\ \text{logaritmo di}\\ \text{un prodotto}}} \quad \ln y = \ln a + \ln b^x \quad \underbrace{\Rightarrow}_{\substack{\text{proprietà del}\\ \text{logaritmo di}\\ \text{una potenza}}} \quad \ln y = \ln a + x \ln b$$

Se ora poniamo $y' = \ln y$ e $x' = x$, l'ultima relazione scritta si può scrivere nella forma:

$$y' = x' \ln b + \ln a \qquad [14]$$

Scopriamo così che sussiste una relazione tipo lineare tra x' e y', ossia tra x e il logaritmo di y.
Similmente a quanto visto per le funzioni potenza, il problema di trovare una funzione di regressione di tipo esponenziale si può allora ricondurre al caso lineare procedendo come segue:

a. si determinano i logaritmi dei valori osservati di y, lasciando invariati i valori di x;
b. si determina l'equazione della retta di regressione che esprime il legame tra le due variabili $x' = x$ e $y' = \ln y$; sia tale equazione:

$$y' = mx' + q \qquad [15]$$

c. si ricavano i valori di a e di b, cioè dei parametri che definiscono l'equazione $y = ab^x$ della funzione esponenziale di regressione, confrontando le equazioni [14] e [15], da cui si deduce che $\ln b = m$ e $\ln a = q$, quindi:

$$\boxed{a = e^q \ \text{ e } \ b = e^m} \qquad [16]$$

ESEMPIO Regressione tramite una funzione esponenziale

Nella seguente tabella sono riportati i valori della pressione atmosferica (y, in millimetri di mercurio) in corrispondenza di alcune altezze sopra il livello del mare (x, in km).

Altezza sopra il livello del mare (x), in km	0	0,5	1	1,5	2	2,5	3
Pressione (y), in mmHg	760	716	674	634	596	560	526

È noto che la funzione che esprime la pressione atmosferica in funzione dell'altezza è di tipo esponenziale.
Cerchiamo una funzione di regressione della forma $y = ab^x$ che interpola i dati.

a. Determiniamo i logaritmi dei valori di y, lasciando invariati i valori di x.

Altezza sopra il livello del mare (x)	0	0,5	1	1,5	2	2,5	3
Logaritmi dei valori della pressione (y)	6,63	6,57	6,51	6,45	6,39	6,33	6,27

b. Determiniamo l'equazione della retta di regressione che esprime y' in funzione di x'. Procedendo come abbiamo visto in uno degli esempi precedenti o avvalendoci delle funzioni predefinite di un software di calcolo si trova che tale retta ha equazione:

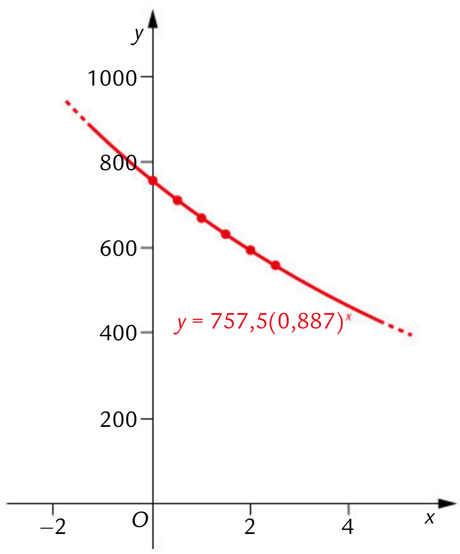
$$y' = -0{,}12x' + 6{,}63 \qquad [17]$$

c. In base alla [16], otteniamo che:

$$a = e^{6{,}63} \simeq 757{,}5 \qquad b = e^{-0{,}12} \simeq 0{,}887$$

La funzione esponenziale che interpola i dati (*vedi* la figura) è dunque:

$$y = 757{,}5\,(0{,}887)^x$$



OSSERVA

Il significato dei coefficienti della funzione esponenziale di regressione è il seguente:
- il coefficiente 757,5 indica approssimativamente la pressione quando $x = 0$, ossia al livello del mare;
- il coefficiente 0,887 indica che la pressione diminuisce circa dell'11,3% al crescere dell'altezza di 1 km.

Talvolta, in assenza di un diagramma cartesiano che indichi chiaramente il tipo di legame che può sussistere tra due variabili, o di modelli teorici che giustifichino la scelta di una particolare funzione di regressione, ci si può trovare nella necessità di stabilire quale modello sia il migliore tra i vari possibili. La scelta preferibile è ovviamente quella del modello che presenta il migliore accostamento ai dati. Per confrontare due possibili modelli che esprimono il legame tra due stesse variabili conviene allora calcolare per ciascuno dei due modelli la somma dei quadrati degli scarti (detti anche **residui**) tra i valori y_i osservati e i valori teorici y_i' stimati dalla funzione di regressione; il modello migliore sarà quello, tra i due, per cui la somma dei quadrati dei residui è *minore*.

Esercizi p. 72

Statistica bivariata, correlazione e regressione

Statistica bivariata

Parte della statistica che si occupa di studiare due caratteri, diciamo X e Y, rilevati congiuntamente sulla popolazione, mettendo in evidenza eventuali relazioni tra X e Y.

Rappresentazione tramite tabella a doppia entrata

differenti modalità osservate di Y

Frequenza *assoluta* con cui la modalità x_i si presenta *congiuntamente* alla modalità y_j

differenti modalità osservate di X

$X \backslash Y$	y_1	....	y_j	.....	y_h	Totale
x_1	....	....	....	....	....	$f(x_1)$
....	....	....	....	....	....	....
x_i	....	....	$f(x_i, y_j)$	....	....	....
....	....	....	....	....	....	....
x_k	....	....	....	....	....	$f(x_k)$
Totale	$f(y_1)$	....	....	....	$f(y_h)$	n

Somme per *riga*: forniscono la distribuzione **marginale** di X

Somme per *colonna*: forniscono la distribuzione **marginale** di Y

Numero complessivo di unità del collettivo

Indipendenza

X e Y risultano indipendenti se e solo se risulta:

$$f(x_i, y_j) = \frac{f(x_i)\, f(y_j)}{n} \qquad \forall i = 1, \ldots\ldots, k \quad \forall j = 1, \ldots\ldots, h$$

Caratteri quantitativi

Caratteri qualitativi

Indice chi-quadrato

La dipendenza tra due caratteri qualitativi viene chiamata **connessione** e viene valutata tramite i seguenti indici:

$$\underbrace{\chi^2 = n\left(\sum_{i=1}^{k}\sum_{j=1}^{h}\frac{f(x_i, y_j)}{f(x_i)\, f(y_j)} - 1\right)}_{\text{indice chi-quadrato}}$$

$$\text{Indice } \chi^2 \text{ normalizzato} = \frac{\chi^2}{n \cdot \min\,(k-1,\ h-1)}$$

Correlazione e covarianza

La dipendenza tra due caratteri quantitativi viene chiamata **correlazione** e viene studiata tramite indici costruiti a partire dalle modalità dei due caratteri X e Y. Un indice per valutare la correlazione è la **covarianza**:

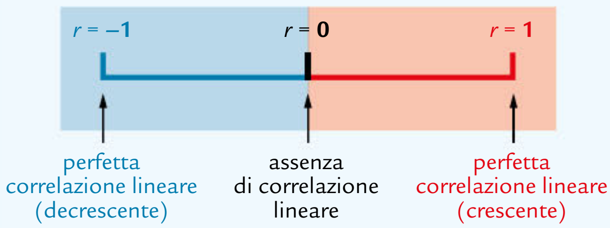
$$\underbrace{\sigma_{XY}}_{\text{covarianza}} = \underbrace{\frac{1}{n}\sum_{i=1}^{n} x_i y_i}_{\substack{n \text{ è il numero di unità del collettivo; } x_i \text{ e } y_i \\ \text{sono i valori di } X \text{ e } Y \text{ osservati sulla } i\text{-esima} \\ \text{unità del collettivo}}} - \underbrace{\bar{x}}_{\substack{\text{media} \\ \text{di } X}} \cdot \underbrace{\bar{y}}_{\substack{\text{media} \\ \text{di } Y}}$$

Coefficiente di correlazione lineare

$$\underbrace{r}_{\substack{\text{coefficiente} \\ \text{di correlazione}}} = \frac{\overbrace{\sigma_{XY}}^{\text{covarianza}}}{\underbrace{\sigma_X}_{\substack{\text{deviazione} \\ \text{standard di } X}} \cdot \underbrace{\sigma_Y}_{\substack{\text{deviazione} \\ \text{standard di } Y}}}$$

Risulta sempre: $-1 \leq r \leq 1$

Significato del coefficiente *r* di correlazione lineare



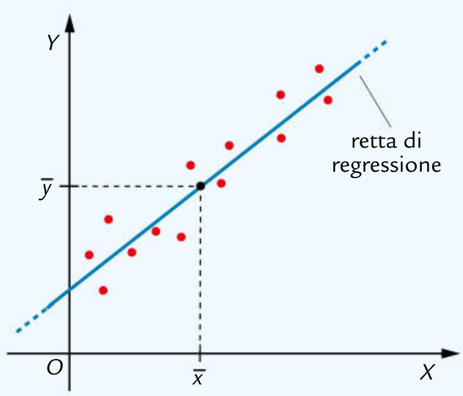
Quanto più r è vicino a ± 1, tanto più il modello *lineare* interpreta bene la relazione tra X e Y; quanto più r è vicino a 0, tanto più la relazione tra X e Y (se c'è) è distante da quella lineare.

Regressione lineare

Se r è vicino a ± 1, tra le variabili X e Y sussiste una relazione di tipo *lineare*. La retta che interpreta meglio tale relazione, esprimendo Y in funzione di X, ha equazione:

$$\underbrace{y - \overbrace{\bar{y}}^{\text{media di } Y} = m(x - \overbrace{\bar{x}}^{\text{media di } X})}_{\substack{\text{equazione della retta} \\ \text{di regressione lineare}}}$$

$$m = \frac{\sigma_{XY}}{\sigma_X^2} = \frac{\sum_{i=1}^{n} x_i y_i - n\,\bar{x}\,\bar{y}}{\sum_{i=1}^{n} x_i^2 - n\,\bar{x}^2}$$



1. Tabelle a doppia entrata

Teoria p. 47

Esercizi introduttivi

1 Vero o falso?

a. in una rilevazione congiunta di due fenomeni X e Y, il fenomeno X viene osservato su alcuni soggetti della popolazione e il fenomeno Y su soggetti diversi da quelli su cui si è osservato X, quindi i dati vengono riuniti in una unica tabella a doppia entrata V F

b. in una distribuzione doppia di frequenze, la somma di tutte le frequenze congiunte è uguale al numero di unità statistiche della popolazione V F

c. in una tabella a doppia entrata, che rappresenta la distribuzione doppia di frequenze di due caratteri X e Y, la somma delle frequenze marginali di X è uguale alla somma delle frequenze marginali di Y V F

d. è possibile ricavare le frequenze marginali conoscendo solo le frequenze congiunte V F

e. è possibile ricavare le frequenze congiunte conoscendo solo le frequenze marginali V F

f. una distribuzione condizionata relativa si ottiene ponendo a rapporto le frequenze della distribuzione condizionata con il numero complessivo di unità della popolazione V F

[3 affermazioni vere e 3 false]

Test

Nei seguenti quesiti, fai riferimento alla distribuzione doppia di frequenze rappresentata dalla tabella qui a fianco. La distribuzione riguarda il rilevamento dei due caratteri «sesso» (X) e «mezzo utilizzato per recarsi a scuola» (Y) sugli studenti di una classe.

X \ Y	Auto	Bici	Treno
Maschio	3	5	3
Femmina	2	6	4

2 A proposito dei due caratteri oggetto della rilevazione si può dire che:

A X è quantitativo e Y è qualitativo

B X è qualitativo e Y è quantitativo

C entrambi i caratteri X e Y sono qualitativi

D entrambi i caratteri X e Y sono quantitativi

3 Quanti sono gli studenti della classe?

A 18 B 20 C 22 D 23

4 La distribuzione marginale di X è:

A

X	Frequenza
Maschio	11
Femmina	12

B

X	Frequenza
Maschio	12
Femmina	11

C

X	Frequenza
Maschio	10
Femmina	14

D Nessuna delle precedenti

5 In quale delle seguenti tabelle è evidenziata la distribuzione di Y, condizionata alla modalità «maschio» di X?

A

X \ Y	Auto	Bici	Treno
Maschio	3	5	3
Femmina	2	6	4

B

X \ Y	Auto	Bici	Treno
Maschio	3	5	3
Femmina	2	6	4

C

X \ Y	Auto	Bici	Treno
Maschio	3	5	3
Femmina	2	6	4

D In nessuna delle precedenti

Problemi

6 ESERCIZIO GUIDATO

Sono stati intervistati i 20 studenti di una classe e su ciascuno sono stati rilevati congiuntamente due caratteri X: il sesso (M = maschio, F = femmina) e Y: il numero di ore dedicate mediamente allo studio in una giornata. Si sono ottenuti i dati riportati nella seguente tabella:

X	M	M	F	M	M	F	M	F	M	F	F	M	F	M	F	M	F	M	M	F
Y	1	2	2	4	3	1	2	3	3	4	2	3	4	2	1	2	3	3	3	4

a. Costruisci la tabella a doppia entrata che rappresenta la distribuzione delle frequenze congiunte e marginali dei due caratteri X e Y.

b. Determina la distribuzione di X, condizionata alla modalità «3 ore di studio al giorno» e la corrispondente distribuzione condizionata relativa.

c. Determina la distribuzione di Y, condizionata alla modalità «femmina» e la corrispondente distribuzione condizionata relativa.

---

a. Completa la tabella a doppia entrata sottostante (la riga e la colonna dei totali rappresentano le distribuzione marginali):

Il numero di maschi che studiano un'ora al giorno

X \ Y	1 ora	2 ore	3 ore	4 ore	Totale
M	1	....	....	....	....
F	2	....	....	....	....
Totale	3	....	....	....	20

Il numero di femmine che studiano un'ora al giorno

Le somme per riga e per colonna devono essere uguali al numero complessivo di studenti intervistati

b. La distribuzione di X, condizionata alla modalità $Y = 3$, è rappresentata dalla quarta colonna della tabella compilata al punto **a.**; la corrispondente distribuzione relativa si ottiene rapportando le frequenze assolute al totale. Le distribuzioni richieste sono quindi rappresentate nella tabella seguente, che devi completare.

$X \mid Y = 3$	Frequenze assolute	Frequenze relative
M	5	....
F	....	....
Totale	7	1

c. La distribuzione di Y, condizionata alla modalità $X = F$, è rappresentata dalla terza riga della tabella compilata al punto **a.**; la corrispondente distribuzione relativa si ottiene rapportando le frequenze assolute al totale. Le distribuzioni richieste sono quindi rappresentate nella tabella seguente, che devi completare.

$Y \mid X = F$	1 ora	2 ore	3 ore	4 ore	Totale
Frequenze assolute	2	....	....	....	9
Frequenze relative	....	....	....	....	1

7 Sono stati intervistati 20 individui e su ciascuno sono stati rilevati congiuntamente due caratteri: il luogo di residenza, indicato con X (Sud = S, Centro = C, Nord = N) e la condizione a livello lavorativo, indicata con Y (Occupato = O, Disoccupato = D). Si sono ottenuti i dati nella seguente tabella:

X	S	N	S	C	S	N	C	C	N	N	N	S	C	S	C	N	S	C	N	N
Y	O	O	D	O	D	D	D	O	O	D	O	D	O	D	D	O	O	D	O	O

Costruisci la tabella che rappresenta la distribuzione doppia di frequenze di X e Y e individua le distribuzioni marginali dei due caratteri X e Y.

8 Ai fini di un controllo della qualità, si è esaminato un campione di 15 pezzi meccanici prodotti in un'azienda. Si sono rilevati congiuntamente due caratteri: la linea di produzione da cui proviene il pezzo, indicata con X (1 = il pezzo proviene dalla linea 1, 2 = il pezzo proviene dalla linea 2), e la conformità del pezzo (C = conforme, N = non conforme), indicata con Y. Si sono ottenuti i risultati nella tabella seguente.

X	1	1	2	1	2	1	1	2	1	2	1	2	1	1	2
Y	C	N	C	C	C	C	N	N	C	C	C	N	N	C	C

Costruisci la tabella che rappresenta la distribuzione doppia di frequenze di X e Y e individua le distribuzioni marginali dei due caratteri X e Y.

9 **Realtà e modelli** Nella seguente tabella sono stati riportati i dati (ufficiali) relativi alle età e al sesso dei deputati della XV legislatura.

Composizione della camera: distinzione dei deputati per età e sesso
[Fonte: Camera dei deputati, 10 marzo 2008]

Età	Donne	Uomini
25-29	0	1
30-39	11	35
40-49	33	143
50-59	45	202
60 e oltre	20	137

a. Determina le distribuzioni marginali dei due caratteri «Età» e «Sesso».
b. Tra i deputati della XV legislatura, qual era la percentuale degli uomini?
c. Tra i deputati donne della XV legislatura, qual era la percentuale di coloro che avevano un'età superiore o uguale a 40 anni?
d. Tra i deputati aventi un'età inferiore a 40 anni, qual era la percentuale delle donne?

[**b.** Circa 83%; **c.** circa 90%; **d.** circa 23%]

10 Nella seguente tabella sono riportati i risultati ottenuti allo scrutinio di giugno di prima superiore dagli ex-studenti della scuola media «Ponti» e i giudizi conseguiti all'esame di terza media.

Alunni provenienti dalla scuola media «Ponti»

	Promossi	Respinti	Sospesi con 1 debito	Sospesi con 2 debiti	Sospesi con 3 debiti
Sufficiente	0	0	0	0	0
Buono	1	0	1	3	5
Distinto	8	1	1	1	0
Ottimo	15	0	0	0	0

a. Determina le distribuzioni marginali dei due caratteri «Giudizio conseguito all'esame di terza media» e «Risultato ottenuto allo scrutinio di giugno della prima superiore».

b. Tra gli studenti provenienti dalla scuola media «Ponti», qual è la percentuale di coloro che sono stati promossi a giugno?

c. Tra gli studenti provenienti dalla scuola media «Ponti» promossi a giugno in prima superiore, qual è la percentuale di coloro che hanno conseguito all'esame di terza media un giudizio «Ottimo»?

d. Tra gli studenti della scuola media «Ponti» che hanno conseguito all'esame di terza media un giudizio «Buono» o «Distinto», qual è la percentuale di quelli che sono stati rinviati agli esami di settembre?

[**b.** Circa 67%; **c.** 62,5%; **d.** circa 52%]

11 Sui cinque componenti di una famiglia sono stati rilevati due caratteri: il peso (X) e l'altezza (Y). Si è ottenuta la distribuzione doppia di frequenze rappresentata in tabella.

X \ Y	172 cm	175 cm	183 cm
65 kg	1	0	1
70 kg	0	1	0
75 kg	0	2	0

a. Determina le distribuzioni marginali di X e di Y.

b. Determina la distribuzione condizionata di X, rispetto alla modalità «175 cm».

c. Calcola la media e la mediana delle altezze.

d. Calcola la media e la mediana dei pesi.

e. Calcola il peso medio dei componenti della famiglia che hanno un'altezza di 175 cm.

[**c.** Media = 176 cm, mediana = 175 cm; **d.** media = mediana = 70 kg; **e.** circa 73,3 kg]

12 Sui 1000 dipendenti di un'azienda si sono rilevati due caratteri: il sesso (X) e lo stipendio mensile, in euro (Y). I risultati ottenuti sono riportati nella seguente tabella.

X \ Y	[1000, 1500)	[1500, 2000)	[2000, 2500)	[2500, 3000)
Uomini	290	200	50	10
Donne	250	180	15	5

a. Determina le distribuzioni marginali di X e Y.

b. Determina la distribuzione di Y, condizionata alla modalità «donne».

c. Determina lo stipendio medio dei dipendenti uomini.

d. Determina lo stipendio medio dei dipendenti donne.

e. Determina lo stipendio medio di tutti i dipendenti dell'azienda. [**c.** 1550 euro; **d.** 1500 euro; **e.** 1527,50 euro]

2. Dipendenza e indipendenza statistica

Teoria p. 50

Esercizi introduttivi

Test

All'uscita di un cinema viene chiesto agli spettatori di esprimere un parere sul gradimento del cast del film (X) e sul gradimento della colonna sonora (Y). Le distribuzioni semplici di X e Y sono riportate nelle tabelle seguenti:

X	Molto	Abbastanza	Poco
Frequenza	120	150	30

Y	Molto	Abbastanza	Poco
Frequenza	100	120	80

Supponendo che X e Y siano *indipendenti*, rispondi ai quesiti 13-14-15.

13 Quanti degli intervistati hanno espresso gradimento «molto» sia per il cast del film sia per la colonna sonora?

A 40 B 50 C 60 D Le informazioni date sono insufficienti per stabilirlo

14 Quanti degli intervistati hanno espresso gradimento «abbastanza» sia per il cast del film sia per la colonna sonora?

A 40 B 50 C 60 D Le informazioni date sono insufficienti per stabilirlo

15 Quanti degli intervistati hanno espresso gradimento «poco» sia per il cast del film sia per la colonna sonora?

A 8 B 9 C 10 D Le informazioni date sono insufficienti per stabilirlo

16 Se due caratteri X e Y sono indipendenti, allora la tabella a doppia entrata che rappresenta la distribuzione doppia delle frequenze dei due caratteri è tale che:

A la distribuzione marginale di X è uguale alla distribuzione marginale di Y
B le distribuzioni condizionate relative di X sono tutte uguali alla distribuzione marginale relativa di X
C la distribuzione marginale relativa di X è uguale alla distribuzione marginale relativa di Y
D le distribuzioni condizionate relative di X sono uguali alle distribuzioni condizionate relative di Y

17 Quale delle seguenti condizioni **non** è sufficiente a garantire l'indipendenza tra due caratteri X e Y?

A Le frequenze teoriche di indipendenza coincidono con le frequenze osservate
B L'indice $\chi^2 = 0$
C Le distribuzioni marginali di X e Y sono uguali tra loro
D L'indice χ^2 normalizzato $= 0$

18 Due caratteri X e Y risultano tra loro indipendenti se e solo se l'indice χ^2 risulta:

A uguale a 0 B uguale a 1 C negativo D maggiore di 1

19 Due caratteri X e Y sono stati rilevati su un campione di 100 persone. X ha assunto 3 modalità, mentre Y ha assunto 5 modalità. Per normalizzare l'indice χ^2 occorre dividerlo per:

A 100 B 200 C 400 D 500

20 **Vero o falso?**

a. se X è indipendente da Y allora Y è indipendente da X V F
b. data la distribuzione doppia di frequenze di due caratteri X e Y, l'indice χ^2 assume valore zero se e solo se X e Y sono indipendenti V F
c. l'indice χ^2 dipende dalla numerosità del collettivo V F
d. per normalizzare l'indice χ^2 è sufficiente dividerlo per il numero complessivo n di unità della popolazione V F
e. l'indice χ^2 può assume valori negativi se X e Y non sono connessi V F
f. l'indice χ^2 normalizzato assume sempre valori appartenenti all'intervallo [0, 1] V F

[4 affermazioni vere e 2 false]

Problemi

21 Completa la tabella, sapendo che X e Y sono indipendenti.

X \ Y	y_1	y_2	y_3	Totale
x_1	.....	.....	.....	60
x_2	.....	.....	.....	30
x_3	.....	.....	.....	.....
Totale	10	50	.....	100

22 Completa la tabella, sapendo che X e Y sono indipendenti.

X \ Y	y_1	y_2	y_3	Totale
x_1	5	.....	.....	50
x_2	.....	.....	.....	40
x_3	.....	.....	.....	.....
Totale	10	60	.....	.....

23 Completa la tabella e verifica che X e Y sono indipendenti.

X \ Y	y_1	y_2	y_3	Totale
x_1	9	54	27	.....
x_2	6	36	18	.....
x_3	5	30	15	.....
Totale	.....	.....	.....	.....

Realtà e modelli

24 **Sondaggio.** Su un campione di abitanti di un comune viene effettuato un sondaggio per sapere se sono favorevoli o meno alla costruzione di una pista ciclabile. Su ogni intervistato viene rilevato il sesso (X) e il parere a proposito della costruzione della pista (Y). I risultati ottenuti sono riassunti nella seguente tabella.

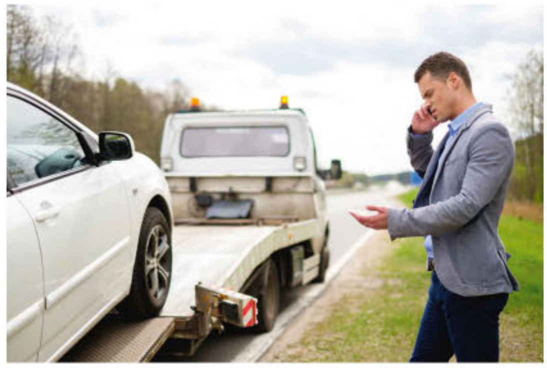
X \ Y	Favorevole	Contrario	Totale
F	30	20	50
M	35	15	50
Totale	65	35	100

a. Costruisci la tabella teorica di indipendenza e verifica che X e Y non sono indipendenti.
b. Valuta il grado di connessione, calcolando l'indice χ^2 normalizzato.
c. Supponi ora che tutti i maschi del campione intervistato si siano dichiarati favorevoli alla pista ciclabile e tutte le femmine contrarie. In tale ipotesi, quanto vale l'indice χ^2 normalizzato?

[**b.** $\chi^2 \simeq 1{,}1$; χ^2 normalizzato $\simeq 0{,}01$: l'indice di connessione è molto basso, prossimo a 0, pertanto X e Y sono molto vicini alla condizione teorica di indipendenza statistica; **c.** l'indice chi-quadrato normalizzato è uguale a 1, ovvero la connessione è la massima possibile]

25 **Incidenti.** Un'agenzia assicurativa ha raccolto in tabella i dati relativi ai 500 incidenti stradali dell'ultimo anno, classificandoli in base al sesso X del responsabile (F = femmina, M = maschio) e alla gravità Y dell'incidente.

X \ Y	Moderato	Grave	Totale
F	180	72	252
M	82	166	248
Totale	262	238	500



a. Costruisci la tabella delle frequenze teoriche di indipendenza e confrontala con la tabella delle frequenze osservate. I due caratteri esaminati X e Y sono indipendenti?
b. Valuta il grado di connessione, calcolando l'indice χ^2 normalizzato.
c. Si può affermare che, rispetto agli uomini, le donne generalmente guidano con maggiore prudenza e sono responsabili di incidenti di minore gravità? E si può sostenere che, rispetto agli uomini, sono responsabili di un minor numero di incidenti?

[**b.** $\chi^2 \simeq 73{,}76$; χ^2 normalizzato $\simeq 0{,}15$]

26 **Traffico e condizioni meteo.** In una città sono stati osservati giornalmente le condizioni meteo (X) e il livello di traffico (Y) per un periodo di un anno. I dati ricavati sono riassunti nella seguente tabella.

X \ Y	Basso	Medio	Alto	Totale
Sereno	80	28	12	120
Variabile	32	92	30	154
Perturbato	8	25	58	91
Totale	120	145	100	365

a. Calcola l'indice χ^2.
b. Normalizza l'indice χ^2 e determina una percentuale che esprima la misura della connessione tra i due caratteri.
c. Puoi sostenere che, statisticamente, quanto più il tempo è bello tanto più il traffico è ridotto?

[**a.** Circa 152,31; **b.** circa 21%]

3. Correlazione e regressione

Teoria p. 54

Esercizi introduttivi

27 **Vero o falso?**

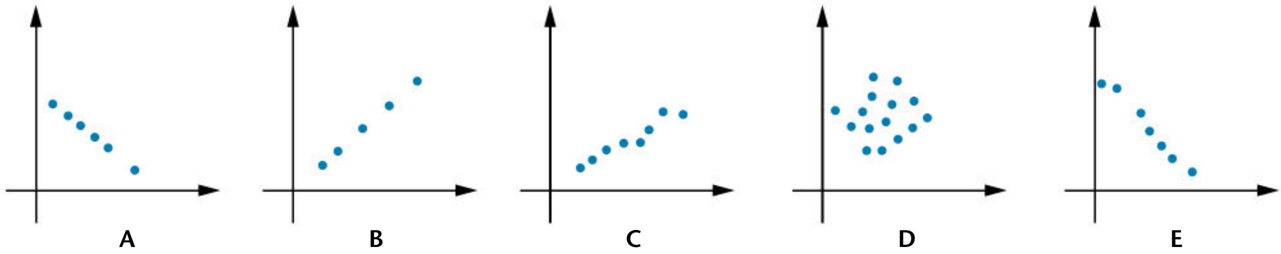
a. la covarianza assume sempre valori non negativi V F
b. il coefficiente di correlazione lineare è sempre compreso tra -1 e 1, potendo anche essere uguale a ± 1 V F
c. se la variabile Y è una funzione lineare della variabile X, allora il coefficiente di correlazione lineare tra X e Y può essere uguale soltanto a 1 V F
d. se il coefficiente di correlazione lineare tra due variabili X e Y è uguale a 1, allora la variabile Y è una funzione lineare della variabile X V F
e. se il coefficiente di correlazione lineare tra due variabili X e Y è 0, allora X e Y non possono essere legate né da una relazione di tipo lineare, né da una relazione di tipo diverso V F
f. la covarianza dipende dall'unità di misura dei dati V F
g. il coefficiente di correlazione lineare dipende dall'unità di misura dei dati V F
h. se la covarianza di una coppia di variabili X e Y è negativa, allora la retta di regressione che modella il comportamento di Y in funzione di X è decrescente V F
i. la bontà di adattamento della retta di regressione dipende dal grado di correlazione tra le due variabili V F

[5 affermazioni vere e 4 false]

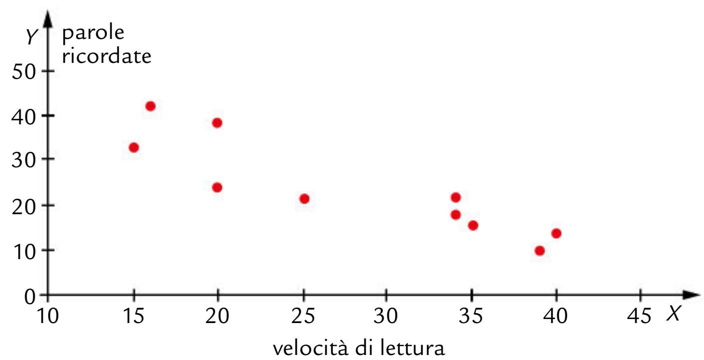
Interpretazione di grafici

28 Associa a ciascuno dei seguenti grafici il corrispondente coefficiente di correlazione lineare, scelto tra i seguenti:

a. $r = 0{,}01$ **b.** $r = 0{,}9$ **c.** $r = 1$ **d.** $r = -0{,}85$ **e.** $r = -1$



29 A un gruppo di studenti del liceo Ariosto è stato chiesto di leggere un brano dai Promessi Sposi. Il grafico riporta i punteggi ottenuti sui test di velocità di lettura (Asse X ascissa) e il punteggio test sul numero di parole ricordate (Asse Y ordinata). Quale tra i valori seguenti potrebbe avere il coefficiente di correlazione per le due variabili X, Y?



A $-0{,}8$ B 0 C $0{,}3$ D -1

(Olimpiadi della statistica 2013)

Test

30 Date due variabili X e Y, in quale dei seguenti casi la retta di regressione che esprime Y in funzione di X si adatterà meglio ai dati?

A Se il coefficiente di correlazione lineare è prossimo a 0
B Se il coefficiente di correlazione lineare è prossimo a 0,5
C Se il coefficiente di correlazione lineare è prossimo a $-0{,}5$
D Se il coefficiente di correlazione lineare è prossimo a 1

31 Se la covarianza di due variabili X e Y è positiva, allora il coefficiente angolare della retta di regressione che esprime Y in funzione di X è:

A certamente positivo
B certamente nullo
C certamente negativo
D le informazioni date non sono sufficienti per stabilirlo

32 Date due variabili X e Y, di deviazione standard σ_X e σ_Y e covarianza σ_{XY}, l'espressione $\dfrac{\sigma_{XY}}{\sigma_X \cdot \sigma_Y}$ rappresenta:

A il coefficiente angolare della retta di regressione che esprime Y in funzione di X
B il coefficiente angolare della retta di regressione che esprime X in funzione di Y
C il coefficiente di correlazione lineare tra le due variabili X e Y
D l'indice di connessione χ^2 tra le due variabili

33 Date due variabili X e Y, di deviazione standard σ_X e σ_Y e covarianza σ_{XY}, l'espressione $\dfrac{\sigma_{XY}}{\sigma_X^2}$ rappresenta:

A il coefficiente angolare della retta di regressione che esprime Y in funzione di X
B il coefficiente angolare della retta di regressione che esprime X in funzione di Y
C il coefficiente di correlazione lineare tra le due variabili X e Y
D l'indice di connessione χ^2 tra le due variabili

34 Il coefficiente di correlazione lineare tra due variabili X e Y, ciascuna delle quali ha media nulla, è $r = 0{,}7$. Calcola l'equazione della retta di regressione che esprime il comportamento di Y in funzione di X sapendo che le varianze di X e di Y sono rispettivamente uguali a 0,25 e a 0,36. $[y = 0{,}84x]$

35 Il coefficiente di correlazione lineare tra due variabili X e Y, aventi rispettivamente medie uguali a 2 e 3, è $-0{,}8$. Inoltre, le varianze di X e Y sono rispettivamente 0,64 e 0,25. Determina l'equazione della retta di regressione di Y in funzione di X.

$\left[y = -\frac{1}{2}x + 4\right]$

Problemi sulla regressione lineare

36 **ESERCIZIO GUIDATO**

Si sono rilevati su cinque individui l'età (X) e la pressione arteriosa (Y), e si sono ottenuti i dati in tabella:

Età (X)	Pressione (Y)
25	124
32	130
40	135
55	145
70	170

a. Calcola il coefficiente di correlazione lineare e valuta se sussiste una relazione lineare tra X e Y.

b. Determina l'equazione della retta di regressione che esprime Y in funzione di X.

---

a. Completa anzitutto la tabella seguente.

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
25	124	3100	625	15376
32	130	....	....	....
40	135	....	....	....
55	145	....	....	....
70	170	....	....	....
$\sum x_i = 222$	$\sum y_i = 704$	$\sum x_i y_i =$	$\sum x_i^2 =$	$\sum y_i^2 = 100\,426$

Per determinare il coefficiente di correlazione lineare r devi calcolare preliminarmente le deviazioni standard σ_X, σ_Y di X e Y e la loro covarianza σ_{XY}. La sequenza di calcoli da svolgere è la seguente:

$$\bar{x} = \frac{\sum x_i}{5} = \frac{.....}{5} = 44{,}4 \qquad \bar{y} = \frac{\sum y_i}{5} = \frac{.....}{5} = 140{,}8$$ **Medie**

$$\sigma_X^2 = \frac{\sum x_i^2}{5} - \bar{x}^2 = \frac{.....}{5} - (.....)^2 = ... \qquad \sigma_Y^2 = \frac{\sum y_i^2}{5} - \bar{y}^2 = \frac{.....}{5} - (.....)^2 = ...$$ **Varianze**

$$\sigma_{XY} = \frac{\sum x_i y_i}{5} - \bar{x} \cdot \bar{y} = \frac{..........}{5} - .... \cdot .... = 255{,}48$$ **Covarianza**

$$r = \frac{\sigma_{XY}}{\sigma_X \cdot \sigma_Y} = \frac{..........}{\sqrt{..........} \cdot \sqrt{..........}} \simeq ....$$ **Coefficiente di correlazione lineare**

Poiché r è prossimo a 1 puoi concludere che ....

b. Calcola il coefficiente angolare della retta di regressione, arrotondando il risultato alla seconda cifra decimale.

$$m = \frac{\sigma_{XY}}{\sigma_X^2} = \frac{..........}{..........} \simeq .... \quad \text{oppure} \quad m = \frac{\sum x_i \cdot y_i - n\bar{x} \cdot \bar{y}}{\sum x_i^2 - n\bar{x}^2} = \frac{..........}{..........} \simeq ....$$

L'equazione della retta di regressione è perciò:

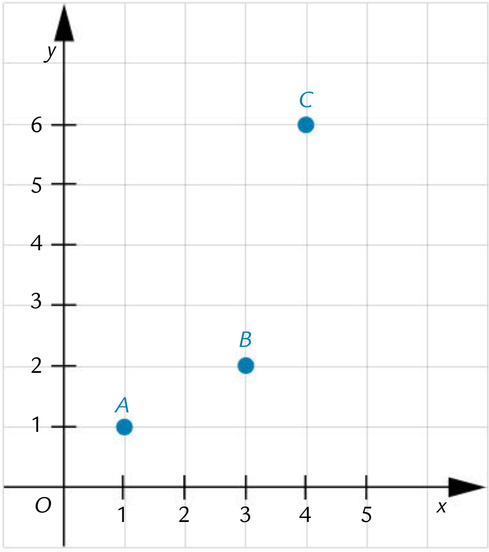
$$\underbrace{y - .....}_{y - \bar{y}} = \underbrace{.........}_{m} \underbrace{(x - ....)}_{x - \bar{x}} \quad \text{da cui:} \quad y = ...\, x + ...$$

[**a.** $r \simeq 0{,}975$; **b.** $y = 0{,}97x + 97{,}74$]

Interpretazione di grafici

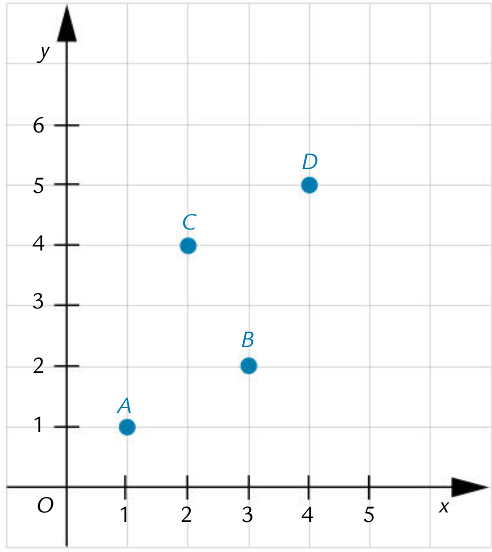
Determina l'equazione della retta di regressione relativa ai punti rappresentati nei seguenti grafici.

37



$\left[y = \frac{3}{2}x - 1\right]$

38



$\left[y = x + \frac{1}{2}\right]$

39 Sono stati rilevati congiuntamente due caratteri X e Y su tre individui e si sono ottenuti i dati rappresentati nella seguente tabella.

x_i	1	2	3
y_i	4	8	10

a. Rappresenta graficamente i punti (x_i, y_i).

b. Calcola le coordinate del loro baricentro.

c. Calcola il coefficiente di correlazione lineare della distribuzione rappresentata. Fornisci il risultato arrotondato a meno di un millesimo.

d. Scrivi l'equazione della retta di regressione che esprime Y in funzione di X.

$\left[\textbf{b.} \left(2, \frac{22}{3}\right); \textbf{c.}\ 0{,}982; \textbf{d.}\ y = 3x + \frac{4}{3}\right]$

40 Sono stati rilevati congiuntamente due caratteri X e Y su quattro individui e si sono ottenuti i dati rappresentati nella seguente tabella.

x_i	1	2	3	4
y_i	2	3	5	6

a. Rappresenta graficamente i punti (x_i, y_i).

b. Calcola le coordinate del loro baricentro.

c. Calcola il coefficiente di correlazione lineare della distribuzione rappresentata. Fornisci il risultato arrotondato a meno di un centesimo.

d. Scrivi l'equazione della retta di regressione che esprime Y in funzione di X.

$\left[\textbf{b.} \left(\frac{5}{2}, 4\right); \textbf{c.}\ 0{,}99; \textbf{d.}\ y = \frac{7}{5}x + \frac{1}{2}\right]$

Realtà e modelli

41 **Visite a un sito Internet.** Nei primi quattro mesi dalla creazione del blog di Paolo il numero di visitatori mensili è stato quello rappresentato in tabella.

Mesi x_i trascorsi a partire dalla creazione del blog	Numero y_i di visitatori
1	130
2	160
3	190
4	230

a. Rappresenta graficamente i punti $(x_i, \; y_i)$.
b. Calcola le coordinate del loro baricentro.
c. Calcola il coefficiente di correlazione lineare della distribuzione rappresentata. Fornisci il risultato arrotondato a meno di un millesimo.
d. Scrivi l'equazione della retta di regressione che esprime il numero di visitatori in funzione dei mesi trascorsi dalla creazione del blog.
e. Sulla base del modello stabilito, determina il primo mese a partire dal quale il numero di visitatori sarà superiore a 500.

[**b.** $\left(\frac{5}{2}, \frac{355}{2}\right)$; **c.** 0,997; **d.** $y = 33x + 95$; **e.** il tredicesimo]

42 **Consumo di elettricità.** In un dato paese il consumo di elettricità nel settore dei trasporti urbani e ferroviari (espresso in TWh, cioè in miliardi di kWh) ha fatto registrare negli anni 2006-2016 i dati rappresentati nella seguente tabella.

Anno	2006	2008	2010	2014	2016
Numero x_i che identifica l'anno di osservazione	1	3	5	9	11
Consumo y_i	8,6	10,4	12,2	11,9	12,1

a. Calcola le coordinate del baricentro della nuvola di punti (x_i, y_i).
b. Calcola il coefficiente di correlazione lineare della distribuzione rappresentata, fornendo il risultato arrotondato a meno di un centesimo.
c. Scrivi l'equazione della retta di regressione che esprime il consumo in funzione del numero che identifica l'anno di osservazione. Fornisci l'equazione con i coefficienti arrotondati a meno di un centesimo.
d. Sulla base del modello trovato al punto precedente, stima il consumo nel settore dei trasporti urbani e ferroviari nel 2018. Fornisci il risultato arrotondato alla prima cifra decimale.

[**a.** $(5{,}8, \; 11{,}04)$; **b.** 0,81; **c.** $y = 0{,}3x + 9{,}28$; **d.** 13,2 TWh]

43 **Vendite di computer portatili.** Un punto vendita di una catena specializzata nel settore dell'informatica ha monitorato il numero di computer portatili venduti in un anno, dividendolo in base al prezzo. I dati raccolti sono riassunti nella seguente tabella.

Prezzo in euro dei portatili (x_i)	Numero di portatili venduti (y_i)
400	250
900	220
1100	200
1200	190
1400	168

a. Calcola le coordinate del baricentro della nuvola di punti (x_i, y_i).
b. Calcola il coefficiente di correlazione lineare della distribuzione rappresentata, fornendo il risultato arrotondato a meno di un centesimo.
c. Scrivi l'equazione della retta di regressione che esprime il numero di portatili venduti in funzione del prezzo. Fornisci l'equazione con i coefficienti arrotondati a meno di un centesimo.
d. Sulla base del modello trovato al punto precedente, stima il numero di portatili aventi il costo di 600 euro che il punto vendita può avere venduto in quell'anno.

[**a.** $(1000, \; 205{,}6)$; **b.** $-0{,}99$; **c.** $y = -0{,}08x + 286{,}29$; **d.** 238]

44 **Clienti di un'azienda.** Un'azienda ha studiato l'evoluzione del numero dei suoi clienti, a partire dall'apertura. Ha così ricavato i dati della seguente tabella.

Anno	2011	2012	2013	2014	2015
Anni x_i trascorsi a partire dall'apertura	1	2	3	4	5
Clienti y_i	120	126	130	135	142

a. Calcola le coordinate del baricentro della nuvola di punti (x_i, y_i).
b. Calcola il coefficiente di correlazione lineare della distribuzione rappresentata, fornendo il risultato arrotondato a meno di un millesimo.
c. Scrivi l'equazione della retta di regressione che esprime il numero di clienti in funzione degli anni trascorsi dall'apertura.
d. Sulla base del modello trovato al punto precedente, stima il numero di clienti che l'azienda potrà avere nel 2020.

[**a.** $(3, 130{,}6)$; **b.** 0,996; **c.** $y = 5{,}3x + 114{,}7$; **d.** 168]

45 **Titolo azionario.** Nella tabella è riportata l'evoluzione del valore di un titolo azionario nel corso del quinquennio 2012-2016.

Anno	2012	2013	2014	2015	2016
Numero x_i che identifica l'anno di osservazione	1	2	3	4	5
Valore (in euro)	13	13,50	13,10	13,70	14

a. Rappresenta in un diagramma i dati. In base al grafico, ritieni che vi sia una significativa correlazione lineare tra gli anni trascorsi e il valore del titolo azionario?
b. Calcola il coefficiente di correlazione lineare della distribuzione rappresentata. Il valore del coefficiente di correlazione lineare conferma la risposta data al punto precedente?
c. Determina l'equazione della retta di regressione lineare che esprime il valore in euro del titolo azionario in funzione degli anni trascorsi.
d. Sulla base del modello trovato al punto precedente, quale prevedi sarà il valore del titolo nel 2022?

[**b.** $r \simeq 0{,}84$; **c.** $y = 0{,}22x + 12{,}8$ (coefficienti approssimati); **d.** 15,22 (circa)]

Regressione non lineare

Determina la funzione potenza del tipo $y = ax^b$ che interpola i dati (i coefficienti dei risultati sono arrotondati alla seconda cifra decimale).

46

x_i	1	2	3	4
y_i	1	7	20	40

[$y = 1{,}04x^{2{,}67}$]

47

x_i	1	4	8	10
y_i	10	15	20	25

[$y = 9{,}65x^{0{,}37}$]

48

x_i	1	2	3	4
y_i	15	8	6	1

[$y = 19{,}62x^{-1{,}68}$]

49

x_i	1	3	5	8	10
y_i	8	12	15	20	25

[$y = 7{,}55x^{0{,}48}$]

50

x_i	1	3	5	8	10
y_i	12	8	6	4	1

[$y = 16{,}13x^{-0{,}87}$]

51

x_i	6	10	20	25	30
y_i	40	52	85	100	120

[$y = 11{,}4x^{0{,}68}$]

Determina la funzione esponenziale del tipo $y = ab^x$ che interpola i dati (i coefficienti dei risultati sono arrotondati alla seconda cifra decimale).

52

x_i	1	2	3	4
y_i	2	8	20	40

[$y = 0{,}89(2{,}69)^x$]

53

x_i	1	2	3	4
y_i	1	6	15	40

[$y = 0{,}39(3{,}31)^x$]

54

x_i	2	5	8	10
y_i	4	15	22	30

[$y = 3{,}11(1{,}27)^x$]

55

x_i	1	2	3	4	5
y_i	4	8	10	15	30

[$y = 2{,}66(1{,}59)^x$]

56

x_i	5	8	10	12	15
y_i	2	4	8	16	20

[$y = 0{,}62(1{,}28)^x$]

57

x_i	1	5	10	15	20
y_i	1	8	12	18	25

[$y = 1{,}87(1{,}16)^x$]

Realtà e modelli

58 **Decadimento di una sostanza radioattiva.** Si ha un campione di 1000 g di una sostanza radioattiva, di cui si vogliono studiare le proprietà di decadimento. Ogni giorno, per una settimana, viene misurata la quantità di sostanza radioattiva rimasta. Si ottengono i dati riportati nella seguente tabella.

Giorno	0	1	2	3	4	5	6	7
Quantità di sostanza radioattiva (in g)	1000	894	798	712	642	573	513	455

a. Rappresenta i dati in un diagramma cartesiano.

b. Determina la funzione di regressione esponenziale che interpola i dati, arrotondando i coefficienti alla seconda cifra decimale.

Utilizzando il modello determinato in **b.**, rispondi ai seguenti ulteriori quesiti.

c. Determina il tempo di dimezzamento della sostanza, cioè il tempo necessario perché rimanga la metà della sostanza radioattiva originaria. Esprimi il risultato in giorni, ore e minuti.

d. Determina quanta sostanza radioattiva rimarrà dopo 10 giorni. Esprimi il risultato arrotondato a un numero intero.

e. Determina dopo quanto tempo la sostanza radioattiva rimanente sarà inferiore a 50 g. Esprimi il risultato in giorni, ore e minuti.

[**b.** $y = 999{,}56(0{,}89)^x = 999{,}56e^{-0{,}11x}$; **c.** circa 6 giorni, 7 ore e 8 minuti; **d.** circa 333 g; **e.** dopo 27 giorni, 5 ore e 31 minuti]

59 **Domanda di un bene.** Si effettua un'analisi statistica allo scopo di stabilire la funzione domanda di un bene che deve essere immesso sul mercato. I risultati dell'indagine sono riassunti nella seguente tabella in cui, per ogni prezzo indicato, viene indicata la corrispondente quantità del bene che i consumatori sarebbero disposti a comprare.

Prezzo (x, in euro)	15	16	17	18	19	20
Quantità domandata (y, in migliaia di unità)	16	15	13	10	7	6

a. Rappresenta i dati in un diagramma cartesiano.

b. Supponendo che la funzione domanda sia una funzione potenza, determina la funzione potenza di regressione che interpola i dati, arrotondando i coefficienti alla seconda cifra decimale.

c. In base al modello della funzione domanda costruito al punto **b.**, stima quale sarà la domanda nel caso in cui venga fissato un prezzo di 21 euro.

d. In base al modello della funzione domanda costruito al punto **b.**, stima in corrispondenza di quale prezzo la domanda sarà di 8000 unità.

[**b.** $y = \dfrac{378\,754{,}92}{x^{3{,}67}}$; **c.** circa 5319 unità; **d.** circa 18 euro e 79 centesimi]

Esercizi di riepilogo

Esercizi interattivi

60 **Vero o falso?**

Nella seguente tabella sono riportati i dati relativi ai laureati di quattro corsi di laurea di tre atenei diversi, suddivisi per corso (Y) e per ateneo (X), nello stesso anno accademico.

X \ Y	Psicologia	Architettura	Matematica	Lettere classiche
A	51	35	15	18
B	47	26	13	14
C	65	23	9	25

a. i caratteri statistici sono entrambi quantitativi V F

b. la media dei laureati per corso di laurea nell'ateneo A è 29,75 V F

c. la media dei laureati in Architettura è 30 V F

d. la distribuzione marginale di X è V F

X	Frequenza
A	119
B	100
C	122

e. la distribuzione di Y, condizionata alla modalità «B» di X, è quella evidenziata in giallo nella seguente tabella V F

X \ Y	Psicologia	Architettura	Matematica	Lettere classiche
A	51	35	15	18
B	47	26	13	14
C	65	23	9	25

f. nell'ateneo B si laurea in Architettura circa il 25% del totale dei laureati in Architettura dei tre atenei V F

g. nell'ateneo C si laurea in Lettere classiche circa il 7% del totale dei laureati dei tre atenei V F

h. i due caratteri statistici sono indipendenti V F

i. possiamo affermare che il grado di connessione tra i due caratteri è di circa il 4,7%. V F

61 **Completamento.**

Completa la seguente tabella, sapendo che i caratteri X e Y sono statisticamente indipendenti.

X \ Y	y_1	y_2	y_3	Totale
x_1				20
x_2				
x_3				50
Totale	10	30		100

62 **Vero o falso?**

Nella seguente tabella sono riportati i dati relativi a un gruppo di 10 studenti, per i quali si mette in relazione il voto Y (in trentesimi) conseguito a un esame universitario con il numero di ore X che ciascuno dichiara di aver dedicato allo studio.

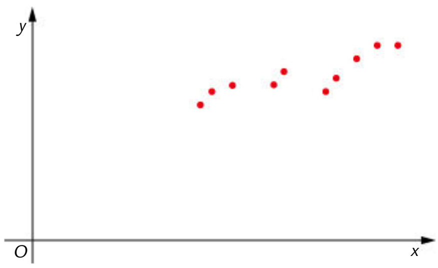
studente	1	2	3	4	5	6	7	8	9	10	TOTALE
X (ore)	120	150	100	85	160	180	125	90	170	145	1325
Y (voto)	24	25	24	21	28	30	26	23	30	23	254

Stabilisci se le seguenti affermazioni sono vere o false.

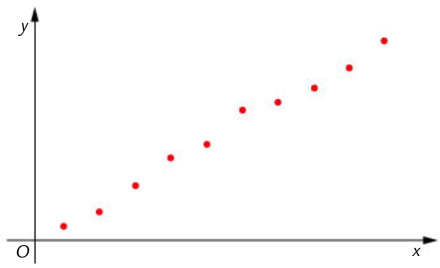
a. ciascuno studente ha studiato in media 132,5 ore V F
b. il voto medio conseguito è 23,8 V F
c. la varianza di X è superiore a 1000 V F
d. la deviazione standard di Y vale circa 2,9 V F
e. la covarianza è molto prossima a 0 V F
f. il coefficiente di correlazione lineare è prossimo all'86% del massimo valore possibile V F
g. la retta di regressione interseca l'asse Y in prossimità dell'origine V F

63 **Associazione.**

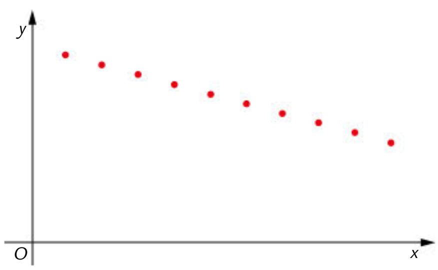
Associa a ogni distribuzione di punti il corrispondente coefficiente di correlazione lineare.



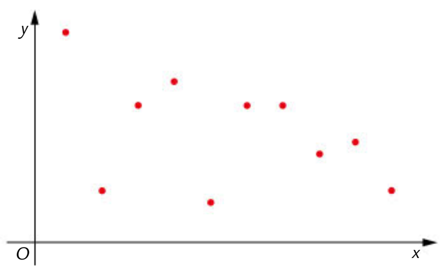
a.



b.



c.



d.

A. $r = -0{,}44$ **B.** $r = -1$ **C.** $r = 0{,}85$ **D.** $r = 0{,}97$

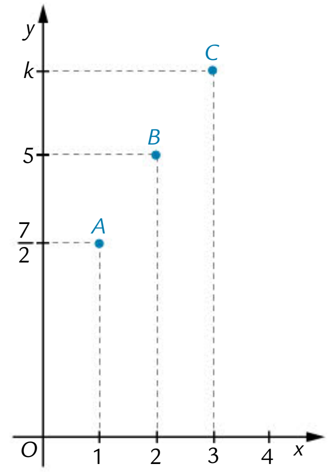
64 **Completamento.**

Completa la seguente tabella, inserendo i valori dei parametri statistici mancanti, con quattro cifre decimali; la cifra intera e quelle decimali devono essere separate dalla virgola.

	X	Y
	1	3
	2	5
	4	9
	6	12
media		
varianza		
deviazione standard		
covarianza		
coefficiente di correlazione lineare		

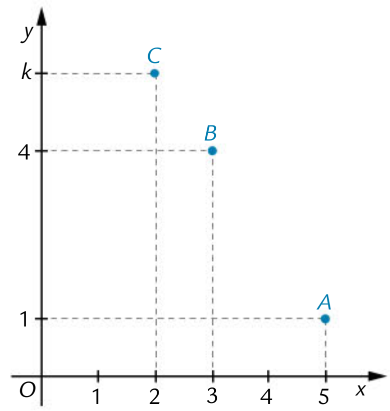
Interpretazione di grafici

65 Per quale valore di k il coefficiente di correlazione lineare del grafico a dispersione in figura risulta uguale a 1?



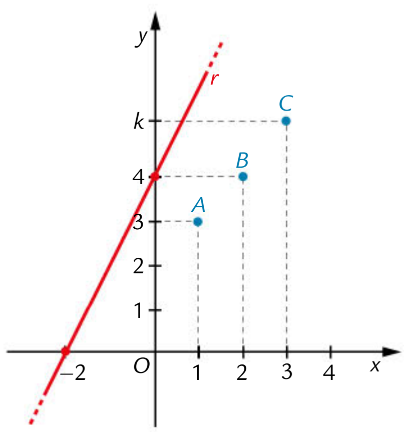
$\left[k = \frac{13}{2}\right]$

66 Per quale valore di k il coefficiente di correlazione lineare del grafico a dispersione in figura risulta uguale a -1?



$\left[k = \frac{11}{2}\right]$

67 Per quale valore di k la retta di regressione relativa ai tre punti A, B e C risulta parallela alla retta r?



$[k = 7]$

68 La tabella che segue si riferisce al numero di addetti X e al fatturato medio mensile Y, espresso in migliaia di euro, di una ditta che opera con 5 filiali (indicate con A, B, C, D, E).

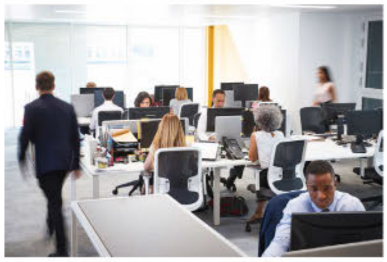
Filiale	A	B	C	D	E
Addetti (X)	50	20	10	5	3
Fatturato (Y)	100	80	45	55	40

a. Calcola la covarianza di X e Y.
b. Stabilisci, tramite un opportuno indice, se un modello lineare interpreta bene il legame tra X e Y.

[**a.** $\sigma_{XY} = 362{,}6$; **b.** il coefficiente di correlazione lineare r vale circa 0,93: essendo prossimo a 1, possiamo concludere che il modello lineare interpreta bene il legame tra X e Y]

69 In cinque aziende sono stati rilevati il numero dei dipendenti (X) e lo stipendio annuo dei dirigenti (Y), espresso in migliaia di euro. I risultati sono riportati nella seguente tabella.

Azienda	1	2	3	4	5
Numero di dipendenti	16	12	30	24	22
Stipendio dei dirigenti	40	35	65	52	48



a. Determina la media e la mediana del numero di dipendenti.
b. Determina le media e la mediana degli stipendi dei dirigenti.
c. Calcola il coefficiente di correlazione lineare tra X e Y. È ragionevole assumere un modello lineare per interpretare la relazione tra X e Y?

[**a.** Media = 20,8, mediana = 22; **b.** media = 48, mediana = 48; **c.** circa 0,99, dunque è ragionevole un modello lineare]

70 Il fatturato (in milioni di euro) di una azienda negli ultimi 4 anni ha avuto l'andamento riportato nella seguente tabella.

Anno	Fatturato
1	8,5
2	12,2
3	15,4
4	14,5

a. Qual è stato l'incremento medio annuo del fatturato, espresso in percentuale (ovvero quale incremento percentuale, applicato uguale tutti e quattro gli anni, lascia invariato il fatturato finale)?
b. Assumendo un modello lineare, quale previsione è possibile fare circa il fatturato nell'anno successivo all'ultimo considerato?
c. Stabilisci, tramite un indice opportuno, se il modello lineare utilizzato è attendibile.

[**a.** Circa 19,5%; **b.** la retta di regressione ha equazione $y = 2{,}12x + 7{,}35$, da cui si stima per il quinto anno un fatturato di 17,95 milioni di euro; **c.** l'indice di correlazione lineare vale circa 0,89, quindi il modello è attendibile]

71 Un'azienda ha acquistato nel 2013 un macchinario del costo di 50 000 euro. Il valore del macchinario negli anni successivi, nel mercato dell'usato, è risultato quello indicato nella seguente tabella.

Anno	2013	2014	2015	2016	2017
Anni x_i trascorsi a partire dall'acquisto	0	1	2	3	4
Valore y_i del macchinario in euro	50 000	42 000	36 000	32 000	26 500

a. Calcola le coordinate del baricentro della nuvola di punti (x_i, y_i).
b. Calcola il coefficiente di correlazione lineare della distribuzione rappresentata, fornendo il risultato arrotondato a meno di un centesimo.
c. Scrivi l'equazione della retta di regressione che esprime il valore del macchinario in funzione degli anni trascorsi dall'acquisto.
d. Sulla base del modello trovato al punto precedente, stima il valore del macchinario nel 2019 e stabilisci a partire da quale anno il valore del macchinario sarà inferiore ai 10 000 euro.

[**a.** $(2,\ 37\,300)$; **b.** $-0{,}99$; **c.** $y = -5700x + 48\,700$; **d.** nel 2019, il macchinario varrà 14 500 euro e il valore scenderà sotto i 10 000 euro nel 2020]

72 Su un campione di 100 pesci di un lago vengono rilevati due caratteri: la lunghezza in cm (X) e la razza (Y). Tra i pesci del campione si rileva che:

- 10 sono tinche di lunghezza minore di 40 cm;
- 20 sono tinche di lunghezza compresa tra 40 e 60 cm;
- 15 sono carpe di lunghezza compresa tra 40 e 60 cm;
- 30 sono carpe di lunghezza compresa tra 60 e 70 centimetri;
- 5 sono orate di lunghezza compresa tra 40 e 60 cm;
- 10 sono orate di lunghezza compresa tra 60 e 70 cm;
- non ci sono orate di lunghezza inferiore a 40 cm;
- non ci sono tinche di lunghezza superiore a 60 cm.

a. Sulla base delle informazioni date, completa la tabella a doppia entrata qui sotto, che rappresenta la distribuzione doppia di frequenze dei due caratteri X e Y (nella tabella sono già stati inseriti alcuni valori come esempio).

Razza (Y) / Lunghezza (X)	Tinche	Carpe	Orate	Totale
$0 < X < 40$			0	
$40 < X < 60$		15		
$60 < X < 70$				
Totale				100

b. Determina le distribuzioni marginali di X e Y.
c. Determina la distribuzione di X, condizionata alla modalità «carpe» di Y.
d. Calcola la lunghezza media dei pesci del campione analizzato.
e. Calcola la lunghezza media delle carpe rinvenute nel campione analizzato.
f. Calcola l'indice di connessione chi-quadrato di X e Y, normalizzalo e stabilisci se X e Y sono indipendenti o connessi.

[**d.** 50 cm; **e.** circa 52,7 cm; **f.** $\chi^2 \simeq 31{,}06$ e χ^2 normalizzato $\simeq 0{,}16$, dunque X e Y non sono indipendenti, ma il grado di connessione è basso, quantificabile nell'ordine del 16% della massima connessione possibile]

73 La seguente tabella rappresenta l'evoluzione della percentuale di automobili monovolume vendute in un anno da un concessionario nel periodo 2013-2017.

Anno	2013	2014	2015	2016	2017
x_i (numero che identifica l'anno di osservazione)	1	2	3	4	5
y_i (numero che esprime la percentuale di monovolume vendute)	10	11	12	15	16

a. Rappresenta la nuvola di punti (x_i, y_i) e determina le coordinate del loro baricentro.
b. Calcola il coefficiente di correlazione lineare della distribuzione rappresentata, fornendo il risultato arrotondato a meno di un centesimo.
c. Scrivi l'equazione della retta di regressione che esprime y in funzione di x.
d. Sulla base del modello trovato al punto precedente:
- stima la percentuale di monovolume vendute dal concessionario nel 2018;
- stima quale sarà il primo anno a partire dal quale la percentuale di monovolume vendute sarà superiore al 20%.

[**a.** (3, 12,8); **b.** circa 0,98; **c.** $y = 1{,}6x + 8$; **d.** 17,6%, nell'anno 2020]

Esercizi *più*

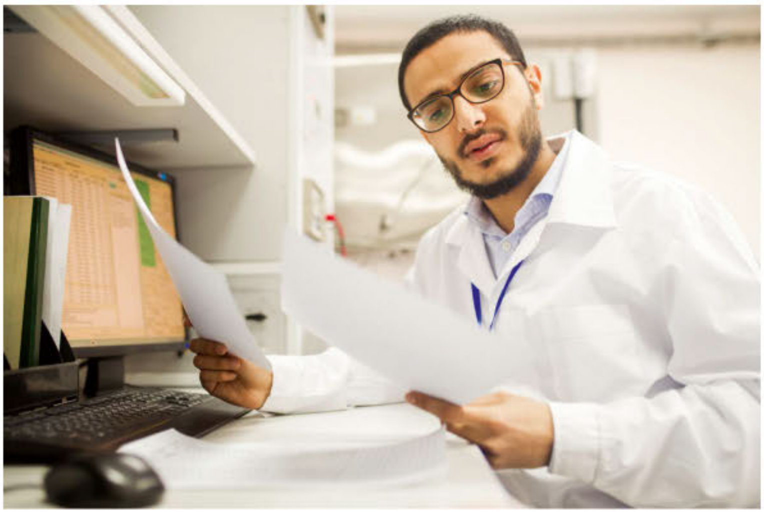
74 **E se?** Considera le due variabili statistiche X e Y. Le deviazioni standard σ_X e σ_Y, la covarianza σ_{XY}, il coefficiente di correlazione lineare r e il coefficiente angolare m della retta di regressione che esprime Y in funzione di X sono reciprocamente dipendenti, cioè è possibile esprimere i cinque parametri in funzione di tre di essi, opportunamente scelti. Verificalo, completando la tabella qui sotto (la prima riga è completata come esempio).

σ_X	σ_Y	σ_{XY}	r	m
σ_X	σ_Y	σ_{XY}	$\dfrac{\sigma_{XY}}{\sigma_X\sigma_Y}$	$\dfrac{\sigma_{XY}}{\sigma_X^2}$
σ_X	σ_Y	.....	r	.....
σ_X	σ_Y	.....	.....	m
σ_X	.....	σ_{XY}	r	.....
σ_X	.....	.....	r	m
.....	σ_Y	σ_{XY}	r	.....
.....	σ_Y	σ_{XY}	.....	m
.....	σ_Y	.....	r	m
.....	.....	σ_{XY}	r	m

▸ Se fossero noti σ_X, σ_{XY} ed m, sarebbe possibile esprimere σ_Y ed r in funzione dei tre parametri noti?

75 In un laboratorio viene ritrovato un foglio, in parte rovinato, che riguarda un'analisi di regressione. Si riesce a leggere che la varianza di X è $\sigma_X^2 = 4$ e che le rette di regressione che esprimono Y in funzione di X e X in funzione di Y hanno rispettivamente equazioni:

$$x - y + 4 = 0 \quad \text{e} \quad 9x - 4y - 24 = 0$$



Determina:
- **a.** i valori medi delle variabili X e Y;
- **b.** la deviazione standard di Y;
- **c.** il coefficiente di correlazione lineare tra X e Y.

$\left[\textbf{a. } \bar{x} = 8,\ \bar{y} = 12;\ \textbf{b. } \sigma_Y = 3;\ \textbf{c. } \dfrac{2}{3}\right]$

Statistica bivariata, correlazione e regressione

1 I gestori di un sito internet hanno chiesto ai loro visitatori di esprimere un giudizio (Y) sui contenuti pubblicati. I giudizi degli utenti, divisi in tre fasce d'età (X), sono sintetizzati nella seguente tabella.

a. Dopo avere determinato le distribuzioni marginali di X e Y, costruisci la tabella teorica di indipendenza.
b. Dopo avere verificato che non sussiste indipendenza tra i due caratteri, valuta con un indice opportuno la connessione tra i due caratteri.

X \ Y	Negativo	Positivo	Ottimo
Giovani	0	0	25
Adulti	0	22	30
Anziani	20	13	0

2 Le tabelle seguenti mostrano la distribuzione degli studenti dell'istituto «Ponti» in base al profitto scolastico e al colore degli occhi:

Rendimento	Insufficiente	Sufficiente	Eccellente	Totale
Numero studenti	65	533	117	715

Colore occhi	Verdi	Azzurri	Marroni	Totale
Numero studenti	55	220	440	715

Tenendo conto che i caratteri «rendimento scolastico» e «colore degli occhi» sono ovviamente indipendenti, determina:
a. il numero degli studenti dell'istituto «Ponti» che hanno un rendimento scolastico eccellente e gli occhi azzurri;
b. la percentuale degli studenti del «Ponti» che hanno gli occhi marroni e il cui rendimento è (appena) sufficiente.

3 Osserva la tabella. Dopo avere rappresentato nel piano cartesiano il corrispondente grafico di dispersione, determina le coordinate del baricentro dei punti e l'equazione della retta di regressione.

x_i	1	2	3	4
y_i	5	4	2	1

4 Su quattro soggetti di sesso maschile sono stati rilevati l'età (X, in anni) e il numero di pulsazioni sotto sforzo al minuto (Y). Sono stati ottenuti i risultati riportati nella seguente tabella.

x_i (età)	20	30	40	50
y_1 (pulsazioni)	195	182	175	170

a. Calcola il coefficiente di correlazione lineare, fornendo il risultato arrotondato a meno di un centesimo.
b. Scrivi l'equazione della retta di regressione che esprime le pulsazioni in funzione dell'età.
c. Sulla base del modello trovato al punto precedente stima il numero di pulsazioni sotto sforzo, al minuto, di un uomo di 65 anni. Arrotonda il risultato a un numero intero.

Valutazione

Esercizio	1	2	3	4	**Totale**
Punteggio massimo	$1{,}5 \cdot 2 = 3$	$1 \cdot 2 = 2$	2	$1 \cdot 3 = 3$	10
Punteggio ottenuto					

Tempo indicativo: 1 h

Risposte p. 218

Risolvere problemi e costruire modelli

1 I docenti di tre diverse commissioni d'esame hanno attribuito rispettivamente i seguenti voti medi: 28, 26 e 25. Determina il voto medio di tutti gli esaminati, sapendo che ciascuna commissione aveva rispettivamente 50, 30 e 80 studenti. [26,125]

2 Per un anno sono monitorati i prezzi al mercato di un particolare tipo di mele. I dati sono stati così sintetizzati:

- per 90 giorni 1,50 euro al kg
- per 60 giorni 1,70 euro al kg
- per 70 giorni 1,75 euro al kg
- per 50 giorni 2,00 euro al kg
- per 95 giorni 1,60 euro al kg

Qual è stato il prezzo medio annuo del tipo di mela preso in considerazione? [Circa 1,68 euro]

3 La media aritmetica di quattro numeri naturali consecutivi è uguale a 12,5. Qual è il più piccolo di essi? [11]

4 La media dei voti in un compito in classe di matematica è stata 7. Le ragazze della classe, in tutto 18, hanno ottenuto una media del 7,5, mentre la media dei voti dei ragazzi è stata 6.

a. Da quanti alunni è composta la classe?

b. Al compito in classe successivo, la media dei voti della classe resta invariata, ma questa volta è la media dei voti dei ragazzi a essere uguale a 7,5. Qual è stata la media dei voti delle ragazze? [**a.** 27; **b.** 6,75]

5 Si è effettuata un'indagine su un campione di persone, chiedendo loro quante volte mangiano fuori casa in una settimana. Dai risultati dell'indagine si è calcolato che in media gli intervistati mangiano fuori casa 3,3 volte alla settimana. Inoltre, il 20% ha risposto che non mangia mai fuori casa, il 30% che mangia fuori casa 3 volte la settimana, il 25% che mangia fuori casa 5 volte la settimana, il 5% che mangia fuori casa 7 volte la settimana; i restanti hanno affermato che mangiano fuori casa n volte la settimana. Quanto vale n? [4]

Interpretare grafici e dati

6 La tabella a doppia entrata qui di seguito riporta le intenzioni di voto alle prossime elezioni politiche di 500 intervistati, suddivisi in tre fasce d'età.

	Partito A	Partito B	Partito C	Totale
Giovani	.....	45	.....	150
Adulti	85	80	35	200
Anziani	.....	25	.....	150
Totale	.....	150	.....	500

a. Completa la tabella, sfruttando le informazioni seguenti e impostando opportunamente un sistema di primo grado in due incognite:

- il numero dei giovani intenzionati a votare per il partito A è uguale al numero degli anziani intenzionati a votare per il partito C;
- il numero degli anziani disposti a votare per il partito A è il doppio del numero dei giovani disposti a votare per il partito C.

b. Tra coloro che intendono votare per il partito C, qual è la percentuale dei giovani? È la stessa percentuale di coloro che, tra i giovani, sono intenzionati a votare per il partito C?

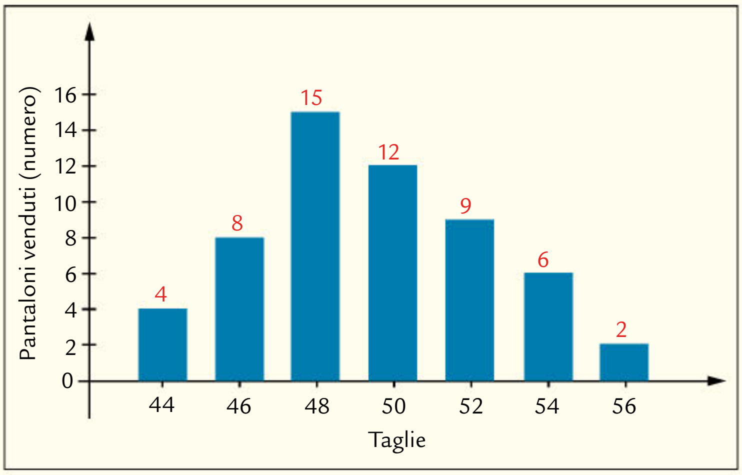
7 **Dal giornale** In un articolo apparso su un quotidiano alcuni anni fa venivano riportati i seguenti dati sul reddito pro capite (cioè sul reddito medio) degli abitanti delle province del Lazio nel 2001.

Provincia	Reddito pro capite (euro)
Frosinone	12 416
Latina	11 343
Rieti	13 399
Roma	16 017
Viterbo	13 300

L'articolo concludeva, calcolando la media aritmetica semplice dei redditi pro capite per provincia, che il reddito pro capite degli abitanti del Lazio nel 2001 era di 13 295 euro. Sei d'accordo con questa conclusione?

8 Un commerciante ha rilevato le taglie di pantaloni venduti in un mese e ha ottenuto i risultati rappresentati nel seguente diagramma a barre.

a. Quanti pantaloni ha venduto il commerciante in quel mese?
b. Qual è la taglia media venduta?
c. Qual è la taglia mediana?
d. Qual è la taglia modale?
e. Qual è la percentuale di pantaloni di taglia superiore alla 48 venduti?



[**a.** 56; **b.** circa 49,4; **c.** 50; **d.**48; **e.** circa il 51,79%]

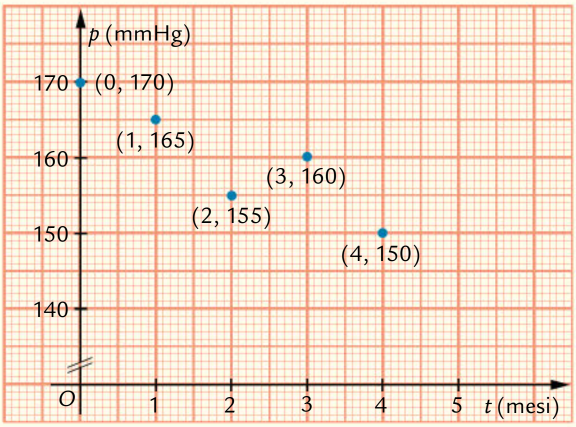
9 In una zona che conta 1400 abitanti, alcuni cittadini sono rimasti senza elettricità in seguito a un violento temporale. Di questi 1400 abitanti, il 70% abita nel centro della città mentre gli altri abitano in periferia. In seguito al temporale, il 5% di coloro che abitano in centro città e il 75% di coloro che abitano in periferia sono rimasti privi di energia elettrica.
Completa la seguente tabella, che rappresenta la distribuzione doppia di frequenze dei due caratteri X (zona di residenza) e Y (comportamento dell'energia elettrica in seguito al temporale) e le distribuzioni marginali di X e Y.

Y (comportamento dell'energia elettrica) / X (zona di residenza)	Regolare	Soggetto a interruzione	Totale
Centro città			
Periferia			
Totale			

a. Determina la distribuzione di X, condizionata alla modalità «Regolare» di Y.
b. Determina la distribuzione di Y, condizionata alla modalità «Periferia» di X.
c. Costruisci la tabella teorica di indipendenza di X e Y.
d. Specifica se sussiste indipendenza tra X e Y; in caso negativo, valuta con un indice opportuno la connessione tra i due caratteri. [**d.** I due caratteri sono connessi e calcolando l'indice di connessione chi-quadrato normalizzato si ottiene che la connessione è circa il 53,5 % della massima possibile]

10 Il signor Bianchi soffre d'ipertensione arteriosa: su prescrizione del medico, da quattro mesi assume un farmaco per stabilizzare la pressione sanguigna. Gli effetti della cura sono deducibili dal grafico: in ascissa è riportato il tempo (in mesi) trascorso dall'inizio della cura, in ordinata la pressione sistolica (in mmHg).

a. Calcola il coefficiente di correlazione lineare e intepreta il risultato: che cosa significa il fatto che il coefficiente di correlazione trovato è negativo? E il fatto che è molto vicino a -1?
b. Determina l'equazione della retta di regressione lineare e rappresenta quest'ultima nel piano cartesiano. Essa conferma (oppure smentisce) la tua valutazione precedente?
c. Stando al modello, quale sarà la pressione sistolica del signor Bianchi dopo 6 mesi dall'inizio della cura? E quando raggiungerà il livello di 130 mmHg?
d. Il modello precedentemente applicato può considerarsi sicuramente adeguato nel breve e medio periodo; ti sembra adatto a prevedere l'evoluzione della pressione anche nel lungo periodo?



[**a.** $r = -0,9$; **b.** $y = -4,5x + 169$; **c.** 142 mmHg, dopo 8 mesi e 20 giorni (considerando un mese di 30 giorni)]

11 In un ospedale si sono rilevati per sei giorni consecutivi (a partire da lunedì) i nuovi casi registrati di infezioni relative a un certo virus, ottenendo i dati riportati qui sotto:

x_i (giorno di osservazione)	1	2	3	4	5	6
y_i (numero di nuove infezioni registrate)	2	1	4	6	6	10

a. Rappresenta la nuvola di punti (x_i, y_i) e determina le coordinate del loro baricentro G.
b. Calcola il coefficiente di correlazione lineare della distribuzione rappresentata, fornendo il risultato arrotondato a meno di un centesimo.
c. Scrivi l'equazione della retta di regressione che esprime il numero di nuovi casi registrati in funzione del giorno di osservazione. Arrotonda i coefficienti dell'equazione della retta a meno di un decimo.
d. Supponendo che il modello trovato sia valido per predire l'andamento delle infezioni nella settimana successiva a quella della rilevazione dei dati, stima il numero di nuove infezioni che si possono prevedere per il mercoledì di tale settimana. Arrotonda il risultato a un numero intero.

$\left[\mathbf{a.}\ G\left(\frac{7}{2}, \frac{29}{6}\right); \mathbf{b.}\ 0{,}94 ; \mathbf{c.}\ y = 1,6x - 0{,}9; \mathbf{d.}\ 15\right]$

Esporre, argomentare e dimostrare

12 Il tasso di disoccupati rispetto alla forza lavoro, nel 2004, era del 6% al Centro, del 4% al Nord e del 15% al Sud. Stabilisci se questi dati sono sufficienti per calcolare il tasso medio nazionale di disoccupazione nel 2004. In caso negativo, spiega come si dovrebbe procedere per effettuare il calcolo correttamente.

13 Si vuole valutare l'amministrazione di un'azienda nel periodo 2011-2016. In particolare, si vuole stabilire se si sono ottenuti risultati migliori durante l'amministrazione del dottor Rossi nel triennio 2011-2013, oppure durante l'amministrazione del dottor Verdi nel triennio 2014-2016. Nella tabella seguente sono riportati fatturato, profitto (in migliaia di euro) e numero di dipendenti in ciascuno dei sei anni del periodo esaminato.

Anno	Fatturato	Profitto	Dipendenti
2011	2351	783	50
2012	2203	762	48
2013	2178	752	49
2014	2098	740	46
2015	2134	761	44
2016	2156	788	44

Di quali indicatori (numeri indice, rapporti statistici ecc.) ti serviresti per sostenere che è stato migliore il dottor Rossi? E di quali, invece, per favorire la riconferma del dottor Verdi?

14 Giulia e Barbara hanno opinioni diverse circa la retta di regressione relativa a quattro punti, A, B, C, D, di cui A, B, C sono allineati. Giulia è convinta che la retta di regressione debba passare per almeno uno dei quattro punti A, B, C, D; Barbara invece non è d'accordo. Chi ha ragione? Giustifica la risposta, fornendo una dimostrazione o un controesempio opportuno.

15 La media delle età di un insieme di 100 persone è 28. L'età mediana delle 100 persone può essere superiore a 56? Se sì, esibisci un esempio; altrimenti dimostra che è impossibile.
(*Suggerimento*: calcola anzitutto qual è la somma delle età di tutte le persone dell'insieme considerato, quindi ragiona per assurdo)

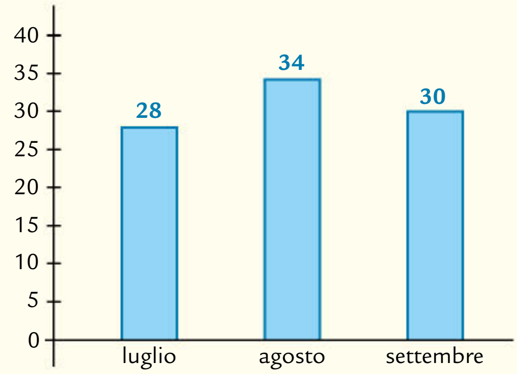
16 In una classe di 20 studenti, la media dei voti in un compito in classe è stata 6,5.

a. È possibile che 15 studenti abbiano preso 9?

b. È possibile che 14 studenti abbiano preso voti minori o uguali a 4?

In caso affermativo, esibisci un esempio; altrimenti dimostra che è impossibile.

1 Il grafico riporta il numero di *e-book reader* (lettori di libri elettronici) venduti nei mesi di luglio, agosto e settembre da un negozio di informatica. Negli altri nove mesi dell'anno lo stesso negozio ha venduto in media 18 e-book reader al mese. Qual è il numero medio mensile di e-book reader venduti in quell'anno dal negozio?



A Circa 31 B Circa 28 C Circa 21 D Circa 24

(Prova Invalsi 2016/2017)

2 Agli alunni di una classe viene chiesto per quanto tempo al giorno, in media, utilizzano la connessione a Internet con i loro dispositivi (PC, tablet, smartphone...). I risultati del sondaggio sono riportati nella seguente tabella.

Minuti di connessione a Internet	Frequenze assolute
Da 0 minuti fino a 60 minuti	2
Più di 60 minuti fino a 120 minuti	4
Più di 120 minuti fino a 180 minuti	12
Più di 180 minuti fino a 300 minuti	8

Quale tra le seguenti espressioni permette di calcolare il tempo medio giornaliero di connessione a Internet degli alunni della classe?

A $\dfrac{30 + 90 + 150 + 240}{4}$

B $\dfrac{60 \cdot 2 + 120 \cdot 4 + 180 \cdot 12 + 300 \cdot 8}{2 + 4 + 12 + 8}$

C $\dfrac{30 \cdot 2 + 90 \cdot 4 + 150 \cdot 12 + 240 \cdot 8}{2 + 4 + 12 + 8}$

D $\dfrac{2 + 4 + 12 + 8}{4}$

(Prova Invalsi 2016/2017)

3 Con una bilancia si è misurata 10 volte la massa di una lastra di alluminio ottenendo le seguenti misure in kilogrammi:

10,55	10,76	10,60	10,87	10,64	10,67	10,84	10,46	10,55	10,70

Quale fra i seguenti indici statistici è quello più adatto a rappresentare la massa della lastra di alluminio?

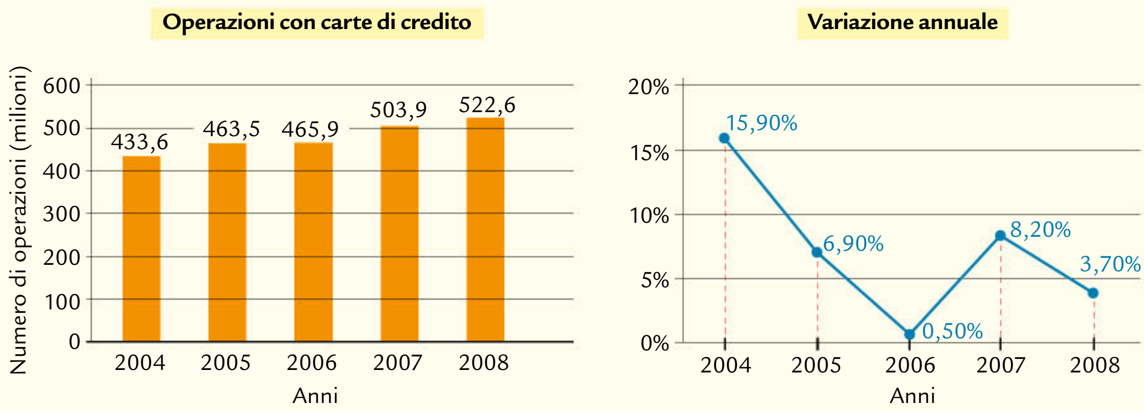
A La moda
B La media aritmetica
C La varianza
D Lo scarto quadratico medio (o deviazione standard)

(Prova Invalsi 2016/2017)

4 Lo sfruttamento medio della capacità ricettiva di un albergo è uguale all'88% durante i tre mesi estivi e al 44% durante i rimanenti mesi dell'anno. Qual è lo sfruttamento medio relativo all'intero anno?

A 55% B 50% C 46% D 80%

5 Osserva i seguenti grafici relativi alle operazioni effettuate con carte di credito dal 2004 al 2008. (*Fonte*: Osservatorio sulle carte di credito. Assofin-Crif Decision Solutions-Gfk Eurisko)



Indica se ciascuna delle seguenti affermazioni è vera (V) o falsa (F).

a. Il numero di operazioni effettuate con carte di credito è diminuito dal 2004 al 2006, poi è aumentato e successivamente è di nuovo diminuito fino al 2008. V F

b. I due grafici sono in contraddizione perché il primo mostra una continua crescita nel tempo, mentre il secondo no. V F

c. L'aumento del numero di operazioni effettuate con carte di credito che si è avuto dal 2006 al 2007 è stato superiore all'aumento che si è avuto dal 2007 al 2008. V F

d. Nel 2006 il numero di operazioni effettuate con carte di credito è quasi azzerato. V F

(Prova Invalsi 2012)

6 La professoressa Rossi vuole verificare il livello delle conoscenze in scienze nelle classi 1A e 1B. Decide di somministrare lo stesso test nelle due classi. Elaborando i punteggi dei test ottiene i seguenti risultati:

	Classe 1A	Classe 1B
Media aritmetica	6,5	6,5
Scarto quadratico medio (o deviazione standard)	1,1	2,3

La professoressa chiede a Martina, una sua alunna di 1B, di commentare i risultati ottenuti dagli alunni delle due classi. Martina afferma che i risultati indicano che gli alunni delle due classi hanno lo stesso livello medio di conoscenze, ma gli studenti della 1A hanno ottenuto complessivamente punteggi più vicini alla media. Martina ha ragione? Scegli una delle due risposte e completa la frase.

☐ Sì, perché .... ☐ No, perché ....

(Prova Invalsi 2012)

7 Un automobilista percorre i primi 120 km di un certo percorso alla velocità media di 60 km/h e i successivi 120 km alla velocità media di 120 km/h. Qual è la sua velocità media durante l'intero percorso?

A 70 km/h B 80 km/h C 90 km/h D 100 km/h

(Prova Invalsi 2013)

8 Le temperature massime giornaliere registrate in una data città nei giorni di una settimana, a partire dal lunedì, sono le seguenti:

Lu 29° C Ma 30° C Me 32° C Gi 31° C Ve 28° C Sa 30° C Do 30° C

Quale delle seguenti affermazioni è *falsa*?

A La temperatura media è quella che si è registrata il martedì
B La temperatura modale è quella che si è registrata il mercoledì
C La temperatura mediana è quella che si è registrata il giovedì
D La temperatura mediana è uguale alla temperatura modale

9 La media degli studenti promossi da una certa scuola, nei quattro anni 1999-2002, è stata di 325 studenti all'anno, mentre nei cinque anni 1999-2003 la media è stata superiore del 20% rispetto al precedente intervallo temporale. Quanti studenti sono stati promossi dalla scuola nel 2003?

A 650 B 600 C 455 D 390

10 La tabella riporta, per alcune regioni, il numero di incidenti stradali.

Regioni	Numero di incidenti	Lunghezza della rete stradale (km)
Umbria	4520	6639
Sicilia	10 283	20 833
Sardegna	5562	12 132

Fonte: Elaborazione su dati ACI

a. Basandoti solo sulle informazioni presenti in tabella, in quale delle tre regioni era più rischioso circolare nel 2010?

Risposta: ....

b. Nel 2010 in Italia si sono verificati 292 762 incidenti e la lunghezza della rete stradale era di 303 365 km. Laura afferma che in Sicilia il rischio di incidenti nel 2010 era maggiore di quello che si aveva in Italia nello stesso anno.

Laura ha ragione?

Scegli una delle due risposte e completa la frase

☐ Laura ha ragione, perché in Sicilia: ....

....

☐ Laura non ha ragione, perché in Sicilia: ....

....

(*Prova Invalsi 2016*)

11 La media dei voti ottenuti in un compito in classe è stata 6 e la mediana 5,5. Il professore decide di alzare tutti i voti di mezzo punto. Allora:

A la media resta invariata e la mediana aumenta di 0,5
B la media aumenta di 0,5 e la mediana resta invariata
C sia la media sia la mediana restano invariate
D sia la media sia la mediana aumentano di 0,5

12 Stabilisci quale delle seguenti affermazioni è *falsa*.
La *deviazione standard* di un insieme di dati:

A è sempre un numero non negativo
B può essere uguale a zero
C non cambia se a tutti i dati si aggiunge uno stesso numero
D non cambia se tutti i dati vengono moltiplicati per uno stesso numero

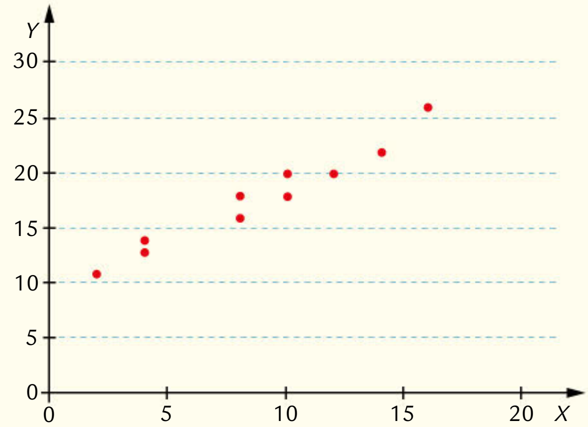
13 Uno studio statistico sulle altezze, misurate in metri, dei componenti di una classe di 20 studenti ha condotto ai seguenti risultati:

moda: 1,87 m   media: 1,76 m   mediana: 1,74 m

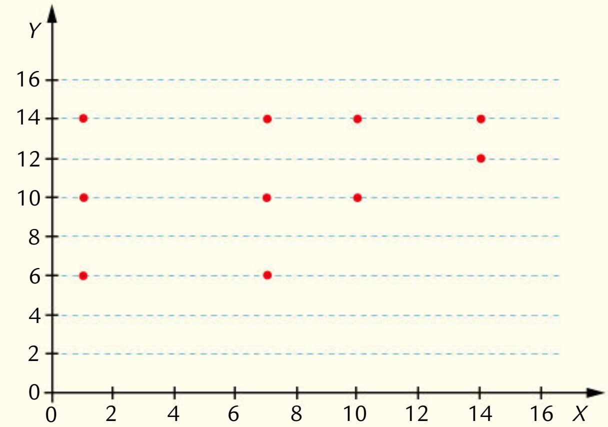
Stabilisci se le seguenti affermazioni sono vere, false o se i dati non sono sufficienti per stabilirlo.

Nessuno studente della classe ha altezza di 1,90 m.	☐ V ☐ F ☐ Non si può stabilire
Nessuno studente della classe ha altezza di 1,85 m.	☐ V ☐ F ☐ Non si può stabilire
La somma delle altezze degli studenti della classe è 35,2 m.	☐ V ☐ F ☐ Non si può stabilire
La metà degli allievi ha altezza minore o uguale a 1,74 m e l'altra metà ha altezza maggiore o uguale a 1,74 m.	☐ V ☐ F ☐ Non si può stabilire
Almeno uno studente della classe ha altezza uguale a 1,76 m.	☐ V ☐ F ☐ Non si può stabilire
Gli studenti della classe aventi altezza superiore a 1,72 m sono meno di 8.	☐ V ☐ F ☐ Non si può stabilire

14 I due grafici riportano i punteggi di due test, uno di inglese (X) e uno di matematica (Y), somministrati a due gruppi di studenti, uno del liceo scientifico e l'altro del liceo artistico. Cosa possiamo dedurre osservando i due grafici?



Liceo scientifico



Liceo artistico

- A La correlazione lineare tra i due test è nulla per gli studenti del liceo scientifico.
- B La correlazione lineare tra i due test è uguale a 1 per gli studenti del liceo scientifico.
- C La correlazione lineare tra i due test è maggiore per gli studenti del liceo scientifico.
- D La correlazione lineare tra i due test è uguale per i due gruppi di studenti.

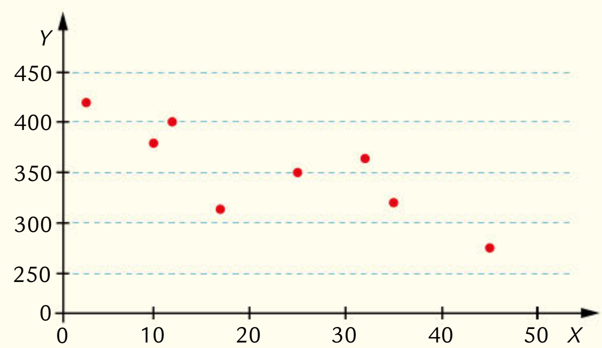
(Adattato da un quesito proposto alle Olimpiadi della Statistica, 2013)

15 In uno studio sulla possibile relazione tra altezza in cm (indicata con X) e peso in kg (indicato con Y) degli scimpanzé adulti, i valori delle due variabili d'interesse, osservate su un campione di 40 animali, hanno dato la seguente espressione della retta di regressione: $Y = 19{,}5 + 0{,}34X$. Di quanto ci aspettiamo che vari il valore di Y se X aumenta di 1? Cioè, di quanti chilogrammi ci aspettiamo che vari il peso di uno scimpanzé, se l'altezza aumenta di un centimetro?

- A 19,5
- B 19,84
- C 0,34
- D Non so
- E Non abbiamo informazioni sufficienti per rispondere

(Olimpiadi della Statistica, 2016)

16 I punti rappresentati nel seguente grafico hanno per coordinate i valori di due variabili rilevate sulle unità di un certo collettivo.

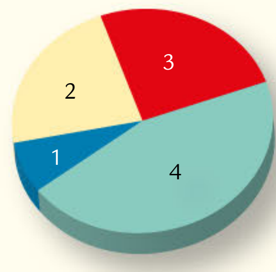


Quale tra i seguenti ritieni che sia il valore del coefficiente di correlazione tra le due variabili? Scegli un'alternativa:

- A 0
- B -1
- C Non so
- D 0,82
- E $-0{,}82$

(Olimpiadi della Statistica, 2016)

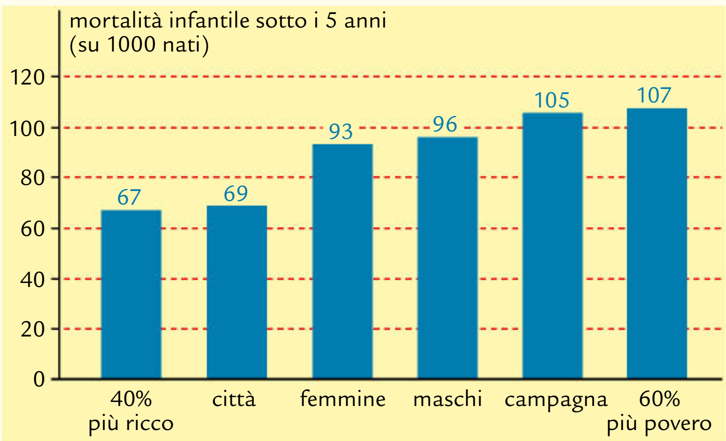
1 A un corso di laurea sono iscritti studenti di 4 nazioni. La composizione percentuale delle varie nazioni è rappresentata nel grafico a torta in figura.
Si sa che i numeri degli iscritti provenienti da tre di queste nazioni sono 12, 36, 40 e che uno dei gruppi costituisce esattamente il 25% del totale. Quanti sono gli studenti del gruppo 4?



A 40 B 48 C 72 D 76

(Test di ingresso, Facoltà di Scienze, 2008)

2 Il seguente diagramma presenta i dati sulla mortalità infantile in base a determinate caratteristiche socio-economiche. Le cifre riportate rappresentano la probabilità di decesso prima del quinto anno di vita, stimate in base a dati raccolti in 63 Paesi in via di sviluppo (anni 1998-2006).



Dall'esame del diagramma **non** si può dedurre che:

A in città la mortalità infantile è del 69 per mille
B in un ambiente rurale la mortalità infantile è maggiore rispetto all'ambiente urbano
C nei Paesi industrializzati si ha una diminuzione della mortalità infantile: da 40 a 6 decessi annui
D la mortalità infantile sotto i 5 anni è maggiore nei maschi rispetto alle femmine
E la mortalità infantile delle femmine è del 9,3%

(Prova di ammissione, corso di laurea in Medicina 2008)

3 Uno studente ha avuto 5 e mezzo ai primi due compiti. Quale voto dovrà raggiungere al terzo compito per ottenere la media del 6?

A 7 B 5 e mezzo C 6 D 6 e mezzo E Non ce la può fare

(Prova di ammissione, corso di laurea in Medicina 2009)

4 Agli studenti di un corso di laurea triennale è stato chiesto di indicare quante lingue straniere sono in grado di comprendere. I risultati dell'indagine sono riportati nella seguente tabella.

Nel complesso degli studenti del primo e secondo anno, qual è la percentuale di quelli che comprendono almeno una lingua straniera?

	Nessuna	Una	Due o più
1° anno	45	51	10
2° anno	41	47	6
3° anno	31	58	11

A 61% B 38% C 49% D 57%

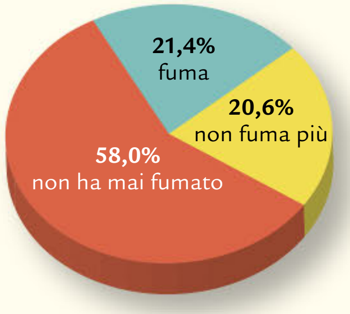
(Test di ingresso per i corsi di laurea scientifici 2008)

5 L'età media dei partecipanti a una festa è di 24 anni. Se l'età media degli uomini è 28 anni e quella delle donne è 18 anni, qual è il rapporto tra il numero degli uomini e quello delle donne?

A $\frac{14}{9}$ B $\frac{9}{14}$ C $\frac{3}{2}$ D $\frac{4}{3}$

(Prova di ammissione, corso di laurea in Ingegneria 2006)

6 Il grafico qui sotto rappresenta l'abitudine al fumo della popolazione di una regione d'Italia nel periodo luglio 1999-giugno 2000, secondo un'indagine ISTAT.



Dall'analisi del grafico si può dedurre che:

A il numero di fumatori nella regione considerata è inferiore percentualmente al resto d'Italia
B la metà della popolazione della regione considerata fuma
C i fumatori rappresentano poco più del 21% della popolazione
D la percentuale di ex fumatori è maggiore della percentuale dei fumatori
E il numero di fumatori è quasi uguale in percentuale a quello delle fumatrici

(Prova di ammissione, corso di laurea in Medicina 2008)

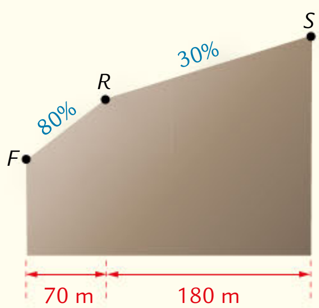
7 Uno studente universitario, dopo aver superato tre esami, ha la media di 28. Nell'esame successivo lo studente prende 20. Qual è la sua media dopo il quarto esame?

A 26 B 24 C 22 D I dati non sono sufficienti per determinare la risposta.

(Test di ingresso per i corsi di laurea scientifici 2008)

---

8 Per arrivare a una cima S si deve percorrere un ripido sentiero FR, con pendenza dell'80%, e poi un sentiero RS, che ha pendenza del 30%. Con riferimento alle misure indicate in figura, determina la pendenza media dell'intero percorso.



A 49% B 55% C 40% D 44% E 52%

(Prova di ammissione, Facoltà di Scienze 2010)

9 Un corridore percorre un sentiero che è per un quarto in piano, metà in salita e un quarto in discesa, impiega in media 6 minuti per km nei tratti in piano e in discesa, 9 minuti per i tratti in salita. La frequenza cardiaca è 110 bpm (battiti per minuto) nei tratti in discesa, 130 in piano e 150 in salita. Qual è la frequenza cardiaca media durante la corsa (all'intero più vicino)?

A 133 B 135 C 138
D 140 E Non ci sono dati sufficienti

(Facoltà di Statistica e informatica per l'Azienda, Università di Trieste, 2016/2017)

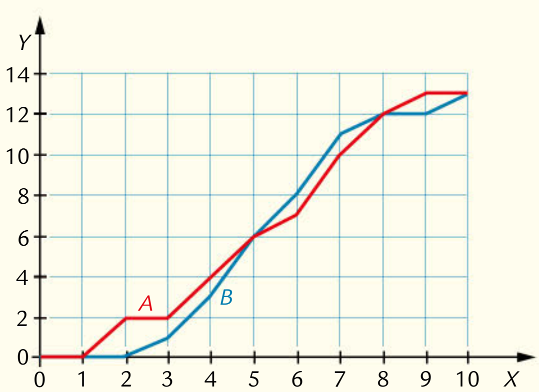
10 Abbiamo 8 numeri di media aritmetica m. Aggiungendo un ulteriore numero x, la media viene aumentata di 3. Possiamo concludere che:

A $x = 3m$ B $x = m + 27$ C $x = \frac{m}{8} + 3$
D $x = 9m + 3$ E $x = 8m + 3$

(Facoltà di Economia, Test Cisia, 2012)

---

11 Tredici concorrenti hanno svolto due prove, ciascuna composta da 10 quesiti, su due temi A e B. Nel grafico è riportato l'esito cumulato relativo a ciascuna delle due prove cioè (sia per il tema A, sia per il tema B), in corrispondenza di ogni valore n sull'asse orizzontale, sull'asse verticale è indicato quanti concorrenti hanno risposto correttamente a un numero di quesiti non superiore a n.



Dal grafico si può dedurre con certezza che:

A il numero di concorrenti che ha risposto correttamente ad almeno 8 quesiti è lo stesso nei due temi
B il numero di concorrenti che ha risposto correttamente a esattamente 8 quesiti è lo stesso nei due temi
C c'è un concorrente che ha risposto correttamente a un solo quesito del tema A
D le risposte corrette sul tema A sono state quante le risposte corrette sul tema B
E il numero di concorrenti che ha risposto correttamente ad al più 5 quesiti è uguale nei due temi

(Facoltà di Economia, Test Cisia, 2013)

Compito di realtà 1

Tasso alcolemico

Una persona che ha assunto alcol viene sottoposta a una misura del tasso alcolemico, dalla quale risulta un valore di 1,2 g/L. Sia x il tempo (misurato in ore) trascorso dalla prima misurazione e y il corrispondente tasso alcolemico misurato al tempo x. I dati relativi a tre misurazioni successive alla prima sono riportanti nella tabella seguente.

x (tempo trascorso, in ore, dalla prima misurazione)	0	1	2	3
y (tasso alcolemico, in g/L)	1,2	1,07	0,96	0,82

1 Stabilisci qual è l'equazione della retta di regressione che esprime y in funzione di x.

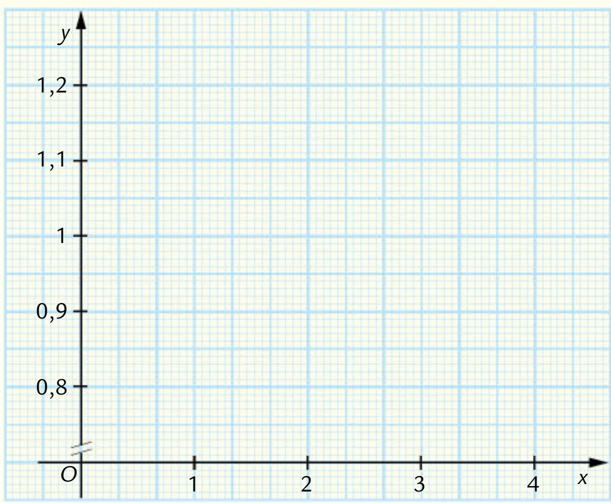
A $y = -0{,}125x + 1{,}2$

B $y = 0{,}125x + 2{,}2$

C $y = -0{,}125x + 0{,}2$

D $y = 0{,}125x + 1{,}2$

2 Rappresenta nel piano cartesiano predisposto in figura i dati riportati in tabella e la retta di regressione individuata al punto **1**.



3 In base al modello stabilito precedentemente, quale sarà il tasso alcolemico dopo 4 ore e mezza?

In Italia il valore limite legale stabilito per la guida è di 0,5 g/L.

4 Dopo quanto tempo la persona sottoposta alle misurazioni potrà rimettersi al volante, secondo il modello stabilito al punto **1**?

La legge impone ai neopatentati il divieto di assumere alcolici prima della guida: limite di tasso alcolemico pari a 0 g/L.

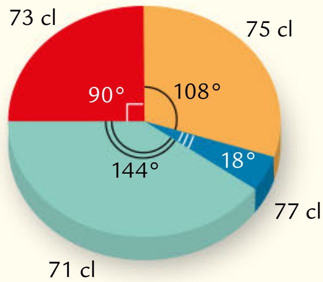
5 Supponi che l'automobilista sottoposto a misurazioni sia un neopatentato. In questo caso, in base al modello individuato al punto **1**, dopo quanto tempo potrà rimettersi alla guida?

Risposte p. 219

Compito di realtà 2

Controllo di qualità

Si vuole stabilire se le bottiglie di Barolo prodotte dall'azienda vinicola «Alticci» contengono effettivamente 75 cl di vino, come riportato nelle etichette: molti consumatori sospettano di no.
Il risultato delle analisi effettuate su un campione di 40 bottiglie è deducibile dal diagramma circolare seguente.



1 Quante bottiglie di vino del campione esaminato contengono meno di 75 cl di vino? Quante contengono più di 75 cl di vino?

2 Qual è la percentuale di bottiglie del campione che contengono una quantità di vino non inferiore a quanto dichiarato sull'etichetta?

3 Calcola la media e la deviazione standard della quantità di vino contenuto nelle 40 bottiglie esaminate.

4 Verifica che, a un livello di confidenza del 95%, non è possibile accettare quanto dichiarato dall'azienda vinicola «Alticci».

5 Cambierebbe la risposta al punto **4**, a un livello di confidenza del 99%?

Supponi ora che in tutte le bottiglie del campione esaminato sia stato imbottigliato 1 cl in più di Barolo.

6 Verifica che, anche in questo caso, a un livello di confidenza del 95%, bisognerebbe rifiutare quanto dichiarato dall'azienda.

Il responsabile del controllo qualità dell'azienda vuole stabilire quanti cl in più di Barolo avrebbero dovuto essere imbottigliati, come minimo, in tutte le bottiglie, per fare in modo che queste venissero dichiarate conformi, a un livello di confidenza del 95%.

7 Aiuta il responsabile: qual è la minima quantità di Barolo che avrebbe dovuto essere imbottigliata in più in tutte le bottiglie?

Risposte p. 220

Calcolo combinatorio e probabilità

TEMA
F

Abbiamo visto come la statistica sia capace di studiare, descrivere e interpretare i fenomeni e le situazioni più svariate. Ma questo all'uomo non è mai bastato e la domanda su che cosa si possa presentare nel futuro è sempre stata presente e sempre lo sarà. Il calcolo delle probabilità può essere impiegato per analizzare gli eventi che si prospettano a partire da una data situazione e per calcolare quanto è più facile che nel futuro si verifichi un evento rispetto a un altro. Ovviamente il calcolo delle probabilità non è una sorta di «premonizione» e non è in grado di dire con assoluta precisione che cosa accadrà a una determinata persona, in una certa località e in un dato istante. Applicati però opportunamente, i metodi di questa disciplina sono in grado di quantificare come certi fenomeni si evolveranno nel tempo, permettendo a chi deve prendere decisioni a riguardo di operare in modo adeguato.
Per esempio, gli enti incaricati del monitoraggio dei fiumi sono dotati di sistemi in grado di calcolare qual è la probabilità che, in conseguenza di determinate condizioni atmosferiche, si abbia una esondazione. Ciò permetterà a chi di dovere di disporre in modo sensato una eventuale evacuazione della popolazione.

Unità 3
Calcolo combinatorio

Unità 4
Calcolo delle probabilità

PREREQUISITI

- Le operazioni tra insiemi
- I connettivi logici

COMPETENZE

- Utilizzare modelli probabilistici per risolvere problemi ed effettuare scelte consapevoli

Unità 3

Tema F

Calcolo combinatorio

- Con GeoGebra
- Videolezioni
- Esercizi interattivi

1. Introduzione al calcolo combinatorio

In questa Unità presentiamo le tecniche di base del **calcolo combinatorio**, ovvero di quella parte della matematica che ha come oggetto il calcolo dei modi con i quali possono essere raggruppati o ordinati, secondo date regole, gli elementi di un insieme finito.

Che cosa si sta contando?

La maggior parte dei problemi di calcolo combinatorio sono riconducibili al seguente: quante «parole» di k caratteri si possono costruire con un alfabeto di n simboli distinti? La risposta dipende dalle due caratteristiche seguenti del problema in esame:

1. è importante l'*ordine* dei caratteri nelle «parole» che si vogliono contare?
2. sono consentite *ripetizioni* dei caratteri, ovvero, uno stesso simbolo può comparire nella «parola» più di una volta oppure no?

Ragioniamo su alcuni problemi per familiarizzare con questi aspetti.

Problema	Analisi	Modello
La mattinata di un corso di formazione prevede quattro ore di lezione: un'ora di diritto, una di inglese, una di gestione del personale e una di informatica. In quanti modi diversi si possono alternare i quattro relatori?	I diversi modi in cui possono succedersi i quattro relatori dipendono dall'*ordine* in cui avvengono le quattro lezioni. I relatori **non** possono *ripetersi*, perché è prevista una sola ora di lezione per ogni materia.	1. Parole ordinate, senza ripetizioni.
Una targa di automobile è costituita da due lettere iniziali, tre numeri e due lettere finali. Quante targhe diverse si possono costruire con questo sistema?	Nelle targhe delle automobili è *importante* l'*ordine* dei numeri e delle lettere. Inoltre sia le lettere sia i numeri possono *ripetersi* (per esempio, nella targa BR 808 RW la lettera R è ripetuta 2 volte, così come il numero 8).	2. Parole ordinate, con ripetizioni.
Quanti sono i possibili terni che si possono giocare al lotto?	Nel gioco del lotto, ai fini della vincita, hanno importanza soltanto i numeri estratti, non l'ordine di estrazione. Nel conteggio dei terni **non** ha quindi importanza l'*ordine* dei numeri. Inoltre tra i numeri che formano un terno **non** possono esserci *ripetizioni*, perché le estrazioni al lotto avvengono senza reimmissione.	3. Parole non ordinate, senza ripetizioni.
Quanti possibili tipi di confezioni diverse di 10 caramelle ai gusti di menta, fragola o limone si possono confezionare?	Ogni possibile confezione si può assimilare a un raggruppamento di 10 lettere, scelte tra M, F, L (M = menta, F = fragola, L = limone). Le lettere possono essere *ripetute* (perché in una confezione possono essere inserite più caramelle dello stesso tipo) e l'*ordine* **non** ha importanza (per stabilire se due confezioni sono o meno uguali confronteremo soltanto il numero di caramelle di ciascun tipo in esse contenute).	4. Parole non ordinate, con ripetizioni.

Una volta individuate le caratteristiche fondamentali del problema in esame in relazione all'ordine e alla possibilità di ripetizione, resta da effettuare il calcolo effettivo. Gli schemi di ragionamento da applicare per effettuare i conteggi in ciascuno dei quattro casi presentati in tabella verranno presentati nei prossimi paragrafi; alla base di tutti questi schemi di ragionamento c'è però *un solo principio fondamentale* che ora illustriamo.

Il principio fondamentale del calcolo combinatorio

Consideriamo il seguente problema.

◆ **PROBLEMA**

Si vuole preparare una colonna sonora per una presentazione multimediale: si è deciso di assemblare tre pezzi di musica classica in successione. Il primo brano sarà scelto fra *primavera*, *estate, autunno o inverno* di Vivaldi; il secondo fra le tre ultime sinfonie di Mozart (sinfonie in *mi bemolle*, in *sol minore* e in *do maggiore*) e l'ultimo tra l'*ottava* o la *nona* sinfonia di Beethoven. Quante colonne sonore si possono ottenere?

Possiamo ragionare secondo uno schema di *scelte successive*, visualizzato nel diagramma ad albero in **Fig. 1**.

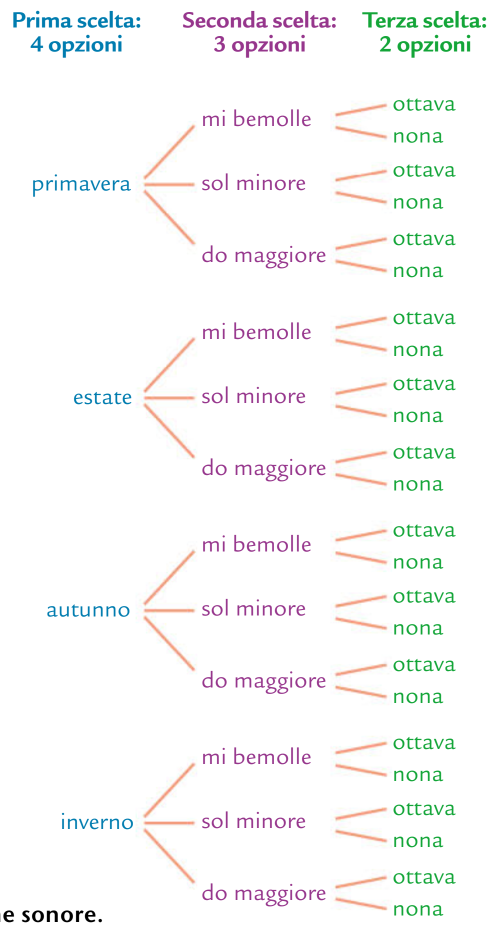


Figura 1 Diagramma ad albero delle colonne sonore.

Il diagramma ci permette di visualizzare come si calcola il numero complessivo di colonne sonore possibili:

- nella prima scelta abbiamo 4 opzioni;
- dopo avere effettuato anche la seconda scelta (3 opzioni) il numero complessivo di possibilità è diventato uguale a $12 = 4 \cdot 3$;
- dopo avere effettuato ulteriormente la terza scelta (2 opzioni), vediamo che il numero complessivo di colonne sonore disponibili diventa $24 = 12 \cdot 2$.

In sostanza, dopo ogni scelta, il numero complessivo di possibilità diviene uguale a quello ottenuto al passo precedente moltiplicato per il numero di opzioni dell'ultima scelta effettuata. In generale vale il seguente principio.

PRINCIPIO | Principio fondamentale del calcolo combinatorio

Se un oggetto è univocamente individuato da una sequenza di n scelte successive, tali che vi siano k_1 possibilità per la prima scelta, k_2 per la seconda, ..., k_n per la n-esima, il numero totale di oggetti che si possono formare con tali scelte è il prodotto:

$$k_1 \cdot k_2 \cdot ..... \cdot k_n$$

ESEMPIO Password

Una password è formata da 4 caratteri, che possono essere cifre (0,1, ..., 9) o lettere minuscole dell'alfabeto italiano (21 lettere). Inoltre cifre e lettere non possono essere ripetuti. Quante password di questo tipo si possono generare?

Per costruire una password dobbiamo effettuare quattro scelte successive:

1. per la scelta del primo carattere abbiamo 31 possibilità (le 10 cifre oppure una delle 21 lettere dell'alfabeto);
2. per la scelta del secondo carattere abbiamo 30 possibilità: infatti, poiché non sono ammesse ripetizioni di cifre e lettere, dobbiamo escludere il carattere scelto al passo precedente;
3. per la scelta del terzo carattere, per le stesse motivazioni del passo precedente, abbiamo 29 possibilità;
4. per la scelta del quarto carattere abbiamo, analogamente, 28 possibilità.

Per il principio fondamentale del calcolo combinatorio, possiamo costruire complessivamente:

$$31 \cdot 30 \cdot 29 \cdot 28 = 755\,160$$

password diverse.

Gli errori più frequenti nell'applicazione del principio fondamentale del calcolo combinatorio derivano dal non tenere conto di una ipotesi fondamentale: le sequenze di scelte devono individuare *univocamente* gli oggetti da contare. Ciò significa che a ogni sequenza di scelte deve corrispondere uno e un solo oggetto e, *viceversa*, a ogni oggetto deve corrispondere una e una sola sequenza di scelte. In modo equivalente, si può dire che per potere applicare il principio fondamentale del calcolo combinatorio deve esserci una *corrispondenza biunivoca* tra le sequenze di scelte e gli oggetti da contare.

ESEMPIO Applicazione impropria del principio di moltiplicazione

Quante squadre di due persone aventi come componenti Antonio, Massimo o Alberto si possono formare?

- Potremmo essere tentati di ragionare come segue. Una squadra può essere costruita tramite due scelte successive: prima scegliamo uno dei tre componenti (abbiamo 3 possibilità), quindi scegliamo il secondo componente tra le due persone rimaste (abbiamo 2 possibilità). Per il principio di moltiplicazione, concludiamo che si possono formare:

  $$3 \cdot 2 = 6 \text{ squadre}$$

- In realtà questo ragionamento non tiene conto del fatto che due squadre formate dagli stessi individui si reputano generalmente uguali indipendentemente dall'ordine degli individui; pertanto è corretto dire che si possono formare soltanto 3 squadre che hanno come componenti Antonio, Massimo o Alberto:

  $$\{\text{Antonio, Massimo}\}, \{\text{Antonio, Alberto}\}, \{\text{Massimo, Alberto}\}.$$

Come mai, applicando il principio di moltiplicazione, abbiamo ottenuto una *sovrastima* del numero di squadre? Il motivo risiede nel fatto che il risultato trovato tiene conto dell'*ordine* delle scelte e quindi ciascuna squadra è stata contata due volte: per esempio, è stato considerato sia il caso della squadra formata, n*ell'ordine*, da Antonio e Massimo, sia il caso della squadra formata, *nell'ordine*, da Massimo e Antonio. Per correggere la sovrastima ottenuta dobbiamo dividere 6 per 2, ottenendo così la risposta corretta: 3.

Riflettendo sull'utilizzo che abbiamo fatto del principio di moltiplicazione, ci rendiamo conto che l'errore è consistito nell'averlo applicato troppo superficialmente, senza controllare *se* una sequenza di scelte definisce *univocamente* una squadra.

In effetti è vero che a ogni sequenza di due scelte corrisponde una e una sola squadra, ma **non** è vero che, *viceversa*, a ogni squadra corrisponde una e una sola sequenza di scelte.
Per esempio, la squadra formata da Antonio e Massimo può ottenersi scegliendo prima Antonio e poi Massimo, *oppure* scegliendo prima Massimo e poi Antonio.
Non erano dunque soddisfatte tutte le ipotesi necessarie per potere applicare il principio fondamentale del calcolo combinatorio.

Esercizi p. 116

2. Disposizioni e permutazioni

In questo paragrafo, vogliamo costruire dei modelli generali di riferimento, per risolvere problemi che si possono assimilare al conteggio di «parole» *ordinate*.

Disposizioni semplici e permutazioni

Consideriamo il seguente problema.

PROBLEMA
A una gara cui partecipano 10 concorrenti, quante sono le possibili classifiche dei primi 3?

Osserviamo che:
- per il primo posto possiamo scegliere tra 10 possibilità;
- per il secondo posto possiamo scegliere tra 9 (10 − 1) possibilità;
- per il terzo posto possiamo scegliere tra 8 (10 − 2) possibilità.

Il numero complessivo di classifiche possibili, per il principio fondamentale del calcolo combinatorio, è quindi:

$$10 \cdot 9 \cdot 8$$

Ogni classifica si può assimilare a una *sequenza ordinata* di 3 atleti scelti tra 10, ammettendo che ogni atleta possa essere scelto una e una sola volta.
Più in generale si pone il problema seguente: «dati n oggetti, quante sono le possibili sequenze ordinate di k oggetti, scelti tra gli n assegnati con il vincolo di non ripetere gli oggetti?».
Alle sequenze di questo tipo si dà un nome particolare.

DEFINIZIONE | Disposizioni semplici

Dati n oggetti distinti, si chiama **disposizione semplice** (o semplicemente **disposizione**) degli n oggetti in k posti, con $k \leq n$, ogni sequenza *ordinata* di k oggetti scelti tra quelli assegnati con il vincolo di non ripetere gli oggetti.

MODI DI DIRE
Le disposizioni di n oggetti in k posti si chiamano anche disposizioni di n oggetti **di classe** k.

Ragionando come nel problema poc'anzi esaminato si giunge a formulare il seguente teorema.

TEOREMA 1 | Numero di disposizioni semplici

Il numero complessivo di **disposizioni di n oggetti in k posti**, indicato con il simbolo $D_{n,k}$, è dato dalla formula:

$$D_{n,k} = \underbrace{n(n-1)(n-2) \cdot \ldots \cdot (n-k+1)}_{\text{prodotto di } k \text{ fattori decrescenti}} \qquad [1]$$

CON LA CALCOLATRICE

Le calcolatrici scientifiche consentono il calcolo delle disposizioni semplici con l'apposito tasto, spesso denotato con nPr, che ha lo stesso significato di $D_{n,r}$.

Per esempio:

$$D_{5,3} = \underbrace{5 \cdot 4 \cdot 3}_{\text{3 fattori decrescenti}} = 60$$

$$D_{7,4} = \underbrace{7 \cdot 6 \cdot 5 \cdot 4}_{\text{4 fattori decrescenti}} = 840$$

ESEMPI Problemi che hanno come modello le disposizioni semplici

a. A una corsa di cavalli ci sono 15 cavalli in gara. Quante classifiche possibili dei primi tre ci possono essere?

Ogni classifica dei primi tre cavalli si può assimilare a una disposizione di 15 oggetti in 3 posti, quindi il numero totale di tutte le possibili classifiche è:

$$D_{15,3} = 15 \cdot 14 \cdot 13 = 2730$$

b. Un'urna contiene 10 palline numerate 1, 2, 3, ..., 10. Se ne estraggono successivamente quattro, senza rimettere la pallina estratta nell'urna prima di estrarre la successiva. Quante diverse estrazioni sono possibili, tenendo conto dell'ordine delle estrazioni?

Ogni estrazione delle quattro palline si può assimilare a una disposizione di 10 oggetti in 4 posti, quindi il numero complessivo di estrazioni possibili è:

$$D_{10,4} = 10 \cdot 9 \cdot 8 \cdot 7 = 5040$$

RICORDA

Le disposizioni (semplici) sono i modelli adatti ad affrontare problemi combinatori in cui:
– *è importante l'ordine*
– *gli oggetti non possono ripetersi.*

Nel caso particolare in cui $n = k$, una disposizione di n oggetti in k posti equivale a un *ordinamento* degli n oggetti e viene chiamata **permutazione**.

DEFINIZIONE | Permutazione

Si chiama **permutazione** di n oggetti distinti ogni **ordinamento** degli n oggetti dati.

Dalla formula [1] segue subito che il numero di permutazioni di n oggetti, indicato solitamente con il simbolo P_n, è dato da:

$$P_n = D_{n,n} = n(n-1)(n-2) \cdot ... \cdot 3 \cdot 2 \cdot 1$$

OSSERVA

Il simbolo $n!$ è definito per ogni $n \in \mathbb{N}$ e non è mai nullo.

Il prodotto $n(n-1)(n-2) \cdot ... \cdot 3 \cdot 2 \cdot 1$ si indica a sua volta con il simbolo $n!$, che si legge «**fattoriale** di n» o «n **fattoriale**». Il simbolo $n!$ resta così definito per ogni n intero positivo, mentre per $n = 0$ si pone, per definizione, $0! = 1$.

TEOREMA 2 | Numero di permutazioni

Il numero complessivo di **permutazioni di n oggetti** distinti è dato dalla formula:

$$P_n = n!$$

Per esempio:

$$3! = 3 \cdot 2 \cdot 1 \quad \text{e} \quad 7! = 7 \cdot 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 5040$$

ESEMPI Problemi che hanno come modello le permutazioni

a. In quanti modi possibili 6 persone possono essere disposte in fila indiana?

Ci sono tanti modi di disporre le persone in fila indiana quante le permutazioni di 6 oggetti; quindi il numero complessivo di modi di disporre le persone è uguale a:

$$6! = 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 720$$

b. Quanti sono i possibili anagrammi (anche privi di significato) della parola «cielo»?

Sono tanti quante le permutazioni delle cinque lettere *c*, *i*, *e*, *l*, *o* quindi sono:

$$5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$$

c. Al corso di balli sudamericani sono iscritti 4 ragazzi e 4 ragazze. Calcola tutti i possibili accoppiamenti per il prossimo samba.

Il primo ballerino può avere come partner una qualsiasi delle 4 ballerine, il secondo ballerino una qualsiasi delle rimanenti 3, e così via. Perciò il numero degli accoppiamenti possibili è dato da:

$$4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$$

Disposizioni con ripetizione

Continuiamo la nostra analisi, ragionando sul seguente problema.

◆ **PROBLEMA**
In quanti modi diversi è possibile riempire una colonna del totocalcio?

Ogni colonna è costituita da 14 caselle, ciascuna delle quali deve essere riempita con uno dei tre simboli 1, 2 o X.
Dunque:
- per la prima casella abbiamo 3 possibilità di scelta (1, 2, X);
- per la seconda casella possiamo ancora scegliere tra 3 possibilità;
- e così via, fino alla quattordicesima casella.

Complessivamente una colonna può essere riempita in:

$$\underbrace{3 \cdot 3 \cdot \ldots \cdot 3 \cdot 3}_{\text{14 volte}} = 3^{14}$$

modi diversi.

La differenza di questo problema rispetto a quelli considerati nel precedente sottoparagrafo, è che qui gli oggetti (i simboli 1, 2, X) *possono essere scelti più di una volta, ossia possono essere ripetuti*.
Più in generale si pone il seguente problema: «dati n oggetti, quante sono le possibili *sequenze ordinate* di k oggetti, scelti tra quelli assegnati ammettendo che *sia possibile ripetere* gli oggetti?».
Alle sequenze di questo tipo si dà un nome particolare.

DEFINIZIONE | Disposizioni con ripetizione

Dati n oggetti distinti, si chiama **disposizione con ripetizione** degli n oggetti in k posti, ogni sequenza *ordinata* di k oggetti, scelti tra quelli assegnati ammettendo che sia possibile ripetere gli oggetti.

OSSERVA
Poiché è ammesso ripetere gli oggetti, può essere $k \geq n$.

Ragionando come nel problema iniziale si giunge al seguente teorema.

TEOREMA 3 | Numero di disposizioni con ripetizione

Il numero complessivo di **disposizioni con ripetizione** di n oggetti in k posti, indicato con il simbolo $D^*_{n,k}$ è dato dalla formula:

$$D^*_{n,k} = \underbrace{n \cdot n \cdot \ldots \cdot n \cdot n}_{\text{prodotto di } k \text{ fattori uguali a } n} = n^k$$

RICORDA
Le disposizioni con ripetizione sono i modelli adatti ad affrontare problemi combinatori in cui:
– *è importante l'ordine*
– *gli oggetti possono ripetersi.*

ESEMPI Problemi che hanno come modello le disposizioni con ripetizione

a. Un numero di telefono di cellulare di dieci cifre inizia con 347. Quanti numeri di telefono di questo tipo ci possono essere?

Le sette cifre rimanenti, dopo 347, possono essere ciascuna un numero qualsiasi tra 0 e 9 (per ogni cifra ci sono dunque dieci possibilità). Esistono perciò tanti numeri di cellulare quante le disposizioni con ripetizione di 10 oggetti in 7 posti, cioè 10^7.

b. Un test è costituito da 10 domande a risposta multipla: A, B, C, oppure D. In quanti modi può essere compilato il test?

Ciascuna delle 10 domande prevede 4 possibili risposte, quindi il numero cercato è $4^{10} = 1\,048\,576$.

Permutazioni con ripetizioni

Consideriamo il seguente problema.

◆ **PROBLEMA**
Quanti sono gli anagrammi della parola «mamma»?

ATTENZIONE!
Salvo avviso contrario, quando parleremo di «anagramma» intenderemo riferirci anche agli anagrammi privi di significato.

Il problema si presenta differente da quello che abbiamo affrontato nel precedente sottoparagrafo, quando abbiamo calcolato gli anagrammi della parola «cielo» perché in quest'ultima parola le lettere sono tutte *distinte*, mentre nella parola «mamma» la lettera «m» è ripetuta 3 volte e la lettera «a» è ripetuta 2 volte.
Più in generale si pone il seguente problema: «quante sono le possibili permutazioni di n oggetti, nell'ipotesi che essi **non** siano tutti *distinti*?».

DEFINIZIONE | Permutazioni con ripetizione
Si chiama **permutazione con ripetizione** ogni permutazione di n oggetti **non** tutti distinti tra loro.

Come possiamo calcolare il numero delle permutazioni con ripetizione? Ragioniamo sul problema da cui siamo partiti di trovare il numero totale, diciamo x, degli anagrammi della parola «mamma».
Cominciamo con il distinguere con un indice le tre lettere «m»:

$$\mathrm{m_1am_2m_3a} \qquad [2]$$

e osserviamo che, per ognuna delle x permutazioni della parola mamma, possiamo costruire 3! permutazioni della parola $\mathrm{m_1am_2m_3a}$, permutando tra loro le 3 lettere $\mathrm{m_1}$, $\mathrm{m_2}$, $\mathrm{m_3}$. Dunque le possibili permutazioni della parola [2] sono $x \cdot 3!$.
Distinguiamo ora con un indice anche le 2 lettere «a»:

$$\mathrm{m_1a_1m_2m_3a_2} \qquad [3]$$

Ragionando come nel caso precedente, concludiamo che le permutazioni della parola [3] saranno $x \cdot 3! \cdot 2!$. D'altra parte le permutazioni della [3] non sono altro che le permutazioni di 5 oggetti distinti, quindi il loro numero totale deve essere 5!. Ne viene l'equazione seguente, da cui è possibile ricavare il numero incognito x:

$$x \cdot 3! \cdot 2! = 5! \ \Rightarrow\ x = \frac{5!}{3! \cdot 2!}$$

Il ragionamento poc'anzi esposto può essere generalizzato e conduce al seguente teorema.

TEOREMA 4 | Numero di permutazioni con ripetizione
Dati n oggetti, di cui a_1 uguali tra loro, a_2 uguali tra loro (e distinti dai precedenti); a_k uguali tra loro (e distinti dai precedenti) con $a_1 + a_2 + \dots + a_k = n$, le permutazioni distinte di questi n oggetti sono:

$$\frac{n!}{a_1! \cdot a_2! \cdot \ldots \cdot a_k!}$$

ESEMPI Problemi che hanno come modello permutazioni con ripetizione

a. Quanti sono gli anagrammi della parola «matematica»?

La parola matematica è formata da 10 lettere, di cui la *m* si ripete 2 volte, la *a* si ripete 3 volte, la *t* si ripete 2 volte, mentre la *e*, la *c* e la *i* compaiono 1 sola volta. Quindi il numero dei possibili anagrammi è

$$\frac{10!}{2! \cdot 3! \cdot 2! \cdot 1! \cdot 1! \cdot 1!} = 151\,200$$

b. In quanti modi diversi una colonna della schedina può essere riempita con 4 segni 1, 6 segni X e 4 segni 2?

Ogni modo di riempire la colonna si può assimilare a una permutazione di 14 elementi di cui 4 coincidenti con 1, 6 coincidenti con X e 4 coincidenti con 2; pertanto il numero complessivo di modi è:

$$\frac{14!}{4! \cdot 6! \cdot 4!} = 2\,101\,210$$

Esercizi p. 118

3. Combinazioni

In questo paragrafo, vogliamo costruire dei modelli generali di riferimento, per risolvere problemi che si possono assimilare al conteggio di «parole» **non** *ordinate*.

Combinazioni

Consideriamo il seguente problema.

◆ PROBLEMA
Gioco al lotto i cinque numeri: 1, 2, 3, 4, 5. In quanti modi posso fare terno?

Rispondere alla domanda posta da questo problema equivale a determinare quanti terni è possibile costruire utilizzando i cinque numeri giocati: 1, 2, 3, 4, 5. I terni devono considerarsi **non** ordinati e in essi non possono esserci numeri ripetuti: per esempio, alcuni possibili terni sono $\{1, 2, 3\}$, $\{3, 4, 5\}$, $\{2, 3, 5\}$, ... Ai raggruppamenti di oggetti non ordinati e senza ripetizioni si dà un nome particolare.

DEFINIZIONE | Combinazione

Dati n oggetti distinti, si chiama **combinazione** degli n oggetti di classe k, ogni raggruppamento **non** *ordinato* di k oggetti scelti tra quelli assegnati, con il vincolo di **non** ripetere gli oggetti.

Confronto tra disposizioni e combinazioni

disposizioni → rappresentano sequenze **ordinate**

combinazioni → rappresentano reggruppamenti **non** ordinati

Come è noto, dato un *insieme* di oggetti (dove il termine «insieme» è da intendersi nell'accezione matematica), gli oggetti **non** possono essere ripetuti e **non** viene preso in considerazione l'ordine degli oggetti: una combinazione di n oggetti di classe k si può perciò identificare con un *sottoinsieme* di k elementi dell'insieme formato dagli n oggetti considerati.

ATTENZIONE!
Il coefficiente binomiale $\binom{n}{k}$ è definito a condizione che $n \geq 1$ e $0 \leq k \leq n$.

DEFINIZIONE | Coefficiente binomiale

Il numero complessivo di combinazioni di n oggetti di classe k viene indicato con il simbolo $C_{n,k}$ o con il simbolo $\binom{n}{k}$ che prende il nome di **coefficiente binomiale** e si legge «n su k».

Quali sono i valori di $\binom{n}{k}$? Proviamo a ragionare inizialmente sul caso particolare proposto all'inizio, cioè sul problema di determinare tutti i sottoinsiemi di 3 elementi dell'insieme $S = \{1, 2, 3, 4, 5\}$; in altri termini ci poniamo l'obiettivo di determinare $C_{5,3}$, ovvero $\binom{5}{3}$.

Non possiamo applicare lo schema delle scelte successive (ossia il principio fondamentale del calcolo combinatorio) perché un insieme è determinato soltanto dai suoi elementi, a *prescindere dall'ordine.* Possiamo però osservare che, costruendo le permutazioni di *tutti* i sottoinsiemi di 3 elementi di S (ci sono 3! permutazioni per ogni sottoinsieme), si ottengono *tutte* le disposizioni dei cinque elementi di S in 3 posti (che sappiamo essere uguali a $D_{5,3}$). Pertanto:

RIFLETTI
Il principio fondamentale del calcolo combinatorio non può essere utilizzato perché tre scelte successive di un elemento da un insieme di 5 elementi non determinano in modo *univoco* un sottoinsieme di 3 elementi dell'insieme originario. È vero infatti che le tre scelte individuano un unico sottoinsieme di 3 elementi, ma **non** è vero che, *viceversa*, ogni sottoinsieme è individuato soltanto dalle tre scelte effettuate (scelte che differiscono solo per l'ordine determinano infatti lo stesso sottoinsieme).

$$\underbrace{C_{5,3}}_{\text{il numero complessivo dei sottoinsiemi di 3 elementi dell'insieme } S} \cdot \underbrace{3!}_{\text{il numero di permutazioni di ciascun sottoinsieme}} \underbrace{=}_{\text{è uguale al}} \underbrace{D_{5,3}}_{\text{numero complessivo di terne ordinate che si possono costruire con gli elementi dell'insieme } S}$$

$$\Rightarrow C_{5,3} = \frac{D_{5,3}}{3!} = \frac{5 \cdot 4 \cdot 3}{3!} = 10$$

Questo ragionamento si può generalizzare e porta a concludere che:

$$\boxed{C_{n,k} = \frac{D_{n,k}}{k!}}$$

Possiamo enunciare pertanto il teorema seguente.

TEOREMA 5 | Numero di combinazioni di n oggetti di classe k

Il numero complessivo di **combinazioni** di n oggetti di classe k, indicato con il simbolo $C_{n,k}$ è dato dalla formula:

$$C_{n,k} = \binom{n}{k} = \frac{n(n-1)(n-2) \cdot \ldots\ldots \cdot (n-k+1)}{k!} \qquad [4]$$

CON LA CALCOLATRICE
Puoi calcolare rapidamente il valore di un coefficiente binomiale tramite la calcolatrice di solito premendo il tasto nCr che ha il significato di $C_{n,r}$.

Per esempio:

$$\binom{7}{3} = \frac{7 \cdot 6 \cdot 5}{3!} = \frac{7 \cdot 6 \cdot 5}{3 \cdot 2 \cdot 1} = 35$$

È importante mettere in rilievo alcune *proprietà* dei coefficienti binomiali.

1. Il simbolo $\binom{n}{0}$ rappresenta il numero dei sottoinsiemi di 0 elementi di un insieme di n elementi; poiché l'unico sottoinsieme di 0 elementi di qualunque insieme è l'insieme vuoto, risulta:

$$\boxed{\binom{n}{0} = 1} \qquad \text{per ogni } n \in \mathbf{N} - \{0\}$$

2. Analogamente, il simbolo $\binom{n}{1}$ rappresenta il numero dei sottoinsiemi di 1 elemento di un insieme di n elementi; poiché di sottoinsiemi siffatti ne esistono esattamente n, risulta:

$$\binom{n}{1} = n \qquad \text{per ogni } n \in \mathbf{N} - \{0\}$$

3. In base alla definizione di fattoriale, è immediato verificare che risulta:

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \qquad [5]$$

Infatti l'espressione al secondo membro della [5] si semplifica e si ottiene la [4]. La [5] è detta *formula dei 3 fattoriali* (di De Moivre).

4. Dalla [5] si deduce immediatamente che:

$$\binom{n}{k} = \binom{n}{n-k} \qquad [6]$$

Questa proprietà è utile per velocizzare il calcolo dei coefficienti binomiali quando $k > \frac{n}{2}$. Per esempio, per calcolare $\binom{10}{8}$ conviene procedere così:

$$\binom{10}{8} = \binom{10}{2} = \frac{10 \cdot 9}{2!} = 45$$

OSSERVA
Per $k = 0$ la [5] fornisce: $\binom{n}{0} = \frac{n!}{0!\,n!} = \frac{1}{0!}$, che è uguale a 1 perché abbiamo posto per definizione $0! = 1$. Si comprende quindi il motivo di questa definizione: è necessaria per far sì che la [5] valga anche per $k = 0$.

MODI DI DIRE
La [6] è anche detta *legge delle classi complementari.*

ESEMPI Problemi che hanno come modello combinazioni

a. Una grossa azienda deve inviare 2 dei suoi 8 ispettori a controllare una filiale lontana. In quanti modi possibili il capo dell'ufficio può determinare la delegazione di 2 ispettori?

Determinare una delegazione equivale a determinare un sottoinsieme di 2 elementi dell'insieme formato dagli 8 ispettori. Quindi il capo dell'ufficio ha $\binom{8}{2}$ possibilità di scelta. Le delegazioni possibili sono quindi:

$$\binom{8}{2} = \frac{8 \cdot 7}{2!} = \frac{56}{2} = 28$$

b. Quanti sono i possibili terni che si possono giocare al gioco del lotto?

Un terno al gioco del lotto equivale a un sottoinsieme di 3 numeri dell'insieme $\{1, 2, 3, ..., 90\}$.
I possibili terni che si possono giocare sono dunque complessivamente:

$$\binom{90}{3} = \frac{90 \cdot 89 \cdot 88}{3 \cdot 2 \cdot 1} = 30 \cdot 89 \cdot 44 = 117\,480$$

RICORDA
Le combinazioni (semplici) sono i modelli adatti a descrivere problemi in cui:
- *non è importante l'ordine*
- *gli oggetti non possono ripetersi.*

Combinazioni con ripetizione

Riconsideriamo uno dei problemi proposti nel **Paragrafo 1**:

◆ PROBLEMA
Quanti possibili tipi di confezioni diverse di 10 caramelle ai gusti di menta, fragola o limone si possono confezionare (ammettendo che le caramelle dello stesso gusto siano tutte dello stesso tipo e quindi indistinguibili l'una dall'altra)?

Ogni possibile confezione si può assimilare a un raggruppamento di 10 lettere, scelte tra M, F, L (M = menta, F = fragola, L = limone). Per esempio, una confezione contenente 5 caramelle alla menta, 2 al limone e 3 alla fragola potrà essere identificata dal raggruppamento:

MMMMMLLFFF **[7]**

Osserviamo che in ogni raggruppamento di lettere di questo tipo:

1. le lettere possono essere *ripetute* (perché in una confezione possono essere inserite più caramelle dello stesso tipo) oppure qualche lettera può non comparire (per esempio è possibile che la confezione non contenga caramelle alla menta);

2. l'*ordine* delle lettere **non** ha importanza; per esempio il raggruppamento:

   MMMLLFFFMM

   rappresenta ancora una confezione contenente 5 caramelle alla menta, 2 al limone e 3 alla fragola, confezione da considerarsi *uguale* a quella rappresentata dal raggruppamento **[7]**.

In sostanza si possono determinare tanti tipi di confezioni quanti i raggruppamenti di 10 lettere, ciascuna scelta nell'insieme $\{M, F, L\}$, considerando *uguali* due raggruppamenti che differiscono *solo per l'ordine* e ammettendo che sia possibile *ripetere* le lettere. Un problema di questo tipo si inquadra nell'ambito del calcolo di *combinazioni* in cui gli elementi possono essere *ripetuti*.

DEFINIZIONE | Combinazioni con ripetizione

Dati n oggetti distinti, si chiama **combinazione con ripetizione** degli n oggetti **di classe *k***, ogni raggruppamento **non** *ordinato* di k oggetti, scelti tra quelli assegnati, ammettendo la possibilità di ripetere gli oggetti.

È importante osservare che, essendo ammesso ripetere gli oggetti, può essere $k \geq n$.
Per stabilire *quante sono* le combinazioni con ripetizione di n oggetti di classe k, ritorniamo a ragionare sul problema da cui siamo partiti, che chiedeva in pratica di determinare le combinazioni con ripetizione di $n = 3$ oggetti (le lettere M, F, L) di classe $k = 10$. Osserviamo che una confezione resta univocamente individuata una volta che vengano stabiliti i numeri x_1, x_2 e x_3, rispettivamente di caramelle alla menta, alla fragola e al limone che deve contenere. Poiché ogni confezione deve contenere 10 caramelle i numeri x_1, x_2 e x_3 devono soddisfare l'equazione:

$$x_1 + x_2 + x_3 = 10 \qquad [8]$$

Dunque il problema di determinare il numero di tutti i possibili tipi di confezioni equivale a quello di stabilire il numero di terne (x_1, x_2, x_3), con $x_1, x_2, x_3 \in \mathbf{N}$, che soddisfano l'equazione **[8]**.
Per risolvere quest'ultimo problema ragioniamo come segue.
Sia (x_1, x_2, x_3) una soluzione dell'equazione **[8]**, per esempio (3, 2, 5).
Rappresentiamo questa soluzione con una sequenza di asterischi e barre come segue:

$$\underbrace{* * *}_{x_1} \,|\, \underbrace{* *}_{x_2} \,|\, \underbrace{* * * * *}_{x_3}$$

Gli asterischi * servono a rappresentare i valori di x_1, x_2, x_3 e le barre | a separarli

Viceversa, ogni sequenza costituita da 10 asterischi e 2 barre verticali individua una terna soluzione; per esempio la sequenza:

$$| * * * * | * * * * * *$$

individua la soluzione (0, 4, 6).
Dunque l'insieme delle terne cercate è in corrispondenza biunivoca con le permutazioni di 12 oggetti, di cui 10 uguali tra loro (gli asterischi) e 2 uguali tra loro e distinti dai precedenti (le barre).
Le terne cercate sono dunque in totale:

$$\frac{12!}{10! \cdot 2!} = \binom{12}{10} = 66$$

Il ragionamento precedente si può generalizzare. Il numero di combinazioni con ripetizione di n oggetti di classe k equivale al numero di n-uple $(x_1, x_2, ..., x_n)$, con $x_1, x_2, ..., x_n \in \mathbf{N}$, che soddisfano l'equazione:

$$x_1 + ... + x_n = k$$

Le n-uple soluzioni di questa equazione sono in corrispondenza biunivoca con le permutazioni di $k + (n - 1)$ oggetti, di cui k uguali tra loro (gli asterischi) e $(n - 1)$ uguali tra loro e distinti dai precedenti (le barre). Il numero totale di queste permutazioni è uguale a:

$$\frac{(n + k - 1)!}{k!(n - 1)!} = \binom{n + k - 1}{k}$$

Possiamo quindi enunciare il seguente teorema.

TEOREMA 6 | Combinazioni con ripetizione di *n* oggetti di classe *k*

Il numero di combinazioni con ripetizione di n oggetti di classe k, che indichiamo con $C^*_{n,k}$, equivale al numero di n-uple di interi non negativi $(x_1, ..., x_n)$ soluzioni dell'equazione:

$$x_1 + ... + x_n = k$$

ed è assegnato dalla formula:

$$C^*_{n,k} = \binom{n + k - 1}{k}$$

RIFLETTI

L'idea alla base della risoluzione del problema delle confezioni di caramelle è la costruzione di opportune *corrispondenze biunivoche* che consentono di tradurre il problema iniziale nel conteggio di oggetti che già sappiamo calcolare; schematicamente abbiamo osservato che:

Il numero di confezioni di caramelle

è in corrispondenza biunivoca con

Le soluzioni dell'equazione $x_1 + x_2 + x_3 = 10$

sono in corrispondenza biunivoca con

Le sequenze ordinate costituite da 10 asterischi e 2 barre

sono in corrispondenza biunivoca con

Le permutazioni di 12 oggetti, di cui 10 uguali tra di loro e 2 uguali tra loro e diversi dai precedenti

ESEMPIO Problemi che hanno come modello combinazioni con ripetizione

Quanti tipi di confezioni diverse di 10 caramelle ai gusti di menta, fragola o limone si possono confezionare, contenenti almeno tre caramelle di gusto diverso?

Diversamente dal caso del problema precedente, ora si vuole assicurare almeno un assaggio di tutti e tre i gusti. Occorre dunque inserire inizialmente nel sacchetto le tre caramelle richieste (una per ciascuno dei tre gusti); a questo punto potremo completare il sacchetto in tanti modi quante sono le possibili confezioni contenenti $10 - 3 = 7$ caramelle ai gusti menta, fragola o limone.
Applicando la formula otteniamo:

$$C^*_{3,7} = \binom{9}{7} = \binom{9}{2} = \frac{9 \cdot 8}{2!} = 36$$

In conclusione le possibili confezioni contenenti 10 caramelle, di cui una almeno per ciascuno dei tre gusti, sono in tutto 36.

RICORDA

Le combinazioni con ripetizione sono i modelli adatti a descrivere problemi in cui:
- ***non è importante l'ordine***
- ***gli oggetti possono ripetersi.***

COLLEGHIAMO I CONCETTI **Formule e problemi di calcolo combinatorio**

▶ Abbiamo via via trovato le formule di conteggio relative a tutti e quattro i modelli di riferimento introdotti nel **Paragrafo 1**.

Numero di «parole» di k caratteri costruite da un alfabeto di n simboli distinti	senza ripetizioni	con ripetizioni
ordinate	$D_{n,k} = n(n-1)(n-2) \cdot \ldots \cdot (n-k+1)$ Nel caso particolare in cui $n = k$: $D_{n,n} = P_n = n!$	$D^*_{n,k} = n^k$
non ordinate	$C_{n,k} = \binom{n}{k}$	$C^*_{n,k} = \binom{n+k-1}{k}$

▶ Il numero di parole ordinate di n caratteri, costruite a partire da un alfabeto di n simboli **non** tutti distinti, di cui a_1 uguali tra loro, a_2 uguali tra loro (e distinti dai precedenti), ..., a_k uguali tra loro (e distinti dai precedenti) con $a_1 + a_2 + \ldots + a_k = n$, è assegnato dalla formula delle permutazioni con ripetizione:

$$\frac{n!}{a_1! \cdot a_2! \cdot \ldots \cdot a_k!}$$

▶ Alcuni problemi di calcolo combinatorio richiedono l'esecuzione di vari conteggi «intermedi» per giungere alla soluzione. In questi casi occorre prestare attenzione a stabilire se i vari conteggi «intermedi» alla fine vanno *moltiplicati* o *sommati*. Rifletti sui seguenti due esempi.

Esempio	Metodo
Supponiamo di estrarre 5 carte (una mano) da un mazzo di 32 carte: quante mani contengono esattamente due donne ed esattamente un fante? Ricorda che un mazzo di 32 carte contiene 8 carte di ogni seme: 7, 8, 9, 10, fante, donna, re e asso.	Una mano come quella richiesta è univocamente determinata effettuando le seguenti scelte: – 2 donne (tra le quattro disponibili); – 1 fante (tra i quattro disponibili); – 2 carte (tra quelli che restano escludendo le donne e i fanti). Le tre scelte possono avvenire rispettivamente in: $\binom{4}{2}, \binom{4}{1}, \binom{24}{2}$ modi. Per il principio fondamentale del calcolo combinatorio, il numero totale di mani che soddisfano le condizioni richieste è: $\underbrace{\binom{4}{2} \cdot \binom{4}{1} \cdot \binom{24}{2}}_{\text{i conteggi iniziali vanno moltiplicati}} = 6 \cdot 4 \cdot 276 = 6624$
Supponiamo di estrarre 5 carte (una mano) da un mazzo di 32 carte: quante mani contengono almeno tre re?	Una mano contiene almeno tre re se e solo se contiene esattamente 3 re o esattamente 4 re. Le mani che contengono esattamente 3 re (e quindi altre 2 carte che non sono re) sono: $\binom{4}{3} \cdot \binom{28}{2}$ Le mani che contengono esattamente 4 re (e quindi 1 altra carta che non sia un re) sono: $\binom{4}{4} \cdot \binom{28}{1}$ L'insieme A costituito dalle mani contenenti esattamente 3 re e l'insieme B costituito dalle mani contenenti esattamente 4 re sono *disgiunti*, quindi il numero totale di mani che contengono almeno tre re è la *somma* degli elementi di A e di quelli di B: $\underbrace{\binom{4}{3} \cdot \binom{28}{2} + \binom{4}{4} \cdot \binom{28}{1}}_{\text{i conteggi iniziali vanno sommati}} = 4 \cdot 378 + 1 \cdot 28 = 1540$

Esercizi p. 124

4. Il teorema del binomio di Newton

Lo sviluppo della potenza del binomio con il triangolo di Tartaglia

Hai già visto nei tuoi studi precedenti la regola per calcolare lo sviluppo della potenza di un binomio basata sul triangolo di Tartaglia e il modo in cui si costruisce tale triangolo. Ricordiamo brevemente i punti fondamentali.

1. Le prime righe del triangolo di Tartaglia sono riportate qui sotto:

1	← riga 0
1 1	← riga 1
1 2 1	← riga 2
1 3 3 1	← riga 3
1 4 6 4 1	← riga 4
1 5 10 10 5 1	← riga 5
1 6 15 20 15 6 1	← riga 6

2. Il procedimento per costruire il triangolo di Tartaglia è molto semplice.
Se indichiamo con x e y due numeri successivi posti su di una stessa riga, l'elemento posto fra di essi, nella riga immediatamente al di sotto è la loro somma:

$$\begin{array}{cc} x & y \\ \searrow & \swarrow \\ \multicolumn{2}{c}{x+y} \end{array}$$

Per esempio, la riga 5 può essere dedotta dalla riga 4 come segue:

	1	4	6	4	1	← riga 4
1	1 + 4	4 + 6	6 + 4	4 + 1	1	
1	5	10	10	5	1	← riga 5

Continuando con questo procedimento si possono costruire tante righe quante si vogliono del triangolo di Tartaglia. Per esempio, puoi ricavare la settima riga e verificare che coincide con:

1 7 21 35 35 21 7 1

3. Lo sviluppo della potenza $(a+b)^n$, può essere eseguito secondo la seguente regola, che chiede la costruzione della n-esima riga del triangolo di Tartaglia:

REGOLA | Potenza *n*-esima di un binomio

Lo sviluppo di $(a+b)^n$ è un polinomio **omogeneo** di grado n, **ordinato** secondo le potenze **decrescenti** di $\boldsymbol{a}$ (a partire da quella di grado n) e **crescenti** di $\boldsymbol{b}$ (a partire da quella di grado 0), i cui **coefficienti** sono quelli della n-esima riga del **triangolo di Tartaglia**.

ESEMPIO Sviluppo di un binomio secondo la regola del triangolo di Tartaglia

Calcoliamo $(a+b)^5$.

Lo sviluppo della potenza sarà un polinomio *omogeneo* di quinto grado, ordinato secondo le potenze decrescenti di a (iniziando da quella di grado 5) e crescenti di b (a partire da quella di grado 0); si tratterà quindi di un polinomio del tipo:

$$\ldots\ldots a^5b^0 + \ldots\ldots a^4b^1 + \ldots\ldots a^3b^2 + \ldots\ldots a^2b^3 + \ldots\ldots ab^4 + \ldots\ldots a^0\,b^5 \qquad [10]$$

DALLA STORIA

Sembra che il triangolo di Tartaglia sia apparso per la prima volta in lavori di matematici islamici e cinesi del secolo XI. Tuttavia esso è diventato noto come «triangolo di Tartaglia», dal soprannome del matematico italiano *Nicolò Fontana* (1500-1559) che per primo lo usò sistematicamente, o come «triangolo di Pascal», dal nome del matematico francese *Blaise Pascal* (1623-1662) che ne scoprì molte proprietà.

Con GeoGebra
Potenza n-esima di un binomio

Restano da determinare i coefficienti che abbiamo provvisoriamente lasciato in sospeso ponendo dei puntini. In base alla regola enunciata poc'anzi, essi coincidono con i numeri della *quinta riga* del triangolo di Tartaglia. Dal momento che la quinta riga del triangolo di Tartaglia è:

1 5 10 10 5 1

possiamo completare il polinomio [10], ottenendo così che:

$$(a+b)^5 = 1 \cdot a^5b^0 + 5 \cdot a^4b^1 + 10 \cdot a^3b^2 + 10 \cdot a^2b^3 + 5 \cdot ab^4 + 1 \cdot a^0 b^5$$

Lo sviluppo della potenza del binomio secondo la formula di Newton

Il metodo dello sviluppo di $(a+b)^n$ basato sul triangolo di Tartaglia è efficiente per piccoli valori di n, ma diventa scomodo al crescere del valore di n, perché la costruzione dell'n-esima riga del triangolo di Tartaglia richiede la costruzione di *tutte* le righe precedenti l'ennesima. Ora che abbiamo presentato le prime nozioni di calcolo combinatorio, possiamo introdurre un nuovo metodo, che permette di superare questo inconveniente. Cominciamo con il ragionare su un caso semplice, lo sviluppo di $(a+b)^3$:

$$(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$$

Ciascun termine dello sviluppo di $(a+b)^3$ si ottiene dalla somma algebrica dei prodotti ottenuti scegliendo o a o b da ciascuno dei tre fattori:

$$(a+b)(a+b)(a+b)$$

e moltiplicando tra loro le variabili scelte.
Il coefficiente di a^3 è 1 perché a^3 è ottenuto solo dal prodotto aaa (corrispondente alla scelta della variabile a da ciascuno dei tre fattori).
Il coefficiente di a^2b, invece, è 3 perché a^2b si ottiene da tre prodotti: aab, aba e baa. In altre parole, i modi in cui può ottenersi a^2b sono tanti quante le permutazioni di 3 lettere, di cui due uguali ad a e una uguale a b: tali modi sono quindi $\dfrac{3!}{2! \cdot 1!} = 3$.

Questo ragionamento può essere generalizzato: lo sviluppo di $(a+b)^n$ ha come termini monomi di grado n, la cui parte letterale è della forma $a^{n-k}b^k$, con $k = 0, 1, \ldots, n$: il coefficiente del monomio $a^{n-k}b^k$ è uguale al numero delle possibili permutazioni di n lettere, di cui $n-k$ uguali ad a e k uguali a b, quindi è uguale a:

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

Ne segue il seguente teorema.

TEOREMA 7 | Formula del binomio di Newton

Sia n un numero intero positivo; allora per ogni $a, b \in \mathbf{R}$:

$$(a+b)^n = \binom{n}{0}a^n + \binom{n}{1}a^{n-1}b + \binom{n}{2}a^{n-2}b^2 + \ldots + \binom{n}{n-1}ab^{n-1} + \binom{n}{n}b^n \quad [11]$$

OSSERVA
Il coefficiente del monomio $a^{n-k}b^k$ è uguale a $\binom{n}{k}$, dove:
– n è il *grado* del monomio;
– k è l'esponente di b.
In modo equivalente, si sarebbe potuto sostituire k con l'esponente di a, in virtù della proprietà:
$\binom{n}{k} = \binom{n}{n-k}$

L'uguaglianza [11] può essere espressa sinteticamente nella forma:

$$\sum_{k=0}^{n} \binom{n}{k} a^{n-k}\, b^k$$

La formula del binomio di Newton spiega perché il numeri $\binom{n}{k}$ vengono chiamati coefficienti *binomiali*: il motivo risiede nel fatto che questi numeri sono i coefficienti dello sviluppo della potenza n-esima del *binomio* $(a+b)$.

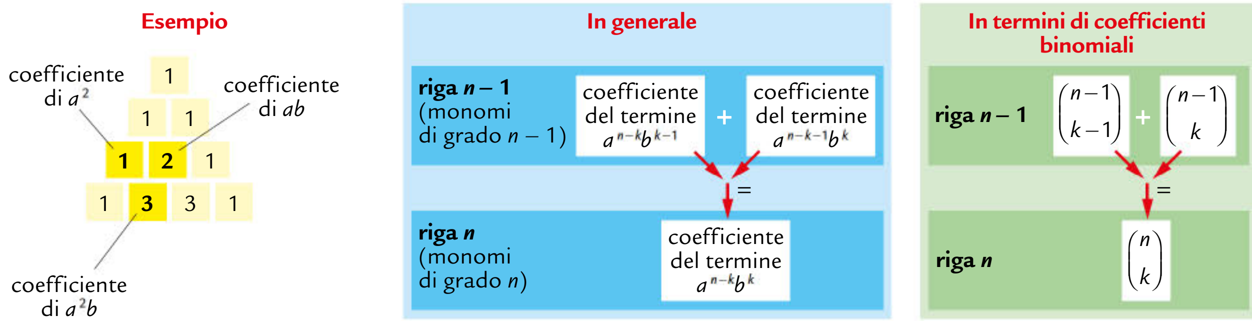
ESEMPIO Sviluppo di un binomio secondo la formula del binomio di Newton

Calcoliamo $(a+b)^5$, secondo la formula del binomio di Newton.

$$(a+b)^5=\binom{5}{0}a^5+\binom{5}{1}a^4b+\binom{5}{2}a^3b^2+\binom{5}{3}a^2b^3+\binom{5}{4}ab^4+\binom{5}{5}b^5=$$

$$=1\,a^5+5a^4b+10a^3b^2+10a^2b^3+5\,ab^4+b^5$$

La regola di costruzione del triangolo di Tartaglia può essere riletta in termini di coefficienti binomiali, come illustrato nel seguente schema.



Se ne deduce la seguente proprietà, che ti invitiamo a dimostrare algebricamente per esercizio, utilizzando la definizione di coefficiente binomiale:

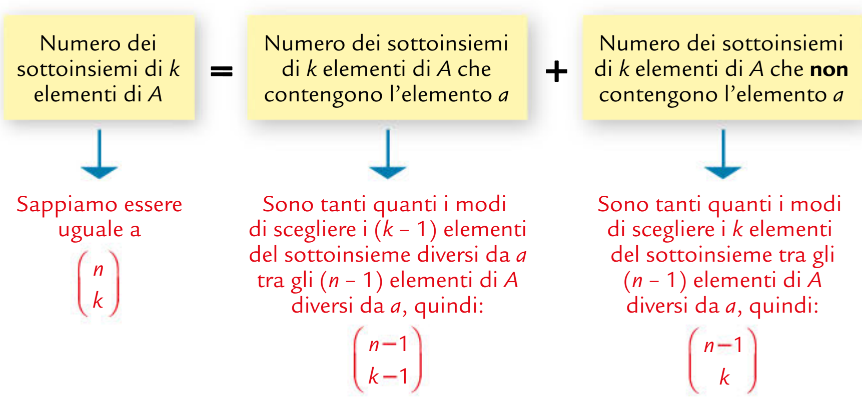
$$\binom{n}{k}=\binom{n-1}{k-1}+\binom{n-1}{k} \qquad [12]$$

MODI DI DIRE

L'uguaglianza [12] è anche nota come «formula di Stifel» in omaggio al matematico tedesco Michael Stifel (1487-1567).

IN UN ALTRO MODO **Una diversa deduzione dell'identità [12]**

Si può dedurre la [12] anche ragionando da un diverso punto di vista che ora illustriamo.
Consideriamo un insieme A di n elementi e fissiamo uno di questi elementi, che chiamiamo a; osserviamo ora che:



Dunque traducendo lo schema in termini di coefficienti binomiali, come spiegato nelle note in rosso, otteniamo l'identità [12].

Esercizi p. 130

Calcolo combinatorio

Principio fondamentale del calcolo combinatorio

Se un oggetto è univocamente individuato da una sequenza di n scelte successive, tali che vi siano k_1 possibilità per la prima scelta, k_2 per la seconda, ... , k_n per la n-esima, il numero totale di oggetti che si possono formare con tali scelte è il prodotto:

$$k_1 \cdot k_2 \ \ \ldots\ldots \cdot k_n$$

Conteggio di oggetti **ordinati**.

Disposizioni semplici

Ogni sequenza *ordinata* di k oggetti, scelti fra n oggetti distinti, con il vincolo di **non** *ripetere* gli oggetti.
Il numero complessivo di disposizioni semplici di n oggetti in k posti, con $k \leq n$, è dato dalla formula:

$$D_{n,k} = \underbrace{n(n-1) \cdot \ldots\ldots \cdot (n-k+1)}_{k \text{ fattori decrescenti}}$$

ESEMPIO

I modi in cui 3 persone possono sedersi su 4 sedie disposte in fila sono:

$$D_{4,3} = 4 \cdot 3 \cdot 2 = 24$$

Disposizioni con ripetizione

Ogni sequenza *ordinata* di k oggetti, scelti fra n oggetti distinti, con la possibilità di *ripetere* gli oggetti.
Il numero complessivo di disposizioni con ripetizione di n oggetti in k posti è dato dalla formula:

$$D^*_{n,k} = n^k$$

ESEMPIO

I possibili esiti ottenuti lanciando 3 volte un dado sono:

$$D^*_{6,3} = 6^3 = 216$$

Permutazioni

Ogni possibile ordinamento di n oggetti distinti.
È un caso particolare di disposizione semplice in cui $n = k$. Le permutazioni di n oggetti sono date dalla formula:

$$P_n = \underbrace{n(n-1) \cdot (n-2) \cdot \ldots\ldots \cdot 1}_{\text{fattoriale di } n \text{, indicato con } n!}$$

ESEMPIO

I modi in cui 4 persone possono disporsi in fila indiana sono:

$$P_4 = 4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$$

Permutazioni con ripetizione

Dati n oggetti di cui a_1 uguali tra loro, a_2 uguali tra loro (e distinti dai precedenti), ... , a_k uguali tra loro e distinti dai precedenti, con:

$$a_1 + a_2 + \ldots\ldots + a_k = n$$

le permutazioni di questi n oggetti sono:

$$\frac{n!}{a_1! \cdot a_2! \cdot \ldots\ldots \cdot a_k!}$$

ESEMPIO

Gli anagrammi, anche senza significato, della parola «cocco» sono:

$$\frac{5!}{3!\ 2!} = \frac{5 \cdot 4 \cdot \cancel{3!}}{\cancel{3!} \cdot 2!} = 10$$

Conteggio di oggetti **NON** ordinati.

Combinazioni semplici

Ogni raggruppamento **non** *ordinato* di k oggetti, scelti fra n oggetti distinti, con il vincolo di **non** *ripetere* gli oggetti.
Il numero complessivo di combinazioni semplici di n oggetti di classe k è dato dalla formula:

$$C_{n,k} = \underbrace{\binom{n}{k}}_{\text{coefficiente binomiale}} = \frac{D_{n,k}}{k!}$$

ESEMPIO

I sottoinsiemi di 3 elementi di un insieme di 5 elementi sono:

$$C_{5,3} = \binom{5}{3} = \frac{D_{5,3}}{3!} = \frac{5 \cdot 4 \cdot 3}{3!} = 10$$

Combinazioni con ripetizione

Ogni raggruppamento **non** *ordinato* di k oggetti, scelti fra n oggetti distinti, con la possibilità *ripetere* gli oggetti. Il numero complessivo di combinazioni con ripetizione di n oggetti di classe k è dato dalla formula:

$$C^*_{n,k} = \binom{n+k-1}{k}$$

ESEMPIO

Una urna contiene 10 palline numerate da 1 a 10. Si effettuano tre estrazioni con reimmissione. Senza tenere conto dell'ordine di estrazione, i possibili modi di estrarre le tre palline sono:

$$C^*_{10,3} = \binom{10+3-1}{3} = \frac{12 \cdot 11 \cdot 10}{3!} = 220$$

Coefficienti binomiali

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$ **Formula dei tre fattoriali**

$$\binom{n}{k} = \binom{n}{n-k}$$ **Legge delle classi complementari**

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$ **Formula di Stifel**

ESEMPI

$$\binom{5}{3} = \frac{5!}{3!\,2!}$$

$$\binom{9}{3} = \binom{9}{6}$$

$$\binom{10}{4} = \binom{9}{3} + \binom{9}{4}$$

Formula di Newton

$$(a+b)^n =$$

$$= \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k =$$

$$= \binom{n}{0}a^n + \binom{n}{1}a^{n-1}b + \ldots\ldots + \binom{n}{n}b^n$$

ESEMPIO

$$(a+b)^4 =$$

$$= \binom{4}{0}a^4 + \binom{4}{1}a^3b + \binom{4}{2}a^2b^2 + \binom{4}{3}ab^3 + \binom{4}{4}b^4$$

1. Introduzione al calcolo combinatorio

Teoria p. 98

Esercizi introduttivi

Per ciascuno dei seguenti problemi, stabilisci se si può assimilare al conteggio di parole ordinate o non ordinate, con o senza ripetizioni. Non è richiesto di risolvere il problema.

1 Un club ha 50 membri, di cui 30 uomini e 20 donne. Si vuole costruire una delegazione di membri del club, costituita da 5 uomini e 5 donne. Quante sono le delegazioni possibili?

2 Quanti numeri interi ci sono, compresi tra 100 e 999, aventi cifre tutte distinte?

3 La combinazione di una cassaforte è costituita da cinque cifre (ciascuna delle quali può essere un numero qualsiasi compreso tra 0 e 9). Quante combinazioni sono possibili?

4 Quanti sono i sottoinsiemi di 3 elementi di un insieme di 10 elementi?

5 Quanti sono i possibili anagrammi (anche privi di significato) della parola «computer»?

6 A una gara partecipano 10 atleti, in quanti modi si può presentare la classifica dei primi tre?

7 In un torneo ogni squadra affronta ciascuna delle altre una e una sola volta. Nel torneo quante partite vengono disputate?

8 Si deve confezionare un sacchetto contenente 20 caramelle, ai gusti di limone, fragola o arancia. In quanti modi si può confezionare il sacchetto, ammettendo che le caramelle in esso contenute possano essere sia di un solo gusto, sia di due gusti, sia di tutti e tre i gusti?

Risolvi i seguenti problemi, utilizzando un diagramma ad albero.

9 Per raggiungere una data località si ha a disposizione il treno o l'aereo. Nel primo caso, per arrivare a destinazione, alla stazione si può scegliere tra un pullman o il taxi o 2 km a piedi. Nel secondo caso all'aeroporto si può noleggiare un'auto o prendere un taxi. In quanti modi diversi si può giungere a destinazione? [5]

10 Il menu di una trattoria offre le seguenti possibilità di scelta. Primi: penne al pesto, risotto ai funghi o gnocchi di patate; secondi: frittura di pesce, braciola di maiale, pollo allo spiedo; infine come dessert si può scegliere tra torta o macedonia. Quanti sono i pranzi completi distinti che la trattoria offre? [18]

11 Il tenente Colombo ama i ragionamenti articolati. O il sospettato confessa o non confessa. Se confessa, si può arrestare; se non confessa, o si contraddice o è coerente. Se si contraddice, si può arrestare; se è coerente o ha un alibi o non ha un alibi. Se ha un alibi si può scagionare; se non ha un alibi si sottoporrà al test del DNA. Se questo è positivo, il sospettato sarà arrestato, altrimenti sarà liberato. In quanti casi si arriva all'arresto? [3]

Problemi sul principio fondamentale del calcolo combinatorio

12 ESERCIZIO GUIDATO

Quante sono le password formate da 5 lettere dell'alfabeto italiano che iniziano con una consonante e finiscono con due vocali, ammettendo che le lettere possano essere ripetute?

Puoi costruire la password con cinque scelte successive:

- per la prima scelta hai 16 possibilità (le 16 consonanti)
- per la seconda scelta hai 21 possibilità (le 21 lettere dell'alfabeto italiano)
- per la terza scelta hai .... possibilità
- per la quarta scelta hai .... possibilità
- per la quinta scelta hai .... possibilità

Per il principio fondamentale del calcolo combinatorio le parole che soddisfano le condizioni richieste sono in totale: $16 \cdot 21^2 \cdot$ .... $= 176\,400$

13 Il comandante di un plotone di 12 soldati deve garantire un turno di guardia all'ingresso principale, all'armeria, all'autorimessa e all'ingresso secondario. In quanti modi diversi può disporre i suoi uomini? [11 880]

14 Il signor Aristide, partendo per una gita, porta con sé quattro magliette di colore diverso, due giacche e tre paia di pantaloni. In quanti modi diversi può vestirsi il sig. Aristide? [24]

15 **E se?** La targa di un motorino è costituita da due lettere, due numeri e due lettere. Supponendo che le lettere possano essere scelte a caso dalle 26 dell'alfabeto anglosassone e che ciascun numero possa essere una cifra qualsiasi tra 0 e 9, quanti motorini diversi si possono immatricolare?

▶ Come cambia la risposta se si considerano le autovetture? Ricorda che la targa di un'auto, rispetto a quella di un motorino, ha tre numeri anziché due. [45 697 600; 456 976 000]

16 Quante password diverse di otto caratteri si possono generare utilizzando per ciascun carattere una cifra compresa tra 0 (incluso) e 9 (incluso)? Aumentando il numero dei caratteri da 8 a 9, di quanto aumenta il numero delle password possibili? [10^8; $9 \cdot 10^8$]

17 Quante diverse password di otto caratteri si possono generare utilizzando per ciascun carattere una lettera (scelta dalle 26 dell'alfabeto anglosassone) o una cifra (compresa tra 0 e 9, incluso 0 e 9)? Aumentando il numero dei caratteri da 8 a 9, di quanto aumenta il numero delle password possibili? [36^8; $35 \cdot 36^8$]

18 Quanti numeri di quattro cifre puoi formare con le cifre 0, 1, 2, 3, 4 e 5?
(*Suggerimento*: presta attenzione al fatto che un numero **non** può iniziare per zero) [$5 \cdot 6^3$]

19 Quanti numeri di tre cifre, tutte dispari, si possono scrivere? E quanti numeri dispari, costituiti da tre cifre? [125; 450]

20 Quanti numeri costituiti da sei cifre distinte possono essere scritti, utilizzando le cifre da 0 a 9? [136 080]

21 Quanti numeri di cinque cifre, tutte pari e diverse da zero, si possono scrivere? [1024]

22 Quanti anagrammi che iniziano con la lettera «G» possono essere composti dalla parola «gesto»? [24]

23 Una segretaria, che ha un'età compresa tra 45 e 50 anni (potendo anche essere uguale a 45 o a 50 anni), dice a tutti che è del segno della Bilancia. Se usa la sua data di nascita come password, quanti tentativi al massimo bisogna fare per scoprirla?
(Ricorda che cadono sotto il segno della Bilancia i nati dal 23 settembre (incluso) al 22 ottobre (incluso)). [180]

24 Si vuole scrivere un numero di tre cifre, con le seguenti caratteristiche: la prima cifra a partire da sinistra deve essere dispari, la seconda pari (0 incluso) e l'ultima deve essere un multiplo di 3 (diverso da zero). Quanti numeri distinti si possono scrivere con queste caratteristiche? [75]

25 **Realtà e modelli** Per evitare che gli utenti usino password insicure, come date di nascita o nomi propri di persona, i sistemi di sicurezza informatica normalmente riconoscono password alfanumeriche solo se contenenti *sia* lettere dell'alfabeto *sia* cifre.

a. Calcola il numero delle password costituite da 6 caratteri (26 lettere dell'alfabeto anglosassone e 10 cifre) e riconosciute come valide (cioè sicure). Non distinguere tra carattere minuscolo e maiuscolo.

b. Nel complesso delle password alfanumeriche costituite da 6 caratteri, qual è la percentuale di quelle respinte perché contenenti caratteri solo alfabetici o solo numerici?

[**a.** $36^6 - (26^6 + 10^6)$; **b.** circa il 14%]

26 Un'urna contiene 10 biglie numerate da 1 a 10. In quanti modi diversi si possono estrarre 3 biglie dall'urna, tenendo conto dell'ordine di estrazione, in ciascuno dei seguenti tre casi:

a. ciascuna biglia, dopo essere estratta, viene rimessa nell'urna prima dell'estrazione successiva;

b. ciascuna biglia, dopo esser estratta, **non** viene rimessa nell'urna prima dell'estrazione successiva;

c. la prima biglia estratta viene rimessa nell'urna prima della seconda estrazione, mentre la seconda biglia estratta **non** viene rimessa nell'urna prima della terza estrazione. [**a.** 1000; **b.** 720; **c.** 900]

27 Un'urna contiene 8 biglie numerate da 1 a 8. In quanti modi diversi si possono estrarre 4 biglie dall'urna, tenendo conto dell'ordine di estrazione, in ciascuno dei seguenti tre casi:

a. ciascuna biglia, dopo essere estratta, viene rimessa nell'urna prima dell'estrazione successiva;

b. ciascuna biglia, dopo esser estratta, **non** viene rimessa nell'urna prima dell'estrazione successiva;

c. la prima biglia estratta viene rimessa nell'urna prima della estrazione successiva, mentre la seconda e la terza biglia estratta **non** vengono rimesse nell'urna prima della estrazione successiva. [**a.** 4096; **b.** 1680; **c.** 2688]

Collegamenti **Calcolo combinatorio e aritmetica**

28 Il numero 189, scomposto in fattori primi, è uguale a $3^3 \cdot 7$. Servendoti del principio fondamentale del calcolo combinatorio, indica quanti sono i divisori di 189. [8]

29 Il numero 3240, scomposto in fattori primi, è uguale a $2^3 \cdot 3^4 \cdot 5$. Servendoti del principio fondamentale del calcolo combinatorio, indica quanti sono i divisori di 3240. [40]

Matematica e informatica

30 Nei calcolatori elettronici la memoria è costituita da tanti «bit», essendo un «bit» un dispositivo che può assumere due stati differenti (per esempio: essere o non essere magnetizzato). Rappresentando i due stati che può assumere un bit con «0» e «1», un bit si può assimilare a «cella» di memoria che può assumere, di volta in volta, o il valore 0 o il valore 1. Una sequenza (ordinata) di 8 bit costituisce un «byte».

a. Quanti diversi «valori» può assumete un byte (ovvero quante diverse sequenze ordinate di otto simboli, costituiti da 0 o 1, è possibile costruire)?

b. Quanti sono i byte il cui primo bit è uguale a 0?

c. Quanti sono i byte che terminano con 11?

d. Quanti sono i byte che iniziano con 0 oppure terminano con 11? [**a.** 256; **b.** 128; **c.** 64; **d.** 160]

31 Ciascuna «informazione» viene rappresentata in un computer tramite una sequenza di bit che la identifica in modo univoco (ossia tramite una sequenza di simboli, ciascuno dei quali può essere soltanto 0 o 1). Il processo che fa corrispondere a una informazione una sequenza di bit prende il nome di *codifica* dell'informazione.

a. Qual è il massimo numero di differenti informazioni che si possono codificare tramite 6 bit (ovvero quante diverse sequenze ordinate di sei simboli, costituiti da 0 o 1, è possibile costruire)?

b. Qual è il massimo numero di differenti informazioni che si possono codificare tramite 2 byte (vedi la definizione di byte nell'esercizio precedente)?

c. Nel sistema di codifica «unicode» (nella sua prima formulazione) ciascun simbolo dell'alfabeto anglosassone era codificato tramite una sequenza di 16 bit; in questo sistema di codifica, quanti byte occupa la parola «matematica»?

d. Quanti bit sono necessari, come minimo, per poter codificare 8000 informazioni differenti?

[**a.** 64; **b.** 2^{16}; **c.** 20 ; **d.** 13]

2. Disposizioni e permutazioni

Teoria p. 101

Esercizi introduttivi

Test

32 **Vero o falso?**

a. le disposizioni di n oggetti in n posti sono le permutazioni degli n oggetti V F

b. le disposizioni di 5 oggetti in 4 posti si indicano con il simbolo $D_{5,4}$ V F

c. gli anagrammi della parola «cielo» sono 5 V F

d. gli anagrammi della parola «cielo» sono 5!, tanti quanti gli anagrammi della parola «terra», perché entrambe le parole sono costituite da cinque lettere V F

e. la scrittura $D_{3,5}$ non ha significato V F

[3 affermazioni vere e 2 false]

33 Quanto vale $D_{5,3}$, ovvero quante sono le disposizioni di 5 oggetti in 3 posti?

A 5^3 B 3^5 C $5 \cdot 4 \cdot 3$ D $5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$

34 Quanto vale 5!?

A 5^5 B $5 \cdot 4 \cdot 3$ C $5 \cdot 5$ D $5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$

35 Sapendo che $9! = 362\,880$, quanto vale 10!?

A 3 628 800 B 725 760 C 36 288 000 D nessuno dei precedenti

36 Quanto vale $\frac{n!}{(n-1)!}$?

A $n-1$ B n C $n+1$ D $2n$

37 Quanto vale $n! + (n+1)!$?

A $(2n+1)!$ B $(n+2)n!$ C $(n+2)!$ D $(2n+1)n!$

Le disposizioni

38 ESERCIZIO SVOLTO

Calcoliamo i seguenti numeri:

a. $D_{7,3}$ **b.** $D_{10,6}$ **c.** $D^*_{4,3}$

a. $D_{7,3} = \underbrace{7 \cdot 6 \cdot 5}_{\text{3 fattori decrescenti}} = 210$ **b.** $D_{10,6} = \underbrace{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6 \cdot 5}_{\text{6 fattori decrescenti}} = 151\,200$ **c.** $D^*_{4,3} = 4^3$

Calcola il valore delle seguenti espressioni.

39 $D_{10,4}$ [5040]

40 $D_{6,4}$ [360]

41 $D^*_{5,3}$ [125]

42 $D^*_{3,4}$ [81]

43 $D^*_{6,5} : D_{6,5}$ [10,8]

44 $D_{10,3} : D_{6,2}$ [24]

45 Si sa che le disposizioni di n elementi in 2 posti sono 132. Quanto vale n? [12]

46 ESERCIZIO GUIDATO

a. **In quanti modi diversi possono essere sistemati su una libreria 6 libri, scelti tra 10, a disposizione?**

b. **Quanti numeri di 4 cifre, tutte dispari, si possono scrivere?**

a. Si tratta di disporre 10 libri in 6 posti, quindi il numero totale di disposizioni possibili è dato da:

$D_{10,6} =$ .... $=$ ....

In alternativa, puoi ragionare applicando il principio fondamentale del calcolo combinatorio, osservando che ci sono 10 possibilità per il primo libro, 9 per il secondo, ... per il terzo, ... per il quarto, ... per il quinto, ... per il sesto.

b. I numeri descritti si possono assimilare alle *disposizioni con ripetizione* dei cinque numeri 1, 3, 5, 7, 9 in quattro posti; quindi sono in totale:

$D^*_{5,4} = 5^4 =$ .....

In alternativa, puoi ragionare applicando il principio fondamentale del calcolo combinatorio, osservando che ci sono 5 possibili scelte per ciascuna delle 5 cifre.

47 In quanti modi quattro insegnanti possono coprire due ore di lezione in una classe (possono esserci anche due ore dello stesso insegnante)? [16]

48 In una società di 30 persone si devono eleggere un coordinatore, un segretario e un tesoriere. Quante sono le scelte possibili? [24 360]

49 **Videolezione** Un club ha 15 membri. In quanti modi possono essere scelti un presidente, un vice-presidente e un segretario (supponendo che nessun membro possa avere più di una carica)? [2730]

50 Lanciando un dado per 5 volte consecutive, quante sono le possibili sequenze ordinate di numeri che si possono ottenere? [6^5]

51 In un cinema, una fila di poltrone ha 20 posti. Arrivano solo 6 persone a sedersi su quella fila. In quanti modi possono disporsi? [27 907 200]

52 In quanti modi si possono sistemare quattro ospiti in un albergo che ha cinque stanze singole libere? [120]

53 **Videolezione** Quanti numeri di 6 cifre, tutte pari e diverse da zero, si possono scrivere? [4096]

54 Quanti numeri di 4 cifre distinte, tutte pari e diverse da zero, si possono scrivere? [24]

Permutazioni

55 **ESERCIZIO SVOLTO**

Calcoliamo i seguenti numeri:

a. $P_2 \cdot P_5$ **b.** $\dfrac{10!}{8!}$

a. $2! \cdot 5! = (2 \cdot 1) \cdot (5 \cdot 4 \cdot 3 \cdot 2 \cdot 1) = 2 \cdot 120 = 240$

b. $\dfrac{10!}{8!} = \dfrac{10 \cdot 9 \cdot 8!}{8!} = 10 \cdot 9 = 90$

Calcola il valore delle seguenti espressioni.

56 $P_8 : P_6$ [56]

57 $P_4 \cdot P_3$ [144]

58 $\dfrac{18!}{16!}$ [306]

59 $\dfrac{7!}{5!2!}$ [21]

60 $\dfrac{10!}{6!4!}$ [210]

61 **Videolezione** $\dfrac{P_6}{D_{5,3}}$ [12]

62 $\dfrac{P_6}{P_4} : D_{6,2}$ [1]

A mente

63 $3!$; $4!$

64 $\dfrac{10!}{9!}$; $\dfrac{8!}{6!}$

65 $0! + 1! + 2!$

66 $\dfrac{5001!}{5000!}$; $\dfrac{999!}{1000!}$

67 M.C.D. $(12!, 15!, 20!)$

68 La cifra delle unità di $999!$

Verifica le seguenti identità.

69 $\dfrac{(n+1)!}{(n-1)!} = n(n+1)$

70 $\dfrac{(2n-2)!}{(2n+1)!} = \dfrac{1}{2n(4n^2-1)}$

71 $\dfrac{(2n+1)!}{(2n-1)!} = 4n^2 + 2n$

72 $\dfrac{(n-1)!}{n!} - \dfrac{(n-1)!}{(n+1)!} = \dfrac{1}{n+1}$

73 $\dfrac{(n-1)!}{n!} - \dfrac{n!}{(n+1)!} = \dfrac{1}{n^2+n}$

74 $\dfrac{(n+2)!}{n!} - \dfrac{(n+2)!}{(n+1)!} = n^2 + 2n$

75 $\dfrac{1}{n!} + \dfrac{1}{(n-1)!} + \dfrac{1}{(n-2)!} = \dfrac{n^2+1}{n!}$

76 ESERCIZIO SVOLTO

Risolviamo l'equazione $(n+2)! - 4(n+1)! = 8(n+1)!$.

- Affinché siano definiti tutti i fattoriali che compaiono nell'equazione n deve essere un numero intero tale che $n \geq -1$.
- In base alla definizione di fattoriale l'equazione equivale a:

$$(n+2)(n+1)! - 4(n+1)! = 8(n+1)!$$

- Poiché il fattoriale di un numero è sempre *diverso da zero*, è possibile dividere entrambi i membri dell'equazione per $(n+1)!$, ottenendo l'equazione equivalente seguente, che risolviamo:

$$(n+2) - 4 = 8 \quad \Rightarrow \quad n = 10$$

- La soluzione trovata è *accettabile*, perché soddisfa la condizione $n \geq -1$.

Risolvi le seguenti equazioni.

77 $(n+1)! - 2n! = 4n!$ [5]

78 $n! - 4(n-1)! = 10(n-1)!$ [14]

79 $(n+2)! - 4(n+1)! = 4(n+1)!$ [6]

80 $(n+3)! - (n+2)! = 9(n+2)!$ [7]

81 $(n+1)! - n! = 16(n-1)!$ [4]

82 $4(n+1)! - 8(n-1)! = 35n!$ [8]

83 $n! - (n-1)! = (n-1)!(n^2 - 6n + 5)$ [1, 6]

84 $\dfrac{[(n+1)!]^2 - (n!)^2}{(n!)^2} = \dfrac{3(n^2+4n)}{4}$ [0, 4]

85 $\dfrac{(n+4)(n+3)!}{(n+5)!} = \dfrac{n-9}{15}$ [10]

86 $\dfrac{3!(n+7)!}{(n+6)!} = 5n + 100$ [58]

87 $\dfrac{2!(n+6)!}{(n+5)(n+4)!} = 3n - 3$ [15]

88 $\dfrac{(n+5)! - (n+3)!}{(n+3)!} = n^2 + 37$ [2]

89 ESERCIZIO GUIDATO

a. Abbiamo 8 palline, di colori tutti diversi tra loro, e altrettante scatole numerate da 1 a 8. In quanti modi si possono disporre le 8 palline nelle 8 scatole, in modo che ogni scatola contenga esattamente una pallina?

b. Quanti anagrammi, anche privi di significato, si possono costruire con la parola «canotto»?

a. Disporre le palline nelle scatole equivale a definire un *ordinamento* delle 8 palline, in modo che la *prima* pallina sia inserita nella scatola 1, la seconda pallina sia inserita nella scatola 2, e così via... Quindi il numero totale di modi di disporre le palline è:

$$P_8 = 8! = \ldots..$$

b. Il problema equivale a calcolare il numero di *permutazioni* di 7 lettere, di cui la «o» è *ripetuta* 2 volte, la «t» è *ripetuta* 2 volte mentre le rimanenti compaiono una sola volta. Il numero totale di anagrammi è dunque:

$$\frac{7!}{2!2!} = \ldots\ldots..$$

90 In quanti modi possono essere disposti su una libreria 6 libri distinti? [720]

91 Quante diverse classifiche finali può avere una gara ciclistica alla quale partecipano 8 atleti (escludendo che ci siano *ex-aequo*)? [40 320]

92 Trova il numero di anagrammi della parola «remo». [24]

93 Trova il numero di anagrammi della parola «Milano». [720]

94 Trova il numero di anagrammi della parola «Toronto». [420]

95 Trova il numero di anagrammi della parola «Sandra». [360]

96 Su uno scaffale di una libreria ci sono 6 libri di cui 2 sono copie identiche dello stesso libro, mentre gli altri libri sono tutti differenti. In quanti modi diversi è possibile disporre i sei libri sullo scaffale? [360]

97 Determina quanti numeri di cinque cifre esistono aventi le stesse cifre del numero 16 306 (ciascuna cifra deve comparire nel numero lo stesso numero di volte con cui compare nel numero 16 306). [48]

98 **E se?** Una partita di calcio tra la squadra A e la squadra B è finita 3 a 2. In quanti modi diversi possono essersi succedute le reti?

▶ Come cambierebbe la risposta sapendo che la squadra A è sempre stata in vantaggio (a parte lo 0-0 iniziale, ovviamente)? E sapendo che nel corso della partita la squadra A si è trovata in svantaggio almeno una volta? [10; 2; 5]

99 **ESERCIZIO GUIDATO**

5 italiani, 4 francesi e 2 tedeschi devono sedersi in fila. Le persone di stessa nazionalità devono rimanere vicine. In quanti modi si possono disporre?

- Ci sono 3! permutazioni possibili dei tre *gruppi* costituiti dagli italiani, dai francesi e dai tedeschi.
- All'interno di ciascun gruppo, si devono poi ordinare le persone: ci sono 5! permutazioni degli italiani, 4! dei francesi e 2! dei tedeschi.
- In totale i modi di disporsi sono: $3! \cdot 5! \cdot 4! \cdot 2! =$ ....

100 Sei amici, tra cui Paola e Marco, si recano al cinema e si dispongono su una stessa fila, in posti adiacenti. In quanti modi possono disporsi, se Paola vuole stare vicina a Marco? [240]

101 Cinque amici, tra cui Valeria e Luisa, si recano al cinema e si dispongono su una stessa fila, in posti adiacenti. Valeria e Luisa però hanno litigato, per cui non vogliono stare vicine. In quanti modi possono disporsi? [72]

102 Tre italiani, due francesi e due inglesi devono sedersi in fila. In quanti modi possono farlo se le persone della stessa nazionalità devono stare vicine? [144]

103 Barbara vuole sistemare su un ripiano vuoto della sua libreria 8 libri (tutti diversi tra loro). Fra gli otto libri ci sono i tre libri della trilogia del Signore degli anelli. Determina in quanti modi Barbara può disporre i libri:

a. se essi possono essere sistemati in ordine qualunque;

b. se i tre libri della trilogia devono essere messi vicini tra loro ed esattamente nell'ordine della trilogia;

c. se i tre libri della trilogia devono essere messi vicini tra loro, ma possono essere disposti in qualsiasi ordine.

[**a.** 40 320; **b.** 720; **c.** 4320]

104 Paolo vuole sistemare su un ripiano vuoto della sua libreria 4 libri di letteratura, 3 libri di storia e 1 libro di matematica. Determina in quanti modi può disporre i libri:

a. se essi possono essere sistemati in qualunque ordine;

b. se i libri di letteratura vanno messi vicini tra loro e i libri di storia vanno messi vicini tra loro;

c. se i libri di letteratura vanno messi vicini tra loro, mentre gli altri libri possono essere sistemati in qualunque ordine. [**a.** 40 320; **b.** 864; **c.** 2880]

105 Un gioco per bambini è costituito da 10 blocchi di legno. Determina in quanti modi il bambino può allineare i 10 blocchi:

a. supponendo che sono tutti di colori diversi;

b. supponendo che, tra i 10 pezzi di legno, 5 sono gialli, 3 sono rossi, 2 sono blu e che i pezzi dello stesso colore sono indistinguibili uno dall'altro;

c. supponendo che i 10 pezzi siano dei colori descritti al punto **b**, e che i pezzi dello stesso colore debbano essere posti vicini tra loro;

d. supponendo che i 10 pezzi siano dei colori descritti al punto **b**, e che soltanto i pezzi gialli debbano essere posti vicini tra loro. [**a.** 3 628 800; **b.** 2520; **c.** 6; **d.** 60]

Esercizi riassuntivi: disposizioni e permutazioni

106 In una società di 20 soci, devono essere scelti un presidente, un vicepresidente e un segretario. In quanti modi si può fare la scelta? [6840]

107 Nonno Anselmo va al mare in treno: ricorda che dopo Firenze ci sono quattro fermate, ha in mente i nomi delle stazioni, ma non ricorda in quale ordine si succedono. In quanti modi possibili nonno Anselmo può incontrare le quattro stazioni? [24]

108 Considera i risultati di 14 partite di calcio: per ogni partita se vince la prima squadra si segna 1, se le due squadre pareggiano si segna X, se vince la seconda squadra si segna 2. Si ottiene, in questo modo, una sequenza di 14 segni che chiamiamo colonna. Quante colonne si possono ottenere? [3^{14}]

109 A una manifestazione si schierano sotto il palco 10 militari: 2 bersaglieri, 3 artiglieri, 2 avieri e 3 marinai. I militari che appartengono alla stessa specialità hanno la stessa altezza e divise identiche, perciò si possono supporre indistinguibili uno dall'altro. In quanti modi i 10 militari si possono disporre? [25 200]

110 Si vogliono dipingere quattro pareti di una stanza con quattro colori diversi, scelti fra azzurro, arancione, verde, giallo, bianco. In quanti modi è possibile farlo? [120]

111 Quanti sono i possibili risultati che si possono ottenere eseguendo cinque lanci successivi di un dado? [7776]

112 Cinque ballerini e altrettante ballerine si accingono a ballare una polka (ballo a due). Quanti accoppiamenti sono possibili? [120]

113 Nella prima fila di un'aula devono sedersi sei studenti: quattro ragazzi e due ragazze. Determina in quanti modi possono disporsi:

a. se possono sistemarsi in ordine qualunque;

b. se i ragazzi devono stare vicini tra loro e le ragazze devono stare vicine tra loro;

c. se le ragazze devono stare vicine tra loro, mentre i ragazzi possono disporsi in ordine qualunque.

[**a.** 720; **b.** 96; **c.** 240]

114 Considera dieci carte, cinque di cuori e cinque di picche, tutte di valori diversi. Quante sono le permutazioni di queste dieci carte tali che due carte consecutive comunque scelte hanno sempre colori differenti? [28 800]

115 Un vigile che vede un automobilista commettere una grave infrazione non riesce a fermare l'autoveicolo ma si accorge che la macchina è di colore grigio, che le prime due lettere della targa sono C e A e che gli ultimi due numeri della targa sono 3 e 5. La targa è costituita da due lettere, seguite da tre cifre, seguite da altre due lettere; ciascuna lettera può essere una qualsiasi dell'alfabeto inglese di 26 lettere a eccezione di I, O, Q e U, mentre ciascuna delle cifre può essere un numero intero qualsiasi compreso tra 0 e 9 (inclusi 0 e 9). Supponendo che le macchine grigie siano un quarto del parco circolante, quanti sono approssimativamente i possibili colpevoli dell'infrazione? [1210]

Collegamenti **Calcolo combinatorio e funzioni**

116 **ESERCIZIO GUIDATO**

a. **Dati gli insiemi $A = \{1, 2, 3, 4\}$ e $B = \{1, 2, 3, 4, 5\}$, quante diverse funzioni $f : A \to B$ è possibile definire?**

b. **Quante funzioni *iniettive* $f : A \to B$ è possibile definire?**

c. **Quante sono le funzioni *biettive* $f : A \to A$?**

a. Abbiamo 5 possibili modi di assegnare a $x_1 = 1$ la sua immagine tramite f; altri 5 modi di assegnare l'immagine a $x_2 = 2$ e così via. Pertanto il numero delle funzioni da A a B è $D^*_{5,4} = ... = ...$

b. Abbiamo 5 possibili modi di assegnare a $x_1 = 1$ la sua immagine tramite f; abbiamo 4 modi (non più 5, dal momento che per definizione una funzione iniettiva associa ad elementi distinti del dominio elementi distinti del codominio) di assegnare a $x_2 = 2$ la sua immagine, e così via. Pertanto il numero delle funzioni iniettive da A a B è $D_{5,4} = ... = ...$

c. Una funzione biettiva $f : A \to A$ deve avere come insieme delle immagini l'insieme A stesso; dunque per definire la funzione è sufficiente ordinare gli elementi di A in modo che il primo elemento sia il corrispondente di x_1, il secondo sia il corrispondente di x_2 e così via. Perciò il numero delle funzioni biettive $f : A \to A$ è uguale al numero delle permutazioni degli elementi di A, cioè: .... = ... = ... [**a.** 625; **b.** 120; **c.** 24]

117 Quante funzioni si possono definire aventi come dominio l'insieme $A = \{1, 2, 3, 4\}$ e come insieme delle immagini un sottoinsieme di $B = \{1, 2, 3, 4, 5\}$? [625]

118 Quante funzioni biettive si possono definire aventi come dominio e codominio l'insieme $A = \{1, 2, 3, 4, 5\}$? [120]

119 Quante funzioni iniettive si possono definire aventi come dominio l'insieme $A = \{1, 2, 3\}$ e come codominio l'insieme $B = \{1, 2, 3, 4, 5, 6\}$? [120]

120 Quante funzioni suriettive si possono definire aventi come dominio l'insieme $A = \{1, 2, 3, 4\}$ e come codominio l'insieme $B = \{1, 2, 3\}$? [36]

3. Combinazioni

Teoria p. 105

Esercizi introduttivi

Test

121 Il numero $\binom{5}{3}$ è uguale a:

A $\dfrac{5 \cdot 4 \cdot 3 \cdot 2 \cdot 1}{3!}$ B $\dfrac{5 \cdot 4 \cdot 3}{3!}$ C $\dfrac{5 \cdot 4}{3!}$ D $\dfrac{5!}{2!}$

122 Quale dei seguenti coefficienti binomiali è uguale a n, per ogni $n \in \mathbf{N} - \{0\}$?

A $\binom{n}{0}$ B $\binom{n}{1}$ C $\binom{n}{n}$ D Nessuno dei precedenti

123 Il coefficiente binomiale $\binom{n}{k}$:

A non è mai uguale a 0

B è uguale a 1 se e solo se $k = n$

C se $k = 0$, non è definito

D se $n > k$, non è definito

124 Il coefficiente binomiale $\binom{n}{k}$ è uguale a:

A $\dfrac{n!}{(n-k)!}$ B $\dfrac{n!(n-k)!}{k!}$ C $\dfrac{n!}{k!}$ D $\dfrac{n!}{(n-k)!k!}$

125 Quale delle seguenti espressioni esprime il numero dei sottoinsiemi di 3 elementi di un insieme di 6 elementi?

A $\dfrac{6!}{3!}$ B $\binom{6}{3}$ C $\dfrac{6!}{3}$ D $\binom{3}{6}$

Combinazioni semplici

126 ESERCIZIO SVOLTO

Calcoliamo:

a. $\binom{6}{3}$ **b.** $C_{10,4}$

---

a. $\binom{6}{3} = \dfrac{D_{6,3}}{3!} = \dfrac{6 \cdot 5 \cdot 4}{3 \cdot 2} = 20$ **b.** $C_{10,4} = \binom{10}{4} = \dfrac{D_{10,4}}{4!} = \dfrac{10 \cdot 9 \cdot 8 \cdot 7}{4 \cdot 3 \cdot 2 \cdot 1} = 210$

Calcola il valore dei seguenti coefficienti binomiali.

127 $\binom{8}{3}$ [56]

128 $\binom{10}{6}$ [210]

129 $\binom{12}{7}$ [792]

130 $\binom{15}{10}$ [3003]

131 $C_{8,2}$ [28]

132 $C_{9,3}$ [84]

133 $C_{12,5}$ [792]

134 **Vero o falso?**

a. $\binom{13}{0} = 0$ V F

b. $\binom{n}{k}$ è definito per ogni $n \in \mathbf{N}$ e per ogni $k \in \mathbf{N}$ V F

c. $\binom{8}{3} = \binom{8}{5}$ V F

d. $\binom{11}{10} = 11$ V F

e. $\binom{7}{3} = \dfrac{7!}{3!4!}$ V F

[3 affermazioni vere e 2 false]

A mente

Calcola il valore delle seguenti espressioni.

135 $\binom{100}{0}$; $\binom{100}{1}$

136 $\binom{10}{9}$; $\binom{157}{156}$

137 $\binom{3}{2}$; $\binom{10}{2}$

138 $\binom{11}{9}$; $\binom{21}{19}$

Calcola il valore delle seguenti espressioni.

139 $\dfrac{C_{6,1} \cdot C_{7,3} - C_{10,7}}{C_{6,3} + C_{8,4}}$ [1]

140 $\dfrac{(C_{4,2})^2 + C_{5,3} - C_{8,2}}{C_{6,2} - C_{4,2}}$ [2]

Metodi a confronto

141 Calcola il valore delle seguenti espressioni in due modi: prima per via diretta, poi riconoscendo che l'espressione **a** rappresenta il numero complessivo di sottoinsiemi di un insieme avente una cardinalità opportuna, mentre l'espressione **b** il numero complessivo di sottoinsiemi propri di un insieme avente una opportuna cardinalità.

a. $\binom{3}{0} + \binom{3}{1} + \binom{3}{2} + \binom{3}{3}$

b. $\binom{5}{1} + \binom{5}{2} + \binom{5}{3} + \binom{5}{4}$

[**a.** 8; **b.** 30]

142 Calcola il valore delle seguenti espressioni in due modi: prima per via diretta, poi riconoscendo che ciascuna di esse rappresenta il numero complessivo di sottoinsiemi aventi almeno 2 elementi di un insieme avente una opportuna cardinalità.

a. $\binom{5}{2} + \binom{5}{3} + \binom{5}{4} + \binom{5}{5}$

b. $\binom{6}{2} + \binom{6}{3} + \binom{6}{4} + \binom{6}{5} + \binom{6}{6}$

[**a.** 26; **b.** 57]

Verifica le seguenti identità.

143 $3\binom{n}{3} = n\binom{n-1}{2}$

144 $\binom{n}{2} + \binom{n}{3} = \binom{n+1}{3}$

145 $5\binom{n+1}{5} = (n+1)\binom{n}{4}$

146 $\binom{n-1}{4} + \binom{n-1}{5} = \binom{n}{5}$

147 **ESERCIZIO SVOLTO**

Risolviamo l'equazione $\binom{n}{2} = \binom{n}{4}$.

- Poniamo anzitutto le condizioni affinché l'equazione abbia significato: deve essere $n \in \mathbb{N}$ e $n \geq 4$, affinché siano definiti i due coefficienti binomiali.
- In base alla definizione di coefficiente binomiale l'equazione si traduce nella seguente:

$$\frac{n(n-1)}{2!} = \frac{n(n-1)(n-2)(n-3)}{4!}$$

Poiché deve essere $n \geq 4$, possiamo supporre $n(n-1) \neq 0$ e dividere entrambi i membri dell'equazione per $n(n-1)$; siamo condotti così all'equazione

$$\frac{1}{2!} = \frac{(n-2)(n-3)}{4!}$$

che fornisce come soluzioni:

$$n = -1 \vee n = 6$$

- Delle due soluzioni trovate, l'unica che soddisfa le condizioni iniziali e che risulta quindi accettabile è $n = 6$.

Risolvi le seguenti equazioni.

148 $\binom{n}{3} = 5\binom{n}{2}$ [17]

149 $\binom{n}{3} = 2\binom{n}{4}$ [5]

150 $\binom{n}{5} = \binom{n-1}{4}$ [5]

151 $5\binom{n}{3} = \binom{n+2}{3}$ [4]

152 $\binom{n+1}{4}+\binom{n}{4}=3n^2-6n$ [7]

153 $\binom{n}{3}-\binom{n-1}{3}=21$ [8]

154 $\binom{n}{3}+\binom{n}{n-3}=2\binom{n}{4}$ [7]

155 $\binom{n}{6}+\binom{n}{n-6}=2$ [6]

156 $\binom{n-1}{4}=\frac{n}{20}\binom{n}{5}$ [10]

157 $\binom{2n}{2}-\binom{n}{2}=n^2+15$ [6]

Problemi sulle combinazioni semplici

158 ESERCIZIO GUIDATO

Un gruppo di 8 dipendenti di un'azienda decide di mandare una delegazione di 3 di loro a esporre delle lamentele alla direzione.

a. **Quante delegazioni diverse sono possibili?**

b. **Supposto che debba fare parte della delegazione il rappresentante sindacale, che è uno degli 8 dipendenti, quante delegazioni sono possibili?**

a. Scegliere una delegazione equivale a scegliere un sottoinsieme di 3 elementi dall'insieme degli otto dipendenti. Le delegazioni possibili sono quindi: $\binom{8}{3}=\ldots$

b. La delegazione deve contenere, oltre al rappresentante sindacale, 2 dei restanti 7 dipendenti, quindi le possibilità sono in totale $\binom{7}{2}=\ldots$ [**a.** 56; **b.** 21]

159 Quanti sono i sottoinsiemi di 3 elementi dell'insieme $E=\{a,b,c,d,e\}$? [10]

160 Un professore decide di interrogare a caso 4 studenti in una classe di 20 studenti. Quanti diversi insiemi di studenti può interrogare? [4845]

161 In quanti modi diversi è possibile scegliere i due rappresentanti degli studenti in una classe di 24 alunni? [276]

162 A un'estrazione del lotto vengono estratti i numeri 5, 6, 8, 60, 74. Verifica che gli ambi vincenti (cioè gli ambi che si possono formare con i numeri estratti) sono tanti quanti i terni vincenti. [10]

163 Quanti diversi incontri di pugilato possono essere organizzati tra 6 pugili? [15]

164 Calcola il numero di strette di mano che possono scambiarsi 8 persone, nell'ipotesi che ciascuno stringa la mano una e una sola volta a tutti gli altri. [28]

165 Una scuola organizza dei corsi pomeridiani di approfondimento in italiano, inglese, matematica, elettronica, informatica e scienze. Se uno studente vuole seguire solo 3 corsi, in quanti modi può sceglierli? [20]

166 Quante coppie non ordinate (ambi) puoi formare con i 90 numeri del lotto? [4005]

167 Quante terne non ordinate (terni) puoi formare con i 90 numeri del lotto? [117 480]

168 Considera cinque punti distinti del piano A, B, C, D, E. Quanti segmenti distinti esistono aventi come estremi due dei cinque punti A, B, C, D, E? [10]

169 Quante sono le diagonali di un poligono avente 10 lati? [35]

170 Considera un pentagono $ABCDE$.

a. Quanti triangoli distinti esistono aventi come vertici tre dei vertici del pentagono?

b. Quanti quadrilateri distinti esistono aventi come vertici quattro dei vertici del pentagono? [**a.** 10; **b.** 5]

171 Dato un parallelepipedo, quanti triangoli possono essere formati aventi come vertici tre dei vertici del parallelepipedo? [56]

Metodi a confronto

172 Determina il numero di tutti i possibili anagrammi della parola «mamma»:

a. utilizzando le permutazioni con ripetizione

b. utilizzando le combinazioni.

(*Suggerimento*: osserva che un anagramma resta univocamente individuato una volta scelti i tre posti dove inserire la lettera «m» oppure i due posti dove inserire la lettera «a») [10]

173 Determina il numero di tutti i possibili anagrammi della parola «pappa»:

a. utilizzando le permutazioni con ripetizione;

b. utilizzando le combinazioni. [10]

174 Ho dieci libri che non ho ancor letto e voglio sceglierne almeno due, ma non più di tre, da portare in vacanza. Quante sono le scelte possibili? [165]

175 I 20 membri di uno sci club decidono di mandare una delegazione di 5 di loro a disputare una gara.

a. Quante delegazioni diverse sono possibili?

b. Supposto che debbano fare parte della delegazione il presidente e il vicepresidente dello sci club, che fanno parte dei 20 membri, quante delegazioni sono possibili? [**a.** 15 504; **b.** 816]

176 **ESERCIZIO GUIDATO**

I 21 studenti di una classe devono essere divisi in 3 gruppi di 7 studenti, ciascuno dei quali lavorerà indipendentemente a una ricerca. Supponendo le tre ricerche diverse, in quanti modi si possono costruire i tre gruppi?

Diciamo A, B, C le tre diverse ricerche che l'insegnante vuole assegnare.

- Osserva che:
  - il gruppo di studenti cui assegnare la ricerca A può essere scelto in $\binom{21}{\dots}$ modi;
  - il gruppo di studenti cui assegnare la ricerca B può essere scelto in $\binom{\dots}{7}$ modi;
  - il gruppo cui assegnare la ricerca C dovrà essere costituito dagli studenti rimanenti.
- Per il principio fondamentale del calcolo combinatorio, le scelte possibili sono in tutto:

$$\binom{21}{\dots} \cdot \binom{\dots}{7} = \dots\dots$$

[399 072 960]

177 Un equipaggio di un treno è costituito da due macchinisti, un capotreno e un bigliettaio. In una certa unità operativa sono in forza 50 macchinisti, 30 capotreno e 80 bigliettai. Quanti equipaggi diversi si possono formare con il personale di quel compartimento? [2 940 000]

178 Due sposi devono scegliere quattro testimoni per il loro matrimonio: due per lui e due per lei. La sposa può scegliere tra 12 amici, lo sposo tra 10. In quanti modi possibili i due sposi possono scegliere i quattro testimoni? [2970]

179 In una compagnia di alpini sono a disposizione 4 ufficiali, 8 sottoufficiali e 20 soldati. In quanti modi si può scegliere un plotone da mandare a una manifestazione se questo deve essere costituito da 1 ufficiale, 2 sottoufficiali e 10 soldati? [20 692 672]

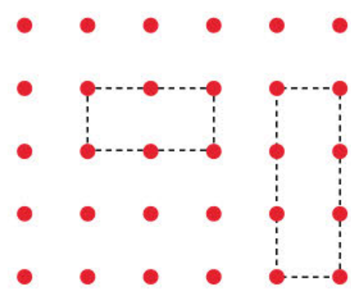
180 In una scuola vi sono 18 insegnanti di materie scientifiche e 30 insegnanti di materie letterarie. In quanti modi si può costituire una commissione di cinque professori, due di materie scientifiche e tre di materie letterarie? [621 180]

181 Un lotto di 15 telefoni cellulari ne contiene 3 difettosi. Viene scelto a caso un campione di 4 telefoni tra i 15 del lotto. Quanti dei possibili campioni contengono almeno un pezzo difettoso? [870]

182 Supponiamo di estrarre 5 carte (una mano) da un mazzo di 32 carte: quante mani contengono almeno tre fanti? [1540]

183 Quante coppie di angoli opposti al vertice sono formate da 8 rette distinte, tutte passanti per uno stesso punto? [56]

184 Considera tutti i possibili rettangoli aventi come vertici quattro dei punti colorati in rosso in figura, tali che i lati dei rettangoli siano orizzontali o verticali (due di tali rettangoli sono rappresentati come esempio). Quanti sono complessivamente tali rettangoli? [150]



185 Ciascuno dei 30 supermercati di una catena ha 2 rappresentanti sindacali. Quante delegazioni di tre rappresentanti sindacali si possono formare, con il vincolo che a una delegazione non possano appartenere due rappresentanti dello stesso supermercato? [32 480]

Problemi sulle combinazioni con ripetizione

186 **ESERCIZIO GUIDATO**

a. **In quanti modi si possono assegnare 15 scrivanie uguali a 4 uffici (ammettendo anche il caso che a qualche ufficio non venga assegnata alcuna scrivania)?**

b. **Come cambia la risposta al problema precedente se a ogni ufficio deve essere assegnata almeno una scrivania?**

a. Si tratta di determinare i numeri x_1, x_2, x_3, x_4 di scrivanie che vengono assegnate a ciascun ufficio, ossia di determinare quante quaterne (x_1, x_2, x_3, x_4) di interi non negativi soddisfano l'equazione:

$$x_1 + x_2 + x_3 + x_4 = 15$$

La suddivisione delle scrivanie può quindi essere fatta in $\binom{18}{15}$ modi.

b. Poiché a ogni ufficio deve essere assegnata almeno una scrivania, si tratta ora di suddividere 11 scrivanie tra 4 uffici.
[**a.** 816; **b.** 364]

187 **E se?** Quanti possibili tipi di confezioni diverse di 20 caramelle ai gusti di menta, fragola o limone si possono confezionare, ammettendo il caso di confezioni costituite da caramelle di un solo gusto o di due gusti?
▸ Come cambierebbe la risposta considerando solo le confezioni contenenti caramelle di tutti e tre i gusti? [231; 171]

188 Si lanciano contemporaneamente 5 dadi. Quante sono le possibili combinazioni di numeri che si possono ottenere? [252]

189 **E se?** Determina il numero di terne (x_1, x_2, x_3), con $x_1, x_2, x_3 \in \mathbf{N}$, che soddisfano l'equazione $x_1 + x_2 + x_3 = 12$. [91]
▸ Che cosa accadrebbe se considerassimo le terne (x_1, x_2, x_3) in cui $x_1, x_2, x_3 \in \mathbf{Z}$?

190 Venti lavagne interattive uguali devono essere suddivise tra 10 scuole. In quanti modi può avvenire la suddivisione (ammettendo anche il caso che a qualche scuola non venga assegnata alcuna lavagna)? E se a ogni scuola deve essere assegnata almeno una lavagna? [10 015 005; 92 378]

191 Il signor Bianchi ha a disposizione la somma di 40 000 euro, che vuole investire, in tranche da 5000 euro, scegliendo tra quattro società. In quanti modi diversi possono essere investiti i soldi del sig. Bianchi, se almeno una tranche deve essere investita in ciascuna delle quattro società?
(*Suggerimento*: si tratta di distribuire in sostanza 8 oggetti identici (le 8 tranche da 5000 euro) tra le 4 possibili società, con il vincolo che almeno 1 tranche deve andare a ciascuna società) [35]

Collegamenti Calcolo combinatorio e algebra

192 ESERCIZIO GUIDATO

Considera il polinomio $(a + b + c)^4$ e determina:
a. quanti termini contiene il suo sviluppo ridotto;
b. qual è il coefficiente del monomio simile ad a^2bc nel suo sviluppo ridotto.

---

a. Lo sviluppo ridotto di $(a + b + c)^4$ è un polinomio omogeneo, i cui termini hanno grado 4. Pertanto ogni termine del polinomio ridotto è simile a un monomio della forma $a^{x_1}b^{x_2}c^{x_3}$ dove x_1, x_2, x_3 sono interi non negativi tali che:

$$x_1 + x_2 + x_3 = 4 \qquad [*]$$

Viceversa, a ogni terna di interi non negativi che risolve l'equazione [*] corrisponde uno e un solo termine del polinomio ridotto. Il problema equivale quindi a determinare il numero di terne di interi non negativi che soddisfano la [*]. In base al **Teorema 6**, tali terne sono in totale $C^*_{3,4}$.

b. Osserva che i monomi *aabc*, *abac*, *cbaa*, ... (che si otterrebbero effettuando la moltiplicazione di quattro fattori uguali ad $a + b + c$) sono simili e quindi nello sviluppo ridotto saranno rappresentati da un *unico* monomio, la loro somma. Il coefficiente di a^2bc, insomma, è quindi uguale al numero complessivo di *tutti* questi monomi simili, cioè al numero di tutti i possibili anagrammi della «parola» *aabc*. [**a.** 15; **b.** 12]

193 Quanti termini può contenere al massimo un polinomio omogeneo di grado 8, nelle variabili *a*, *b* e *c*? [45]

194 Supponiamo di sviluppare la potenza $(a + b + c + d)^7$ e di ridurre i termini simili. Quanti termini contiene il polinomio così ottenuto? [120]

195 Qual è il coefficiente di AB^2C^2 nello sviluppo ridotto di $(A + B + C)^5$? [30]

196 Qual è il coefficiente di x^3y^2 nello sviluppo ridotto di $(2x - y + 3)^7$?
(*Suggerimento*: poni $2x = X$, $-y = Y$ e $3 = Z$ e determina inizialmente il coefficiente di $X^3Y^2Z^2$ nello sviluppo ridotto di $(X + Y + Z)^7$) [15 120]

Esercizi riassuntivi: combinazioni semplici o con ripetizioni

197 Un'associazione ha 20 membri. Se tra essi deve essere formato un comitato di 5 persone, quanti comitati diversi sono possibili? [15 504]

198 Quante partite diverse possono disputare 22 ragazzi, suddividendosi in due squadre da 11 e supponendo che ogni coppia di squadre giochi una sola partita? [352 716]

199 Dei 40 dipendenti di una piccola azienda, 20 sono laureati, 10 sono diplomati e 10 hanno la licenza di scuola media. Per un'indagine interna si vuole utilizzare un campione in cui sia rappresentato casualmente il 10% di ognuno dei gruppi individuato dal titolo di studio. Quanti sono i possibili campioni distinti? [19 000]

200 Quante commissioni d'esame formate da 3 professori si possono formare scegliendo i professori da un insieme di cinque insegnanti? [10]

201 Si hanno vernici di colore rosso, blu e verde. Mescolando 3 parti uguali di tali vernici, con la possibilità di ripetere eventualmente uno o più colori, quante diverse sfumature di colore si possono ottenere? [10]

202 Quanti sono i possibili risultati che si potrebbero ottenere in una estrazione del Lotto, se l'estrazione delle cinque palline avvenisse con reimmissione? [54 891 018]

---

203 **Zoom sull'enunciato** Lanciamo *contemporaneamente* 4 dadi; quante sono le possibili combinazioni di numeri che si possono ottenere? Cambierebbe la risposta se, invece di lanciare contemporaneamente 4 dadi, effettuassimo 4 lanci *successivi* di uno stesso dado? [126; 1296]

Guida alla comprensione del testo
L'espressione «contemporaneamente» indica che non è possibile distinguere un ordine tra i numeri ottenuti; se invece si effettuano quattro lanci «successivi» è definito un preciso ordine nei risultati, che va tenuto in considerazione.

204 Quanti sono i monomi di grado 8, nelle tre variabili x, y e z? [45]

205 Per stabilire il migliore a briscola in quattro, sei amici al bar intendono giocare tante partite quanti sono i possibili accoppiamenti tra di loro (due coppie si sfidano, una riposa): chi vincerà più partite sarà riconosciuto come il migliore del gruppo. Quante partite giocheranno in tutto, i sei? E quante partite ciascuno? [45; 30]

206 Una sede produttiva di una casa automobilistica ha nella propria catena produttiva 5 tipi diversi di automobili. In quanti modi diversi può giungere alla sede un ordine di 50 automobili? [316 251]

207 Si deve formare un comitato costituito da 2 uomini e 3 donne, scegliendone i componenti da un gruppo di 6 uomini e 5 donne.

a. In quanti modi si può formare il comitato?

b. In quanti modi si può formare il comitato, se tra le cinque donne ce ne sono due che hanno litigato e perciò non vogliono appartenere al comitato insieme? [**a.** 150; **b.**105]

208 Cinque amici devono ripartirsi 10 caramelle uguali.

a. In quanti modi possono farlo (ammettendo anche il caso in cui qualcuno non riceva nessuna caramella)?

b. In quanti modi possono farlo se ciascuno deve ricevere almeno una caramella? [**a.** 1001; **b.** 126]

209 **Caccia all'errore.** Barbara deve risolvere il problema seguente: «Supponiamo di estrarre 5 carte (una mano) da un mazzo di 32 carte: quante mani contengono almeno un asso?». Ragiona così:

«Posso costruire una mano contenente almeno un asso effettuando due scelte successive:

- prima scelgo uno dei quattro assi: ho 4 possibilità;
- poi scelgo 4 carte tra le 31 restanti: ho $\binom{31}{4}$ possibilità.

Il numero totale di mani che contengono almeno un asso è allora:

$$4 \cdot \binom{31}{4} = 125\,860$$

in forza del principio fondamentale del calcolo combinatorio»

Individua l'errore che ha commesso Barbara nel suo ragionamento e determina la risposta corretta del problema.

[103 096]

210 **Inventa tu.** Scrivi il testo di un problema che può essere risolto calcolando $\binom{15}{5} \cdot \binom{5}{3}$.

211 **Inventa tu.** Scrivi il testo di un problema che può essere risolto calcolando $\binom{10}{3} \cdot \binom{7}{2}$.

4. Il teorema del binomio di Newton

Teoria p. 111

212 ESERCIZIO SVOLTO

Sviluppiamo la potenza $(2x + y)^4$.

Utilizziamo la formula del binomio di Newton relativa ad $(a + b)^4$, sostituendo $2x$ al posto di a e y al posto di b:

$$(2x + y)^4 = \binom{4}{0}(2x)^4 + \binom{4}{1}(2x)^3 y + \binom{4}{2}(2x)^2(y)^2 + \binom{4}{3}(2x)y^3 + \binom{4}{4}y^4 =$$

$$= 16x^4 + 32x^3y + 24x^2y^2 + 8xy^3 + y^4$$

Sviluppa le seguenti potenze di binomi.

213 $(2x - 1)^4$

214 $(x + 2)^5$

215 $\left(\frac{1}{2}x^2 - y^3\right)^4$

216 $(x - 3y)^5$

217 ESERCIZIO SVOLTO

Calcoliamo il coefficiente di x^8y^2 nello sviluppo di $(x - 2y)^{10}$.

- Per la formula del binomio di Newton:

$$(x - 2y)^{10} = \sum_{k=0}^{10} \binom{10}{k} x^{10-k}(-2y)^k =$$

$$= \sum_{k=0}^{10} \binom{10}{k} (-2)^k x^{10-k} y^k$$

- Confrontando x^8y^2 con $x^{10-k}y^k$, ci accorgiamo che siamo interessati a calcolare il coefficiente del termine corrispondente a $k = 2$. Il coefficiente di tale termine è allora:

$$\binom{10}{2}(-2)^2 = 180$$

218 Calcola il coefficiente di a^6b^2 nello sviluppo di $(a+b)^8$. [28]

219 Calcola il coefficiente di a^7b^3 nello sviluppo di $(a+b)^{10}$. [120]

220 Calcola il coefficiente di x^6 nello sviluppo di $(x+3)^8$. [252]

221 Calcola il coefficiente di x^3 nello sviluppo di $(x-1)^{10}$. [-120]

222 Calcola il coefficiente di x^4y nello sviluppo di $(2x-3y)^5$. [-240]

223 Calcola il coefficiente di x^4y^2 nello sviluppo di $(2x-3y)^6$. [2160]

224 Calcola il coefficiente di x^3y^2 nello sviluppo di $(3x+4y)^5$. [4320]

225 Utilizzando lo sviluppo di $(a+b)^5$ e sostituendo in esso opportuni valori di a e b, verificare che:

$$\binom{5}{0}+\binom{5}{1}+\binom{5}{2}+\binom{5}{3}+\binom{5}{4}+\binom{5}{5}=2^5$$

226 Utilizzando lo sviluppo di $(a+b)^6$ e sostituendo in esso opportuni valori di a e b, verificare che:

$$\binom{6}{0}-\binom{6}{1}+\binom{6}{2}-\binom{6}{3}+\binom{6}{4}-\binom{6}{5}+\binom{6}{6}=0$$

227 Utilizzando il teorema del binomio di Newton, dimostra che:

$$\binom{n}{0}+\binom{n}{1}+\ldots+\binom{n}{n}=2^n.$$

228 Utilizzando il teorema del binomio di Newton, dimostra che:

$$\binom{n}{0}-\binom{n}{1}+\binom{n}{2}-\ldots+(-1)^n\binom{n}{n}=0.$$

Determina il valore delle seguenti somme.

229 $\sum_{k=0}^{n}\binom{n}{k}x^k(1-x)^{n-k}$ [1]

230 $\sum_{k=0}^{n}\binom{n}{k}x^{k+3}(1-x)^{n-k}$ [x^3]

231 $\sum_{k=0}^{n}(-1)^k\binom{n}{k}$ [0]

232 $\sum_{k=0}^{n}\binom{n}{k}3^k$ [4^n]

233 $\sum_{k=0}^{n}\binom{n}{k}5^n$ [10^n]

234 $\sum_{k=0}^{n}\binom{n}{k}5^k$ [6^n]

Esercizi di riepilogo

Esercizi interattivi

235 Vero o falso?

a. $D_{10,3} > D^*_{8,3}$ V F

b. $C_{20,4} > P_7$ V F

c. se $D_{n,2} = C_{n,3}$, allora $n = 8$ V F

d. con le lettere della parola "invincibili" si possono formare 166 320 anagrammi V F

e. nel gioco del poker, il numero di modi in cui si può servire a un giocatore cinque carte dello stesso seme, estratte da un mazzo di 52, è 1287 V F

f. il quarto termine dello sviluppo della potenza $\left(a^2 - \frac{1}{a^3}\right)^6$ è $\frac{15}{a^8}$ V F

Test

236 Nel magazzino di un'azienda, per contrassegnare i prodotti, viene utilizzato un codice formato, in sequenza, da quattro caratteri alfabetici, tre numerici (da 0 a 9) e due alfabetici. Quale calcolo è necessario eseguire per trovare il numero di prodotti diversi che si possono così contrassegnare? Supponi che i caratteri possano essere ripetuti.

A $D_{26,4} \cdot D_{10,3} \cdot D_{26,2}$
C $D^*_{26,4} \cdot D^*_{10,3} \cdot D^*_{26,2}$
B $C_{26,4} \cdot C_{10,3} \cdot C_{26,2}$
D $C_{26,4} \cdot D_{10,3} \cdot C_{26,2}$

237 Dello sviluppo della potenza $\left(2x^2 - \frac{3}{x}\right)^n$ sappiamo che il nono termine è un numero puro. Quanto vale questo termine?

A 7920 B $7920 \cdot 3^8$ C $7920 \cdot 2^3$ D 9!

238 Un fioraio ha a disposizione rose gialle, arancione e rosse. In quanti modi diversi può confezionare un mazzo con 13 rose (supponendo indistinguibili le rose dello stesso colore)?

A $\frac{13!}{3!}$ B 13! C 105 D $D^*_{13,3}$

239 Quale delle seguenti uguaglianze è vera?

A $\binom{n+1}{k} - \binom{n}{k-1} = \binom{n}{k}$

B $(n+1)\binom{n}{k} - \binom{n-1}{k-1} = \binom{n-1}{k}(n+1)$

C $\binom{n+1}{k} = n\binom{n}{k-1} + \binom{n}{k}$

D $\binom{n}{k} - \binom{n-1}{k-1} = n\binom{n-1}{k}$

240 Un tavolo rettangolare ha 10 posti, 5 su un lato e 5 sul lato opposto. In quanti modi possibili vi si possono sedere 3 uomini e 3 donne in modo tale che le donne stiano su di un lato, gli uomini sull'altro e che di fronte a ogni donna ci sia sempre un uomo?

A $D_{5,3} \cdot P_3$
C $2 \cdot C_{5,3} \cdot P_3$
B $C_{5,3} \cdot P_3$
D $2 \cdot D_{5,3} \cdot P_3$

241 Una scatola contiene 8 palline numerate. Se ne estraggono 3 e si vuole sapere quante sono le possibili configurazioni a seconda che: l'estrazione sia simultanea, l'estrazione avvenga in sequenza con reimmissione, l'estrazione avvenga in sequenza senza reimmisione.

A $D_{8,3}$; $D^*_{8,3}$; $C_{8,3}$

B $C_{8,3}$; $D^*_{8,3}$; $D_{8,3}$

C $C_{8,3}$; $D_{8,3}$; $D^*_{8,3}$

D $C_{8,3}$; $C^*_{8,3}$; $D_{8,3}$

242 A una corsa podistica partecipano 500 persone. In quanti modi sarà possibile in teoria redigere la classifica dei primi 20 (escludendo ex-aequo)?

A $\frac{500!}{20! \cdot 480!}$

B $500 \cdot 499 \cdot 498 \cdot \ldots \cdot 481 \cdot 480$

C $500 \cdot 499 \cdot 498 \cdot \ldots \cdot 481$

D $\frac{500!}{20!}$

243 Facendo riferimento all'esercizio precedente e supponendo che i partecipanti alla gara siano 180 donne e 320 uomini, in quanti modi possibili la classifica dei primi 20 potrebbe vedere occupate da uomini le posizioni dispari e da donne quelle pari?

A $\frac{320!}{10!} \cdot \frac{180!}{10!}$

B $\frac{320!}{310!} \cdot \frac{180!}{170!}$

C $\binom{320}{10} \cdot \binom{180}{10}$

D $500 \cdot 499 \cdot 498 \cdot \ldots \cdot 481$

244 Nel gioco del poker ogni giocatore riceve 5 carte, estratte da un mazzo di 32. In quanti modi diversi il giocatore può ricevere le carte? [201 376]

245 Quanti sono i numeri di cinque cifre che si possono costruire scegliendo le cifre nell'insieme $\{0, 1, 2, 5, 6, 7\}$ e ammettendo di poter ripetere le cifre? [6480]

246 Sono di più i sottoinsiemi di 2 elementi in un insieme di 10 elementi oppure i sottoinsiemi di 4 elementi in un insieme di 8 elementi? [I sottoinsiemi di 4 elementi di un insieme di 8]

247 Ci sono dieci ambasciatori senza sede e tre posti da coprire: Città del Vaticano, Parigi e Vienna. In quanti modi possibili i tre ambasciatori possono essere assegnati? [720]

248 Un professore di storia non chiede le date degli eventi, ma è molto severo nella richiesta dell'ordine cronologico: per ogni domanda elenca cinque fatti e chiede che essi siano ordinati cronologicamente. Quante sono le risposte possibili per ogni domanda? [120]

249 Quanti numeri di 6 cifre è possibile costruire, aventi cifre tutte diverse da zero e multiple di 3? [729]

250 Un magazzino di una casa editrice ha in giacenza 10 titoli di libri (il numero di copie di ciascuno è più che sufficiente a far fronte a qualunque richiesta). In quanti modi possibili quel magazzino può ricevere un ordine di 15 volumi? [1 307 504]

251 Quattro amici partono per un viaggio con un'automobile a quattro posti. Solo tre dei quattro amici hanno la patente. In quanti modi diversi possono disporsi i quattro amici all'interno dell'auto? [18]

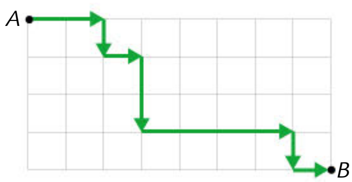
252 Quanti sono i numeri di 3 cifre che si possono scrivere senza mai utilizzare lo zero? Fra di essi, sono in numero maggiore quelli in cui compare almeno 1 volta la cifra 1 o quelli in cui non compare? [9^3; i numeri in cui compare la cifra 1 almeno una volta sono 217 e sono minori di quelli in cui non compare, che sono 512]

253 Quanti sono gli anagrammi della parola «anagramma»? [7560]

254 Sette persone hanno a disposizione cinque sedie numerate da 1 a 5: cinque persone si siedono e due restano in piedi. In quanti modi diversi possono occupare le cinque sedie? [2520]

255 Una classe è formata da 20 alunni. In quanti modi la classe può essere suddivisa in due gruppi ugualmente numerosi (considerando irrilevante l'ordine dei due gruppi)? [92 378]

256 In riferimento alla figura, calcola il numero dei cammini lungo la quadrettatura che conducono dal punto A al punto B, supponendo che sia possibile muoversi soltanto verso destra o verso il basso. Un possibile cammino è, per esempio, quello indicato in verde in figura.
(*Suggerimento*: ci si può ricondurre al calcolo di una permutazione con ripetizioni) [495]



Realtà e modelli

257 **Esame di stato.** La commissione d'esame di stato (o maturità) è costituita da 7 membri: tre docenti interni (cioè in servizio presso lo stesso istituto dei candidati), tre docenti esterni e un presidente. Nel corso delle prove orali, i sette commissari si dispongono in fila lungo il lato di un tavolo, riservando l'altro allo studente esaminato.

a. In quanti modi possono accomodarsi i sette membri della commissione?
b. In quanti modi, se i commissari interni vogliono sedere affiancati e così pure i commissari esterni?
c. In quanti modi, se il presidente intende occupare la posizione centrale della fila, mentre gli altri sei possono accomodarsi liberamente?
d. In quanti modi, se il presidente vuole sedere al centro della fila, i membri interni affiancati e così pure i membri esterni? [**a.** 5040; **b.** 216; **c.** 720; **d.** 72]

258 **E se?** La squadra di calcio di Pietro conta esattamente 2 portieri, 8 difensori, 7 centrocampisti e 5 attaccanti. Per una questione di equità, l'allenatore seleziona i giocatori titolari casualmente, ma in coerenza con lo schema di gioco: per esempio, scegliendo lo schema 4-5-1 giocano 4 difensori, 5 centrocampisti e 1 attaccante, oltre a 1 portiere. Calcola tutte le possibili formazioni secondo lo schema 4-3-3. Occorre distinguere i ruoli all'interno dello stesso reparto (per esempio, terzino sinistro e terzino destro vanno considerati ruoli diversi, anche se appartenenti allo stesso reparto difensivo).

▶ Come cambierebbe la risposta considerando tutti i possibili schieramenti secondo lo schema 5-3-2? [42 336 000; 56 448 000]

259 Nel gioco del poker ogni giocatore riceve 5 carte (una «mano»), estratte da un mazzo di 32.

a. Quante mani contengono esattamente tre fanti?

b. Quante mani contengono 2 donne e 3 re? [**a.** 1512; **b.** 24]

260 Una password è costituita da sei caratteri, ciascuno dei quali può essere una delle 21 lettere dell'alfabeto italiano.

a. Quante password diverse si possono costruire, formate da lettere tutte distinte tra loro?

b. Quante password diverse si possono costruire, ammettendo di poter ripetere le lettere?

c. Quante password si possono costruire, che iniziano con una consonante, terminano con una vocale e sono formate da lettere tutte distinte tra loro? [**a.** 39 070 080; **b.** 21^6; **c.** 7 441 920]

261 L'investigatore Colombo, per concludere un'indagine su un omicidio, deve scoprire un certo numero di telefono. Due testimoni hanno sentito il numero, ma non concordano su di esso. Sono però entrambi d'accordo su quanto segue:

a. il numero è composto da cinque cifre e termina con la cifra 0;

b. la seconda cifra è dispari;

c. tutte le cifre sono diverse;

d. la cifra più grande è 6.

Quanti numeri telefonici deve controllare l'investigatore Colombo? [108]

262 Una cassetta contiene 20 mele, di cui 5 sono marce.

a. Quanti campioni diversi di 3 mele possono essere presi dalla cassetta?

b. Quanti campioni diversi di 3 mele possono essere presi, in cui tutte e tre le mele siano marce?

c. Quanti campioni diversi di 3 mele possono essere presi, in cui 2 mele siano sane e 1 sia marcia? [**a.** 1140; **b.** 10; **c.** 525]

263 Anna ha dieci amici.

a. In quanti modi può invitarne 5 a pranzo?

b. In quanti modi può invitarne 5 a pranzo, se due dei dieci amici sono sposati e partecipano alla cena solo insieme?

c. In quanti modi può invitarne 5 a pranzo, se due di essi hanno litigato e partecipano solo separatamente?

[**a.** 252; **b.** 112; **c.** 196]

264 Sia A un sottoinsieme di un insieme X. Supponiamo che $|A| = k$ e $|X| = n$, con $k \leq n$. Quanti sono i sottoinsiemi di X che contengono A? [2^{n-k}]

265 In una classe di 24 studenti, di cui 10 femmine e 14 maschi, si deve formare un gruppo per una ricerca costituito da 3 maschi e 3 femmine. In quanti modi può essere costituito il gruppo se tra i 14 maschi ci sono due gemelli e si decide che non possano stare insieme? [42 240]

266 In quanti modi è possibile suddividere 12 penne in 6 cassetti, ammettendo che le 12 penne siano indistinguibili e che qualche cassetto possa restare vuoto? [6188]

Realtà e modelli

267 **Comitato direttivo.** In un gruppo di 26 persone, 10 donne e 16 uomini, deve essere scelto un comitato direttivo costituito da un presidente, un vicepresidente e un segretario. Uno dei 16 uomini è il sig. Bianchi.

a. In quanti modi diversi è possibile scegliere il comitato?

b. In quanti modi diversi è possibile scegliere il comitato, se il posto di segretario deve essere occupato da una donna?

c. In quanti modi diversi è possibile scegliere il comitato, se si decide di assegnare il posto di presidente al sig. Bianchi?

d. In quanti modi diversi è possibile scegliere il comitato, se si decide di assegnare il posto di presidente a un uomo e il posto di vicepresidente a una donna?

e. In quanti modi diversi è possibile scegliere il comitato, se si vuole che presidente e vicepresidente siano di sesso diverso? [**a.** 15 600; **b.** 6000; **c.** 600; **d.** 3840; **e.** 7680]

268 **Giuria.** Si vuole formare una giuria composta da 6 persone. I membri della giuria sono da scegliere tra 6 uomini e 7 donne. Una delle 7 donne è la signora Verdi.

a. In quanti modi diversi è possibile comporre la giuria?

b. In quanti modi diversi è possibile comporre la giuria, se essa deve essere composta da 3 uomini e 3 donne?

c. In quanti modi diversi è possibile comporre la giuria, se essa deve essere composta da 3 uomini e 3 donne e tra le tre donne deve necessariamente essere presente la signora Verdi? [**a.** 1716; **b.** 700; **c.** 300]

269 Un'urna contiene 10 palline: 5 nere numerate da 1 a 5, 3 rosse numerate da 6 a 8, e 2 bianche numerate 9 e 10. Supponendo di estrarre dall'urna cinque palline, successivamente e rimettendo nell'urna l'ultima pallina estratta prima dell'estrazione successiva, determina:

a. in quanti modi è possibile estrarre le cinque palline;

b. in quanti modi è possibile estrarre cinque palline, di cui esattamente 2 rosse.

Rispondi poi nuovamente alle domande **a** e **b**, sia nel caso in cui le cinque palline siano estratte successivamente ma senza reimmissione sia nel caso in cui siano estratte contemporaneamente. [**a.** 10^5; **b.** 30 870; nel caso di estrazioni successive senza reimmissione le risposte ad **a.** e **b.** diventano rispettivamente 30 240 e 12 600; nel caso di estrazione contemporanea, le risposte ad **a.** e **b.** diventano rispettivamente 252 e 105]

---

270 Si estraggono contemporaneamente 5 carte da un mazzo di 32: un mazzo di questo tipo è costituito da 8 carte per ciascuno dei quattro semi (cuori, quadri, picche e fiori): 7, 8, 9, 10, fante, donna, re, asso. In quanti modi diversi è possibile estrarre 5 carte contenenti:

a. nessun asso;

b. esattamente 2 donne;

c. almeno 3 fanti;

d. 2 carte di picche e 3 di cuori;

e. 2 carte di un colore e 3 di un altro;

f. almeno un fante;

g. esattamente 3 carte di cuori ed esattamente 2 re.

[**a.** 98 280; **b.** 19 656; **c.** 1540; **d.** 1568; **e.** 134 400; **f.** 103 096; **g.** 1428]

271 Un'urna contiene 10 palline: tre bianche, numerate da 1 a 3 e sette nere, numerate da 4 a 10.
Si estraggono successivamente senza reimmissione 4 palline. In quanti modi diversi è possibile estrarre:

a. 4 palline nere;

b. 3 palline nere e 1 bianca, in quest'ordine;

c. 3 palline nere e 1 bianca, in ordine qualsiasi;

d. 2 palline bianche e 2 palline nere, in ordine qualsiasi;

e. almeno 3 palline nere;

f. al massimo 3 palline nere.

[**a.** 840; **b.** 630; **c.** 2520; **d.** 1512; **e.** 3360; **f.** 4200]

272 Un'associazione di n persone deve eleggere un comitato direttivo di k persone (k fissato, con $k \leq n$), una delle quali sarà il presidente dell'associazione. In quanti modi diversi l'associazione può eleggere il comitato direttivo e il presidente? Rispondi a questa domanda determinando nei due modi diversi seguenti il numero delle possibili scelte dell'associazione:

a. supponendo che l'associazione scelga prima il comitato direttivo e poi il comitato direttivo elegga, all'interno di quest'ultimo, il presidente;

b. supponendo che l'associazione elegga prima il presidente, poi il presidente scelga gli altri membri del comitato direttivo.

Quale identità puoi dedurre dal confronto dei risultati ottenuti in **a** e in **b**? Dimostra questa identità algebricamente. $\left[\textbf{a. } k\binom{n}{k}; \textbf{ b. } n\binom{n-1}{k-1}\right]$

273 Un'associazione di n persone deve eleggere un comitato direttivo formato da un numero qualunque di persone e un presidente del comitato. In quanti modi diversi l'associazione può eleggere il comitato direttivo e il presidente? Rispondi a questa domanda determinando in due modi diversi il numero delle possibili scelte dell'associazione:

a. supponendo che l'associazione scelga prima il comitato direttivo e poi il comitato direttivo elegga, all'interno di quest'ultimo, il presidente;

b. supponendo che l'associazione elegga prima il presidente, poi il presidente scelga gli altri membri del comitato direttivo.

Quale identità puoi dedurre dal confronto dei risultati ottenuti in **a** e in **b**? $\left[\textbf{a. } \sum_{k=1}^{n} k\binom{n}{k}; \textbf{ b. } n \cdot 2^{n-1}\right]$

Giustificare e argomentare

Per ciascuna delle seguenti uguaglianze stabilisci se è un'identità; in caso affermativo, dimostrala, altrimenti fornisci un controesempio.

274 $(2n)!(3n)! = (6n^2)!$

275 $3\binom{n}{3} = n\binom{n-1}{2}$

276 $\dfrac{(4n)!}{n!} = 4!$

277 $\binom{2n}{2} = 2\binom{n}{2} + n^2$

278 $\binom{n}{3}\binom{5}{2} = \binom{5n}{6}$

279 $\binom{n}{5} = \binom{n-1}{4} + \binom{n-1}{5}$

Esercizi *più*

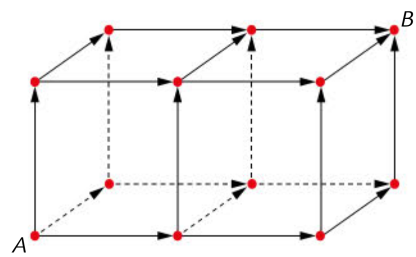
280 Vogliamo disporre 5 pennarelli di colori diversi (giallo, verde, rosso, blu, viola) in quattro cassetti vuoti numerati da 1 a 4.

a. In quanti modi diversi è possibile farlo (ammettendo anche i casi in cui uno o più cassetti rimangano vuoti)?
b. In quanti modi diversi è possibile farlo, se si vuole lasciare il quarto cassetto vuoto?
c. In quanti modi diversi è possibile farlo, se si vogliono lasciare vuoti sia il terzo che il quarto cassetto?
d. In quanti modi diversi è possibile farlo, se si vuole che il terzo o il quarto cassetto non siano vuoti?
e. In quanti modi diversi è possibile farlo, se si vuole che il terzo e il quarto cassetto non siano vuoti?
f. In quanti modi diversi è possibile farlo, se si vuole che nessun cassetto resti vuoto?

[**a.** $4^5 = 1024$; **b.** $3^5 = 243$; **c.** $2^5 = 32$; **d.** 992; **e.** 570; **f.** 240]

Dalle gare

281 Osserva la figura. Quanti diversi cammini consentono di andare da A a B muovendosi lungo gli spigoli e rispettandone il verso indicato?



A 6   B 8   C 9   D 10   E 12

(Kangourou 2013) [E]

282 Dividiamo il numero $(1! + 2! + 3! + \dots + 100!)^2$ per 5: qual è il resto?

A 0   B 1   C 2   D 3   E 4

(Kangourou 2013) [E]

283 Carla si è dimenticata la password di accesso al suo nuovissimo computer! Si ricorda però che è una sequenza di 4 vocali, non necessariamente distinte, di cui due sono maiuscole e due sono minuscole. Quante passwords diverse deve provare Carla, al massimo, per accedere al suo computer?

A $3 \cdot 5^4$ B 5^5 C $6 \cdot 5^4$ D 5^6 E $3 \cdot 5^6$

(Giochi di Archimede 2009) [C]

284 Dieci amici decidono di giocare una partita di calcetto, cinque contro cinque. Sapendo che vi sono due terne di fratelli e che i tre fratelli Ambrosio desiderano giocare tutti nella squadra A mentre i tre fratelli Bianchi desiderano giocare tutti nella squadra B, in quanti modi si possono formare le due squadre?

A 3   B 6   C 15   D 24   E 30

(Giochi di Archimede 2004) [B]

285 Un ladro spia Marco mentre chiude la sua valigia con un lucchetto con una combinazione di 3 cifre (ciascuna cifra va da 0 a 9). Non ha potuto vedere la combinazione, ma è riuscito a capire che due cifre consecutive sono uguali e la terza è diversa. Qual è il massimo numero di combinazioni che il ladro dovrà provare per aprire la valigia di Marco?

A 180   B 190   C 200   D 210   E 220

(Giochi di Archimede 2000) [A]

286 Quante parole di quattro lettere (anche prive di senso compiuto) si possono scrivere utilizzando solo le lettere A, B, E, M, O (ammettendo che le lettere possano essere ripetute) in modo che nessuna delle lettere successive a una B (andando da sinistra verso destra) sia una M? (Quindi, per esempio, *ABEB* deve essere contata, ma *OBAM* no)

A $4^3 \cdot 5$ B $4^2 \cdot 5^2$ C $4 \cdot 5^3$ D 2^9 E 5^4

(Giochi di Archimede 2005) [D]

287 In un torneo di tennis, 8 persone decidono di giocare degli incontri di doppio (cioè due contro due) in tutti i modi possibili. Quanti incontri ci sono nell'intero torneo?

A 1680   B 126   C 1260   D 210   E 64

(Olimpiadi della Matematica, gara senior 1993) [D]

Calcolo combinatorio

1 **Vero o falso?**

a. $\binom{n}{k} = \binom{n}{n-k}$ per ogni n e k numeri naturali con $1 \leq k \leq n$ V F

b. $D^*_{7,6} = D^*_{6,7}$ V F

c. $(n+1)! - n! = n \cdot n!$ per ogni $n \in \mathbf{N}$ V F

d. L'insieme $\{1, 2, 3, 4, 5\}$ ha 10 sottoinsiemi di cardinalità 2 V F

2 Risolvi l'equazione $3\binom{n+2}{n} = \binom{n+3}{n+1}$.

3 L'insegnante di matematica vuole interrogare tre dei 24 studenti della classe di Paolo.

a. Quanti sono i possibili gruppi?

b. E quanti, se l'insegnante ha annunciato che uno dei «prescelti» sarà proprio Paolo?

4 Cinque soci del circolo di tennis intendono eleggere il migliore giocatore di doppio tra di loro. Convengono che la maniera più equa per farlo è disputare tante partite quanti sono tutti i modi possibili di accoppiare 4 dei 5 amici (le prime 2 coppie scenderanno in campo mentre il quinto riposerà). Quante partite in tutto verranno disputate? E quante giocate da ciascun tennista?

5 Martina deve svolgere un test costituito da 8 domande a risposta multipla: A, B, C o D.

a. Calcola in quanti modi possibili Martina potrebbe compilare il test.

b. Come cambierebbe la risposta, sapendo in anticipo che le risposte giuste sono equamente divise: due «A», due «B» e così via?

6 L'allenatore di calcio di Nicola ha le idee piuttosto confuse sullo schema di gioco da adottare: il solo punto certo è che, portiere escluso, ciascuno dei tre reparti (difesa, centrocampo, attacco) deve essere occupato da almeno due dei 10 giocatori in campo. Per esempio: 4-4-2, 3-5-2 ecc.
Quanti schemi di gioco sono possibili per la squadra di Nicola? Attenzione: è richiesto il numero degli schemi di gioco, non tutte le possibili disposizioni dei giocatori in campo. In altre parole: i calciatori sono indistinguibili, senza ruolo.

7 Calcola il coefficiente del monomio x^3 nello sviluppo di $(2x-1)^7$.

Valutazione

Esercizio	1	2	3	4	5	6	7	**Totale**
Punteggio	1	1,25	$0{,}75 \cdot 2 = 1{,}5$	1,5	$0{,}75 + 1 = 1{,}75$	1,5	1,5	10
Punteggio ottenuto								

Tempo indicativo: 1 h

Risposte p. 219

Unità 4

Tema F

Calcolo delle probabilità

- Approfondimenti
- Con GeoGebra
- Videolezioni
- Esercizi interattivi

1. Richiami di calcolo delle probabilità

Molte situazioni che si presentano usualmente sono caratterizzate dall'*incertezza* su ciò che accadrà nel futuro. Il *calcolo delle probabilità* è la parte della matematica che si occupa di elaborare dei modelli per descrivere queste situazioni.
Oggigiorno le tecniche proprie di questa disciplina, inizialmente nate dallo studio dei giochi d'azzardo, trovano applicazione in svariati settori: fisica, ingegneria, informatica, statistica, controllo della qualità, gestione della sicurezza delle comunicazioni, affidabilità dei sistemi ecc.
Prima di approfondire lo studio del calcolo della probabilità iniziato nel **Volume 2** è opportuno rivedere il linguaggio da essa adottato.

Esperimento aleatorio, spazio campionario ed eventi

Il lancio di una moneta, l'estrazione di un numero al lotto, il lancio di un dado sono esempi di fenomeni il cui esito dipende in modo imprevedibile dal caso: fenomeni di questo tipo, il cui risultato non può essere previsto con certezza, vengono detti **esperimenti aleatori** (o **casuali**).
A proposito degli esperimenti aleatori si introducono le seguenti definizioni.

ATTENZIONE!
Nel prosieguo, talvolta, parleremo per brevità di «spazio» intendendo «spazio campionario».

DEFINIZIONE | Spazio campionario

Si dice **spazio campionario** (o **spazio dei campioni** o **spazio dei risultati**), e si indica con il simbolo Ω, l'insieme di tutti i possibili esiti di un esperimento aleatorio.

OSSERVA
Uno spazio campionario può essere *finito* (vedi i primi due esempi in tab. 17.1) o *infinito* (vedi il terzo esempio in tab. 17.1).

DEFINIZIONE | Evento

Dato uno spazio campionario Ω, si chiama **evento** ogni sottoinsieme di Ω.

Un *evento* viene solitamente indicato con una lettera maiuscola dell'alfabeto; oltre a essere rappresentato dal sottoinsieme di Ω che lo individua, può essere descritto a parole, come messo in evidenza nei seguenti esempi.

Tabella 1

Esperimento aleatorio	*Lanciamo un dado regolare a 6 facce e osserviamo quale numero esce.*	*Lanciamo una moneta regolare successivamente per due volte e prendiamo nota, in ciascuno dei due lanci, se esce «testa» (T) o «croce» (C).*	*Lanciamo successivamente una moneta regolare e osserviamo il numero del lancio in cui esce «testa» per la prima volta.*
Spazio campionario	I possibili esiti dell'esperimento sono sei: 1, 2, 3, 4, 5, 6 Pertanto lo spazio campionario è l'insieme: $\Omega = \{1, 2, 3, 4, 5, 6\}$	L'esperimento ha quattro possibili esiti: esce «testa» entrambe le volte, esce «testa» la prima volta e «croce» la seconda, esce «croce» la prima volta e «testa» la seconda, esce «croce» in entrambi i lanci; quindi: $\Omega = \{(T, T), (T, C), (C, T), (C, C)\}$	Il risultato di questo esperimento aleatorio può essere qualsiasi numero naturale diverso da zero, dunque lo spazio campionario è: $\Omega = \mathbf{N} - \{0\}$
Esempio di evento	*E*: «esce un numero pari» Rappresentazione insiemistica: $E = \{2, 4, 6\}$	*E*: «esce 'testa' almeno una volta» Rappresentazione insiemistica: $E = \{(T, T),\ (T, C),\ (C, T)\}$	*E*: «esce 'testa' per la prima volta dopo il quinto lancio» Rappresentazione insiemistica: $E = \{x \in \Omega \mid x > 5\}$

Come emerge dall'analisi degli esempi precedenti, lo spazio campionario può essere *finito* o *infinito*; in questo volume ci limiteremo a considerare (salvo avviso contrario) casi in cui lo spazio campionario Ω è *finito* (analizzeremo il caso in cui Ω è infinito nel **Volume 5**).
Se Ω è finito, si danno ad alcuni particolari eventi dei nomi specifici (**Tab. 2**).

Tabella 2

Un evento si dice...	**... elementare,** se è rappresentato da un sottoinsieme di Ω costituito da un solo elemento	**... certo,** se è rappresentato dall'intero spazio campionario	**... impossibile,** se è rappresentato dall'insieme vuoto
Esempio, relativo al lancio di un dado	l'evento: «esce il numero 1»	l'evento: «esce un numero minore di 7»	l'evento: «esce un numero maggiore di 8»

Poiché abbiamo definito gli *eventi* come particolari *insiemi* (sottoinsiemi dello spazio campionario), possiamo definire, mediante le ordinarie operazioni tra insiemi, delle operazioni tra eventi.

DEFINIZIONI | **Operazioni tra eventi**

Dati due eventi A e B appartenenti a uno spazio campionario Ω:

a. si definisce evento **unione** di A e B, e si indica con $A \cup B$, l'evento che si realizza quando si realizzano A o B (o entrambi);
b. si definisce evento **intersezione** di A e B, e si indica con $A \cap B$, l'evento che si realizza quando si realizzano entrambi gli eventi A e B;
c. si definisce evento **contrario** di A, e si indica con $\overline{A}$, l'evento che si realizza quando **non** si realizza A, ossia l'evento rappresentato dal complementare di A: $\overline{A} = \Omega - A$.

Per esempio, consideriamo l'esperimento che consiste nel lancio di un dado e i due eventi:

A: «è uscito un numero maggiore di 2»
B: «è uscito un numero minore di 4».

Allora vale quanto riassunto nella **Tab. 3**.

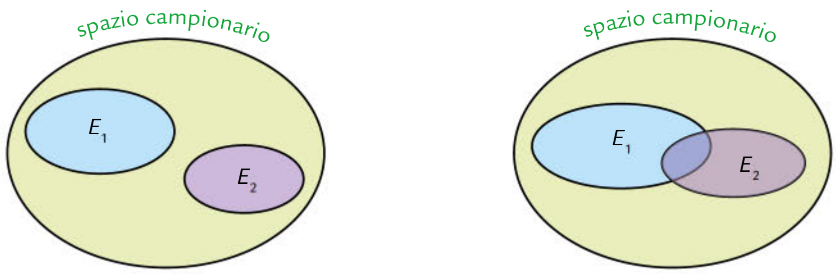
Tabella 3

Evento	**A parole**	**Notazione insiemistica**
Unione di A e B	È uscito un numero maggiore di 2 o minore di 4	$A \cup B = \Omega$
Intersezione di A e B	È uscito un numero maggiore di 2 e minore di 4	$A \cap B = \{3\}$
Contrario di A	È uscito un numero *minore* o *uguale* a 2	$\overline{A} = \{1, 2\}$
Contrario di B	È uscito un numero *maggiore* o *uguale* a 4	$\overline{B} = \{4, 5, 6\}$

Due eventi tali che il verificarsi dell'uno *esclude* il verificarsi dell'altro si dicono **incompatibili**; tramite il linguaggio degli insiemi la definizione di eventi incompatibili può esprimersi rigorosamente come segue.

DEFINIZIONE | Eventi incompatibili

Due eventi si dicono **incompatibili** se la loro intersezione è vuota; si dicono **compatibili** in caso contrario (**Fig. 1**).



a. I due eventi E_1 ed E_2 sono *incompatibili* perché la loro intersezione è vuota.

b. I due eventi E_1 ed E_2 sono *compatibili* perché la loro intersezione **non** è vuota.

Figura 1

Esempi	Controesempi
Nel lancio di un dado, i due eventi «esce il numero 5» ed «esce un numero pari» sono incompatibili.	Nel lancio di un dado, i due eventi «esce un multiplo di 3» ed «esce un numero pari» **non** sono incompatibili.
Nel lancio successivo di due monete, i due eventi «esce due volte 'testa'» ed «esce due volte 'croce'» sono incompatibili.	Nel lancio successivo di due monete, i due eventi «esce almeno una volta 'testa'» ed «esce almeno una volta 'croce'» **non** sono incompatibili.

SINTESI

Il linguaggio in teoria degli insiemi e nel calcolo della probabilità

Notazione	Teoria degli insiemi	Calcolo della probabilità
Ω	Insieme universo	Spazio campionario
$a \in \Omega$	a è un elemento di Ω	$\{a\}$ è un evento elementare
$A \subseteq \Omega$	A è un sottoinsieme di Ω	A è un evento
$A = \Omega$	A è l'insieme universo	A è l'evento certo
$A = \varnothing$	A è l'insieme vuoto	A è l'evento impossibile
$\overline{A}$	$\overline{A}$ è il complementare di A	$\overline{A}$ è l'evento contrario di A
$X = A \cup B$	X è l'unione di A e B	X è l'evento «A o B»
$X = A \cap B$	X è l'intersezione di A e B	X è l'evento «A e B»
$A \cap B = \varnothing$	A e B sono disgiunti	A e B sono incompatibili

Il concetto di probabilità

Ognuno di noi possiede un'idea, almeno vaga, del concetto di probabilità. Intuitivamente, possiamo dire che la **probabilità** di un evento E è un numero che esprime il *grado di fiducia* attribuito al verificarsi di E.
Resta però da capire *come* attribuire a un evento la sua probabilità, cioè come determinare il numero che esprime il suddetto «grado di fiducia». Nei casi più semplici il modo di attribuire la probabilità a un evento è in realtà del tutto naturale e intuitivo.

◆ **PROBLEMA**

Estraiamo una carta da un mazzo di 52 carte; qual è la probabilità di ottenere una figura?

Sappiamo che in un mazzo di 52 carte ci sono in tutto 12 figure (3 per ogni seme), quindi abbiamo in tutto 12 possibilità su 52 di estrarre una figura; supponendo che tutte le carte abbiano la stessa possibilità di essere estratte, siamo portati intuitivamente ad affermare che la probabilità di estrarre una figura è uguale a $\frac{12}{52}$, cioè $\frac{3}{13} \simeq 23\%$. Questo tipo di approccio si può formalizzare nella seguente definizione.

DEFINIZIONE | Definizione classica di probabilità

Consideriamo un evento E relativo a uno spazio campionario Ω in cui tutti gli eventi elementari hanno la stessa possibilità di verificarsi; supponiamo che l'evento E sia formato da k eventi elementari (brevemente detti «casi favorevoli») e lo spazio campionario Ω sia formato da n eventi elementari (brevemente detti «casi possibili»). Si definisce **probabilità** dell'evento E, e si indica con $p(E)$, il **rapporto tra il numero dei casi favorevoli e il numero dei casi possibili**:

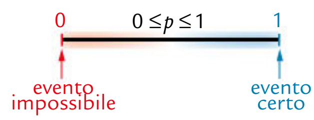
$$p(E) = \frac{k}{n}$$

MODI DI DIRE

Per riferirsi alla probabilità di un evento calcolata secondo la definizione classica si parla talvolta di probabilità *teorica* o *razionale* dell'evento.

Osserva che, poiché è $k \leq n$, risulta sempre:

$$0 \leq p(E) \leq 1$$



In particolare, se $p(E) = 0$ (cioè se $k = 0$) l'evento E è **impossibile**; se $p(E) = 1$ (cioè se $k = n$) l'evento E è certo.
Nei prossimi paragrafi ci occuperemo del calcolo della probabilità secondo la definizione *classica* esposta poc'anzi (che è la più intuitiva e che, storicamente, è stata la prima a essere data, a opera del matematico francese Laplace). È bene tuttavia osservare fin d'ora che la definizione classica ha anche dei *limiti*, dovuti al fatto che è applicabile soltanto a spazi campionari *finiti* in cui tutti gli eventi elementari sono *equipossibili* (**ipotesi di equiprobabilità**). Torneremo ampiamente su questo punto nell'ultimo paragrafo di questa Unità, presentando altre definizioni di probabilità che si adattano ai casi in cui la definizione classica è inapplicabile.

Esercizi p. 166

2. Valutazione della probabilità secondo la definizione classica

In questo paragrafo approfondiamo alcuni aspetti legati alla definizione classica di probabilità:

1. discuteremo l'importanza di valutare con attenzione l'ipotesi di *equiprobabilità*;
2. ci soffermeremo sulle tecniche di base per calcolare la probabilità secondo la definizione classica.

L'ipotesi di equiprobabilità

Per indicare che in un dato problema si sta supponendo l'ipotesi di *equiprobabilità* degli esiti dell'esperimento si usa di solito l'espressione «a caso» (per esempio: si estrae *a caso* una pallina da un'urna); inoltre, nei comuni giochi di estrazione di biglie e di lanci di monete o dadi, si usa specificare che i dadi o le monete si suppongono regolari (ossia non truccati) e che le palline estratte sono indistinguibili al tatto. Anche se queste parole «spia» devono subito farci pensare all'ipotesi di equiprobabilità, occorre non dimenticarsi mai di *valutare con attenzione* se tale ipotesi è effettivamente verificata *in relazione allo spazio campionario scelto*, cioè se tutti gli eventi elementari sono effettivamente equiprobabili. Assumere indebitamente l'equiprobabilità degli eventi elementari è una delle più frequenti cause di errore.

DALLA STORIA

L'errore descritto qui a fianco era stato commesso anche dal matematico e filosofo francese Jean-Baptiste d'Alembert.

ESEMPIO Importanza dell'ipotesi di equiprobabilità

Barbara sostiene che, lanciando due monete, la probabilità che escano 2 «testa» è $\frac{1}{3}$, in base al seguente ragionamento: ci sono tre casi possibili (escono 0 «testa» oppure esce 1 «testa» oppure escono 2 «testa») e di questi tre casi uno solo è quello favorevole (l'uscita di 2 «testa»). Quale errore sta commettendo Barbara?

Barbara ha assunto come spazio campionario dell'esperimento l'insieme:

$$\Omega = \{0, 1, 2\} \qquad [1]$$

costituito da tutti i possibili numeri di «testa» che possono uscire, quindi ha calcolato la probabilità dell'evento richiesto secondo la *definizione classica*. Ma per poter applicare la definizione classica gli eventi elementari dello spazio campionario devono essere tutti *equiprobabili*, mentre in questo caso **non** lo sono! Infatti, immaginiamo per aiutare la nostra intuizione che le due monete siano di due colori diversi, per esempio una bianca e una nera: allora l'evento «esce 1 'testa'» può realizzarsi in *due* modi diversi (quando esce «testa» sulla moneta bianca e «croce» su quella nera oppure quando esce «testa» sulla moneta nera e «croce» su quella bianca), quindi ha una *maggiore* probabilità di realizzarsi rispetto ai due eventi «escono 0 'testa'» ed «escono 2 'testa'» (ciascuno dei quali può realizzarsi in un solo modo). L'errore di Barbara è stato quindi quello di non rendersi conto che, assunto come spazio campionario l'insieme [1], gli eventi elementari **non** sono equiprobabili, perciò la definizione classica **non** è applicabile.
Per procedere correttamente occorre assumere come spazio campionario:

$$\Omega = \{(\mathrm{T}, \mathrm{T}), (\mathrm{T}, \mathrm{C}), (\mathrm{C}, \mathrm{T}), (\mathrm{C}, \mathrm{C})\}$$

Così facendo gli eventi elementari sono *equiprobabili* e la definizione classica è applicabile; si conclude così che la probabilità che escano 2 «testa» è $\frac{1}{4}$.

RIFLETTI

Come è emerso dall'esempio a fianco, nella risoluzione di un problema di calcolo della probabilità occorre prestare particolare attenzione nella *scelta* dello *spazio campionario*: se è possibile, per evitare errori e facilitare i calcoli, conviene sempre scegliere spazi campionari i cui eventi elementari risultano *equiprobabili*.

COLLEGHIAMO I CONCETTI **Probabilità e geometria**

In questa unità ci occupiamo di calcolo della probabilità nel caso di spazi campionari *finiti*. Esistono però situazioni in cui lo spazio campionario è *infinito*, le quali possono essere affrontate con approccio simile a quello della *probabilità classica*. Per esempio, se Ω è un segmento AB e scegliamo *a caso* un punto P del segmento AB, la probabilità che il punto P appartenga a un segmento CD contenuto in AB (**Fig. 2**) è data dal rapporto tra la lunghezza di CD e quella di AB.

$$p(P \in CD) = \frac{CD}{AB}$$

Figura 2

In modo analogo, se Ω è una superficie piana (oppure una figura solida) e scegliamo un punto P a caso di Ω, la probabilità che P appartenga a un sottoinsieme $E \subseteq \Omega$ è data dal rapporto tra l'area (oppure il volume) di E e l'area (oppure il volume) di Ω. In questi esempi, l'ipotesi di *equiprobabilità* consiste nel fatto che il punto P sia scelto «a caso», cioè in modo che nessuna parte di Ω risulti «privilegiata».

ESEMPIO

Si sceglie un punto P a caso all'interno di un quadrato di lato a (**Fig. 3**). Qual è la probabilità dell'evento E: «il punto P si trova all'interno del cerchio inscritto nel quadrato»?

$$p(E) = \frac{\text{area(cerchio inscritto)}}{\text{area quadrato}} = \frac{\pi\left(\frac{a}{2}\right)^2}{a^2} = \frac{\pi}{4} \simeq 0{,}7854 = 78{,}54\%$$

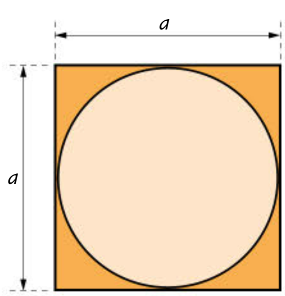


Figura 3

Utilizzo di diagrammi ad albero e tabelle a doppia entrata

La definizione «classica» traduce sostanzialmente problemi di probabilità in problemi di *conteggio*: per determinare la probabilità di un evento E dobbiamo calcolare il numero degli elementi di E, il numero degli elementi dello spazio campionario Ω e infine costruirne il rapporto. Per eseguire questi conteggi, almeno nei casi più semplici, può essere utile ricorrere a rappresentazioni tramite *diagrammi ad albero* o *tabelle a doppia entrata*.

ESEMPIO Utilizzo di un diagramma ad albero

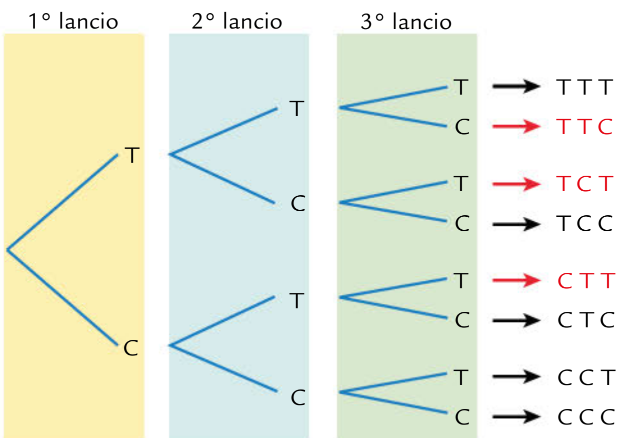
Lanciamo una moneta regolare, successivamente, per tre volte. Determiniamo la probabilità che esca «testa» esattamente *due* volte.

- **Analisi preliminare**

Lo spazio campionario è chiaramente *finito* e l'ipotesi che la moneta sia regolare ci consente di affermare che siamo in una situazione in cui tutti gli esiti sono *equiprobabili*; possiamo quindi valutare la probabilità richiesta come rapporto tra casi favorevoli e casi possibili.

- **Calcolo della probabilità**

Per individuare tutti i casi possibili possiamo utilizzare il seguente diagramma ad albero:



Ne deduciamo che ci sono complessivamente 8 casi possibili, di cui 3 (colorati in rosso) favorevoli alla realizzazione dell'evento «esce 'testa' esattamente due volte».
La probabilità di tale evento è quindi $\frac{3}{8}$.

ESEMPIO Utilizzo di una tabella a doppia entrata

Lanciamo due dadi regolari, uno rosso e uno blu, ciascuno con le facce numerate da 1 a 6. Determiniamo la probabilità che la somma dei due numeri ottenuti sia maggiore di 8.

Con GeoGebra
Simulazione del lancio di due dadi

- **Analisi preliminare e scelta dello spazio campionario**

Ecco un altro problema in cui occorre prestare particolare attenzione all'ipotesi di *equiprobabilità*!
Lanciando due dadi, la somma dei due numeri ottenuti può essere:

2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12

Si potrebbe quindi essere tentati di scegliere come spazio campionario l'insieme:

$$\Omega = \{2,\ 3,\ 4,\ 5,\ 6,\ 7,\ 8,\ 9,\ 10,\ 11,\ 12\} \qquad [2]$$

quindi osservare che gli esiti favorevoli sono quattro (9, 10, 11, 12), quelli possibili in tutto 11 (gli elementi di Ω), e concludere che la probabilità richiesta è $\frac{4}{11}$.

Questa soluzione però è **errata**; infatti gli eventi elementari dello spazio campionario [2] **non** sono tutti *equiprobabili*. Per esempio, la probabilità di ottenere il numero 2 (che può ottenersi solo dall'uscita del numero 1 su entrambi i dadi) deve chiaramente essere *inferiore* alla probabilità di ottenere il numero 6 (che può ottenersi in tanti modi diversi: sia dall'uscita di 1 e 5, sia dall'uscita di 2 e 4, sia dall'uscita di due 3).
Per avere l'equiprobabilità degli eventi elementari conviene scegliere come spazio campionario semplicemente l'insieme che rappresenta tutti i possibili esiti dei due lanci (senza considerare momentaneamente la somma dei numeri ottenuti); lo spazio campionario sarà quindi l'insieme:

$$\Omega' = \{(1,1),(1,2),(1,3),(1,4),(1,5),(1,6),(2,1),(2,2),(2,3),\ldots,(6,4),(6,5),(6,6)\}$$

RICORDA
Dato un insieme finito X, indichiamo il numero dei suoi elementi, ovvero la sua cardinalità, con il simbolo $|X|$.

I suoi eventi elementari, essendo i dadi regolari, si possono considerare tutti *equiprobabili*.
L'insieme Ω' non è altro che il prodotto cartesiano dell'insieme $\{1, 2, 3, 4, 5, 6\}$ per se stesso, quindi $|\Omega'| = 36$, ovvero ci sono in totale 36 *casi possibili*.

- **Calcolo dei casi favorevoli**

Possiamo rappresentare tutte le possibili somme ottenute dai numeri usciti nel lancio dei due dadi nella seguente tabella. Per esempio, la casella all'incrocio della seconda riga e della seconda colonna rappresenta la somma ottenuta in seguito all'uscita dei due numeri 1 e 1, quindi in corrispondenza di tale casella abbiamo scritto il numero $2 = 1 + 1$.

Dado rosso / Dado blu	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

In tabella abbiamo evidenziato i casi in cui la somma dei numeri ottenuti è *maggiore di* 8: deduciamo che abbiamo in tutto 10 *casi favorevoli*.

- **Conclusione**

Poiché ci sono 36 casi possibili e 10 favorevoli, la probabilità richiesta è uguale a $\dfrac{10}{36} = \dfrac{5}{18}$.

Utilizzo del calcolo combinatorio

Esclusi casi particolarmente semplici, come quelli esaminati negli ultimi due esempi, per calcolare il numero dei casi favorevoli e di quelli possibili è necessario fare ricorso alle regole del calcolo combinatorio, che abbiamo introdotto nell'Unità precedente.

ESEMPIO Terno al lotto

Qual è la probabilità di fare un terno al lotto, cioè che si giochino tre numeri e questi facciano parte dei cinque estratti?

I *casi possibili* sono tanti quanti le cinquine **non** ordinate che è possibile costruire con i numeri 1, 2, ..., 90; esse sono:

$$\binom{90}{5}$$

I casi *favorevoli* sono le cinquine **non** ordinate che contengono i tre numeri giocati; esse sono tante quanti i modi di scegliere due numeri tra gli 87 rimanenti, dunque sono:

$$\binom{87}{2}$$

La probabilità p di fare terno è dunque uguale a

$$p = \frac{\binom{87}{2}}{\binom{90}{5}} = \frac{\frac{87 \cdot 86}{2}}{\frac{90 \cdot 89 \cdot 88 \cdot 87 \cdot 86}{5 \cdot 4 \cdot 3 \cdot 2}} = \frac{5 \cdot 4 \cdot 3}{90 \cdot 89 \cdot 88} = \frac{1}{11\,748} \simeq 0{,}000085$$

Sebbene ci sia 1 possibilità di vincere su 11 747, la realizzazione di un terno al lotto viene pagata solo 4500 volte: si tratta quindi di un gioco molto a favore del banco (il gioco sarebbe invece *equo* se la vittoria venisse pagata 11 747 volte).

ESEMPIO 100 metri piani

Gli otto centometristi in finale sono ai blocchi di partenza. Qual è la probabilità di indovinare i primi tre all'arrivo, nell'esatto ordine? Supponi che la scelta dei tre atleti sia casuale (e non dettata da considerazioni di tipo tecnico-sportivo).

Primo modo
L'ordine d'arrivo è essenziale. Conviene quindi considerare come spazio campionario quello di tutti i possibili ordini d'arrivo degli 8 velocisti: con questa scelta, i casi possibili sono 8!. I casi favorevoli sono 5!, cioè tanti quanti i possibili ordini d'arrivo per i corridori dal quarto all'ultimo: le prime tre posizioni infatti sono «obbligate» dal pronostico. La probabilità cercata è quindi $\frac{5!}{8!} = \frac{1}{336}$.

Secondo modo
Consideriamo come spazio campionario quello di tutti i possibili ordini d'arrivo dei *primi tre* velocisti soltanto (visto che le posizioni dal quarto all'ultimo sono ininfluenti): con questa scelta i casi possibili sono $D_{8,3} = 336$. Poichè vi è 1 solo caso favorevole (la terna pronosticata), ritroviamo la stessa probabilità calcolata in precedenza, cioè $\frac{1}{336}$.

Esercizi p. 168

3. I primi teoremi sul calcolo delle probabilità

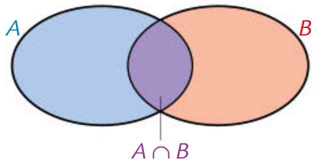
Probabilità dell'unione di due eventi

Il primo teorema di calcolo della probabilità che dimostriamo riguarda l'*unione* di due eventi. Esso si basa sul cosiddetto **principio di addizione e sottrazione**.

DEFINIZIONE | Principio di addizione e sottrazione

Dati due insiemi A e B, il numero degli elementi dell'insieme unione di A e B è dato dalla formula:

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



La giustificazione di questo principio è ovvia: per calcolare il numero di elementi dell'insieme $A \cup B$ è necessario sottrarre, dalla somma tra il numero degli elementi di A e il numero di B, il numero degli elementi di $A \cap B$, altrimenti questi ultimi verrebbero contati due volte. Da questo principio segue immediatamente il seguente teorema.

TEOREMA 1 | Probabilità dell'unione di due eventi

Siano A e B due eventi; allora risulta:

$$p(A \cup B) = p(A) + p(B) - p(A \cap B) \qquad [3]$$

In particolare, se A e B sono incompatibili:

$$p(A \cup B) = p(A) + p(B) \qquad [4]$$

DIMOSTRAZIONE

a. Indichiamo con Ω lo spazio campionario. Abbiamo:

$$p(A \cup B) = \frac{|A \cup B|}{|\Omega|} = \qquad \text{Definizione di probabilità in senso classico}$$

$$= \frac{|A| + |B| - |A \cap B|}{|\Omega|} = \qquad \text{Per il principio di addizione e sottrazione}$$

$$= \frac{|A|}{|\Omega|} + \frac{|B|}{|\Omega|} - \frac{|A \cap B|}{|\Omega|} = \qquad \text{Proprietà distributiva (a destra) della divisione}$$

$$= p(A) + p(B) - p(A \cap B) \qquad \text{Definizione di probabilità in senso classico}$$

b. Nel caso particolare in cui gli eventi A e B sono incompatibili, risulta $A \cap B = \emptyset$, quindi $p(A \cap B) = 0$, perciò la [3] si riduce alla [4].

ESEMPIO Probabilità dell'unione di due eventi

In un lotto di pezzi meccanici, il 3% ha il peso sbagliato, il 5% ha il diametro sbagliato e il 2% ha sia il peso sia il diametro sbagliati. Qual è la probabilità che, estratto a caso un pezzo dal lotto, esso risulti difettoso (cioè abbia peso sbagliato o diametro sbagliato)?

Sia A l'evento: «il pezzo estratto ha il peso sbagliato» e B l'evento: «il pezzo estratto ha il diametro sbagliato». In base ai dati sappiamo che:

$$p(A) = 0{,}03 \qquad p(B) = 0{,}05 \qquad p(A \cap B) = 0{,}02$$

Dobbiamo calcolare la probabilità dell'evento $A \cup B$; in base alla [3] abbiamo:

$$p(A \cup B) = \underbrace{0{,}03}_{p(A)} + \underbrace{0{,}05}_{p(B)} - \underbrace{0{,}02}_{p(A \cap B)} = 0{,}06$$

Dunque la probabilità che un pezzo estratto a caso dal lotto sia difettoso è del 6%.

Probabilità dell'evento contrario

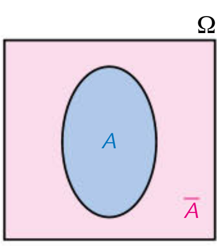
Vediamo ora quale relazione lega la probabilità di un evento con la probabilità dell'evento contrario.

TEOREMA 2 | Probabilità dell'evento contrario

Se A è un evento e $\overline{A}$ è il suo evento contrario, allora:

$$p(\overline{A}) = 1 - p(A) \qquad [5]$$

DIMOSTRAZIONE



Sia Ω lo spazio campionario cui appartiene l'evento A.

Poiché $\Omega = A \cup \overline{A}$ e inoltre i due eventi A e $\overline{A}$ sono incompatibili, abbiamo:

$$1 = p(\Omega) = p(A \cup \overline{A}) = p(A) + p(\overline{A})$$

da cui possiamo ricavare $p(\overline{A}) = 1 - p(A)$.

ESEMPIO

Un individuo su 10 000 è allergico a un certo vaccino. Qual è la probabilità di assumere quel vaccino senza manifestare fenomeni di allergia?

La probabilità di essere allergici al vaccino è $\frac{1}{10\,000}$; la probabilità di non presentare problemi dopo la somministrazione è $1 - \frac{1}{10\,000} = \frac{9999}{10\,000} = 99{,}99\%$.

Per calcolare la probabilità di un dato evento, in certi casi è più facile calcolare la probabilità dell'evento contrario e poi risalire alla probabilità dell'evento originario mediante la [5].

ESEMPIO Passaggio all'evento contrario

Si lancia successivamente per 6 volte una moneta non truccata; qual è la probabilità che esca «testa» almeno una volta?

- **Analisi preliminare**

Indichiamo con E l'evento: «esce 'testa' almeno una volta». Esso equivale a: «esce 'testa' esattamente 1 volta o esattamente 2 volte o esattamente 3 volte o esattamente 4 volte o esattamente 5 volte o esattamente 6 volte». Calcolare la probabilità dell'evento E seguendo questa via è possibile ma sarebbe laborioso: dovremmo calcolare la probabilità di tutti gli eventi in cui abbiamo scomposto E, ossia «esce 'testa' esattamente 1 volta», «esce 'testa' esattamente 2 volte» ecc., e poi, dato che tali eventi sono *incompatibili*, sommare le loro probabilità.
Il problema si risolve molto più agevolmente se si considera l'evento contrario $\overline{E}$: «non esce mai 'testa'» e poi si deduce mediante la [5] la probabilità di E.

SUGGERIMENTO

In generale, in molti problemi di calcolo delle probabilità in cui compare la parola «*almeno*», è conveniente passare all'evento contrario. Ogni qualvolta compare in un problema questa parola, chiediti quindi se può essere utile seguire questa via.

- **Calcolo**

Poiché la moneta non è truccata, possiamo supporre che tutti gli esiti siano equiprobabili e dunque calcolare la probabilità dell'evento $\overline{E}$ secondo la definizione classica:

– ci sono 2^6 *casi possibili*: infatti effettuiamo 6 lanci e per ogni lancio ci sono 2 possibili esiti («testa» o «croce»), quindi, per il principio fondamentale del calcolo combinatorio, i casi possibili sono complessivamente $2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 = 2^6$;
– c'è 1 solo *caso favorevole*: quello in cui esce «croce» in tutti i sei lanci.

Concludiamo che:

$$p(\overline{E}) = \frac{1}{2^6} = \frac{1}{64}$$

quindi, per la [5]:

$$p(E) = 1 - p(\overline{E}) = 1 - \frac{1}{64} = \frac{63}{64}$$

Esercizi p. 177

4. Probabilità composte ed eventi indipendenti

Probabilità condizionata

Vogliamo introdurre un importante concetto, quello di **probabilità condizionata**. Esso interviene ogni qualvolta si vuole calcolare la probabilità di un evento A, *sapendo che si è verificato un evento B*. L'informazione aggiuntiva che deriva dal sapere che si è verificato B può modificare la probabilità che si verifichi A, come mostra l'esempio seguente.

ESEMPIO Introduzione intuitiva al concetto di probabilità condizionata

Supponiamo di avere lanciato per due volte consecutive una moneta non truccata.

a. Sapendo che nel primo lancio è uscita «testa», qual è la probabilità che sia uscita «testa» in entrambi i lanci?

b. Sapendo che in almeno uno dei due lanci è uscita «testa», qual è la probabilità che sia uscita «testa» in entrambi i lanci?

Lo spazio campionario che rappresenta i possibili esiti dell'esperimento consistente nel lancio per due volte consecutive di una moneta non truccata è:

$$\Omega = \{(\mathrm{T, T}); (\mathrm{T, C}); (\mathrm{C, T}); (\mathrm{C, C})\}$$

e la probabilità di ottenere «testa» due volte, ossia dell'evento $\{(\mathrm{T, T})\}$, in assenza di ulteriori informazioni, è $\frac{1}{4}$. Le informazioni aggiuntive fornite nei due casi **a.** e **b.** modificano tale probabilità: vediamo come.

a. Sapere che *nel primo lancio* è uscita «testa» ci permette di «restringere» lo spazio campionario, eliminando le due coppie ordinate $(\mathrm{C, T})$ e $(\mathrm{C, C})$. Possiamo quindi assumere come nuovo spazio campionario:

$$\Omega' = \{(\mathrm{T, T}); (\mathrm{T, C})\}$$

I casi possibili sono ora soltanto 2, di cui uno solo favorevole, quindi la probabilità che siano uscite due «testa», sapendo che nel primo lancio è uscita «testa», diviene $\frac{1}{2}$.

b. Sapere che *in almeno uno dei due lanci* è uscita «testa» ci permette ancora di «restringere» lo spazio campionario, ma questa volta possiamo eliminare soltanto la coppia ordinata $(\mathrm{C, C})$. Possiamo quindi assumere come nuovo spazio campionario:

$$\Omega'' = \{(\mathrm{T, T}); (\mathrm{T, C}); (\mathrm{C, T})\}$$

I casi possibili sono ora 3, di cui uno solo favorevole, quindi la probabilità che siano uscite due «testa», sapendo che in almeno uno dei due lanci è uscita «testa», diviene $\frac{1}{3}$.

ATTENZIONE!

Il risultato ottenuto nel punto b. appare spesso controintuitivo, perché a un primo approccio viene naturale il seguente ragionamento: «sapendo che almeno uno dei due lanci ha dato come risultato 'testa', vi sono due sole possibilità: che sia uscita 'testa' in entrambi i lanci o che sia uscita 'testa' in uno solo dei due lanci. Ciascuno di questi ultimi due eventi può verificarsi *con la stessa probabilità*, quindi la probabilità che siano uscite due 'testa', sapendo che ne è uscita almeno una, è $\frac{1}{2}$». L'errore consiste nell'aver assunto indebitamente l'*equiprobabilità*: infatti, come abbiamo visto, lo spazio campionario nel caso b. è $\Omega'' = \{(\mathrm{T, T}); (\mathrm{T, C}); (\mathrm{C, T})\}$, dunque la probabilità che sia uscita «testa» in uno solo dei due lanci è *il doppio* della probabilità che sia uscita «testa» in entrambi i lanci.

Per fissare le idee, concentriamoci sul caso **a.** dell'esempio precedente. Poniamo:

A: «è uscita 'testa' in entrambi i lanci»
B: «è uscita 'testa' nel primo lancio»

La probabilità dell'evento A, sapendo che si è verificato B, viene indicata con il simbolo $p(A|B)$. Per calcolare tale probabilità abbiamo

- assunto come spazio campionario $\Omega' = \{(\mathrm{T, T}); (\mathrm{T, C})\}$, ossia l'evento B stesso;
- contato il numero dei casi possibili, ossia il numero degli elementi di B;
- contato il numero dei casi favorevoli, cioè il numero degli elementi di B che realizzano A, ossia il numero degli elementi di $A \cap B$;
- costruito il rapporto tra il numero dei casi favorevoli e il numero dei casi possibili.

In altre parole, è risultato che:

$$p(A|B) = \frac{\text{numero di elementi di } A \cap B}{\text{numero di elementi di } B} = \frac{|A \cap B|}{|B|}$$

Dividendo numeratore e denominatore dell'ultima frazione scritta per il numero degli elementi dello spazio campionario Ω, cioè per $|\Omega|$, otteniamo:

$$p(A|B) = \frac{|A \cap B|}{|B|} = \frac{\dfrac{|A \cap B|}{|\Omega|}}{\dfrac{|B|}{|\Omega|}} = \frac{p(A \cap B)}{p(B)}$$

Questo risultato motiva la seguente definizione.

DEFINIZIONE | Probabilità condizionata

Siano A e B due eventi, con B di probabilità non nulla. Si definisce **probabilità condizionata** dell'evento A dato l'evento B, e si indica con il simbolo $p(A|B)$, il rapporto tra la probabilità di $A \cap B$ e la probabilità di B (**fig. 4**):

$$p(A|B) = \frac{p(A \cap B)}{p(B)} \qquad [6]$$

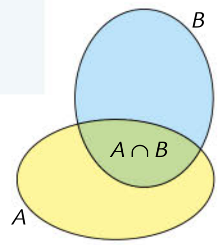


Figura 4 **Supposto che si sia realizzato l'evento B, i soli casi che realizzano l'evento A sono quelli di $A \cap B$. Inoltre, l'insieme B diviene lo spazio campionario. Pertanto la probabilità condizionata di A dato B viene calcolata come rapporto tra la probabilità di $A \cap B$ e quella di B.**

Dalla definizione stessa di probabilità condizionata segue l'uguaglianza seguente, detta **formula delle probabilità composte**, che si rivela spesso molto utile per il calcolo della probabilità dell'intersezione di due eventi:

$$p(A \cap B) = p(B) \cdot p(A|B) \qquad [7]$$

Scambiando nella [6] i ruoli di A e B otteniamo:

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

perciò la probabilità di $A \cap B$ può essere espressa anche mediante la seguente formula, simmetrica della [7]:

$$p(A \cap B) = p(A) \cdot p(B|A) \qquad [8]$$

ESEMPIO Applicazione della formula delle probabilità composte

Da un sacchetto contenente i 90 numeri della tombola si estraggono successivamente due numeri, senza rimettere nel sacchetto il primo numero estratto. Qual è la probabilità che il primo numero estratto sia divisibile per 10 e il secondo sia pari?

Indichiamo con A l'evento: «il primo numero è divisibile per 10» e con B l'evento: «il secondo numero è pari»; dobbiamo calcolare $p(A \cap B)$.
Allo scopo di utilizzare la [8] osserviamo che:

- $p(A) = \dfrac{9}{90} = \dfrac{1}{10}$
- $p(B|A)$ è la probabilità di estrarre un numero pari dopo che dal sacchetto è stato tolto un numero divisibile per 10 (cioè pari); dunque risulta $p(B|A) = \dfrac{44}{89}$.

Possiamo ora applicare la [8], ottenendo così:

$$p(A \cap B) = p(A) \cdot p(B|A) = \frac{1}{10} \cdot \frac{44}{89} = \frac{22}{445}$$

La formula delle probabilità composte può essere generalizzata al caso di più di due eventi. Per esempio, nel caso di tre eventi si può dimostrare che:

$$p(E_1 \cap E_2 \cap E_3) = p(E_1) \cdot p(E_2|E_1) \cdot p(E_3 \mid E_1 \cap E_2)$$

Più in generale, se un evento è l'intersezione degli n eventi $E_1, E_2, ..., E_n$, la probabilità che si verifichi E è data dal *prodotto* che ha come fattori:

- la probabilità di E_1;
- la probabilità di E_2, supposto verificato E_1;
- la probabilità di E_3, supposti verificati E_1 ed E_2;
- ....
- la probabilità di E_n, supposti verificati gli $n - 1$ eventi precedenti.

In formule:

$$p(E_1 \cap E_2 \cap \ldots \cap E_{n-1} \cap E_n) =$$
$$= p(E_1) \cdot p(E_2 \mid E_1) \cdot p(E_3 \mid E_1 \cap E_2) \cdot \ldots \cdot p(E_n \mid E_1 \cap \ldots \cap E_{n-1})$$

ESEMPIO Applicazione della formula delle probabilità composte generalizzata

Da un mazzo di 40 carte se ne estraggono quattro successivamente (senza rimettere le carte estratte nel mazzo). Qual è la probabilità che le carte estratte siano quattro Assi?

Si tratta di calcolare la probabilità dell'evento $p(E_1 \cap E_2 \cap E_3 \cap E_4)$, essendo E_1, E_2, E_3, E_4 rispettivamente gli eventi: «è uscito un Asso alla prima estrazione», «è uscito un Asso alla seconda estrazione», «è uscito un Asso alla terza estrazione», «è uscito un Asso alla quarta estrazione». Per applicare la formula delle probabilità composte generalizzata, osserviamo che:

- la probabilità di estrarre un Asso alla prima estrazione è $\frac{4}{40} = \frac{1}{10}$;
- la probabilità di estrarre un Asso alla seconda estrazione, se un Asso è stato già estratto nella prima estrazione, è $\frac{3}{39} = \frac{1}{13}$;
- la probabilità di estrarre un Asso alla terza estrazione, se due Assi sono stati già estratti nelle prime due estrazioni, è $\frac{2}{38} = \frac{1}{19}$;
- la probabilità di estrarre un Asso alla quarta estrazione, se tre Assi sono stati già estratti nelle prime tre estrazioni, è $\frac{1}{37}$.

La probabilità di estrarre quattro Assi è quindi: $\frac{1}{10} \cdot \frac{1}{13} \cdot \frac{1}{19} \cdot \frac{1}{37} = \frac{1}{91\,390}$.

Valgono infine le proprietà espresse nel teorema seguente, che estendono alle probabilità condizionate i teoremi relativi alla probabilità dell'evento complementare e dell'unione di due eventi.

TEOREMA 3 | Proprietà delle probabilità condizionate

Dati tre eventi A, B, H, con H di probabilità non nulla, risulta:

a. $p(\overline{A} \mid H) = 1 - p(A \mid H)$

b. $p(A \cup B \mid H) = p(A \mid H) + p(B \mid H) - p(A \cap B \mid H)$

In particolare, se A e B sono incompatibili:

$$p(A \cup B \mid H) = p(A \mid H) + p(B \mid H)$$

ATTENZIONE!

Se $p(H) \neq 1$, cioè se $p(\overline{H}) \neq 0$, ha senso considerare, oltre a $p(A|H)$, anche $p(A|\overline{H})$.
In generale, tuttavia, **non** è vero che:

$p(A|\overline{H}) = 1 - p(A|H)$ errato!

Eventi indipendenti

Intuitivamente, due eventi A e B si dicono indipendenti se il verificarsi dell'uno non altera la probabilità che si verifichi l'altro.
Formalmente, utilizzando la definizione di probabilità *condizionata*, possiamo dare la seguente definizione.

DEFINIZIONE | Eventi indipendenti

Due eventi A e B (con B di probabilità non nulla) si dicono **indipendenti** se la probabilità condizionata dell'evento A, dato l'evento B, è uguale alla probabilità di A.
In simboli, l'indipendenza tra A e B può esprimersi tramite la condizione:

$$p(A \mid B) = p(A) \qquad [9]$$

La definizione di eventi indipendenti è simmetrica, nel senso che può esprimersi in modo equivalente scambiando i ruoli di A e B; dunque si può dire, in modo equivalente, che due eventi A e B (con A di probabilità non nulla) sono indipendenti se e solo se:

$$p(B \mid A) = p(B)$$

Se A e B sono due eventi **indipendenti**, la formula delle probabilità composte diventa:

$$p(A \cap B) = p(A) \cdot p(B) \qquad [10]$$

nota anche come **regola del prodotto**.
Viceversa, se è vera la [10], allora i due eventi A e B sono indipendenti.

REGOLA | **Regola del prodotto per eventi indipendenti**

Due eventi A e B sono indipendenti se e solo se:

$$p(A \cap B) = p(A) \cdot p(B)$$

La regola del prodotto può essere utilizzata per stabilire se due eventi sono indipendenti, oppure per il calcolo della probabilità di $A \cap B$, nel caso in cui si sappia a priori che A e B sono indipendenti (l'indipendenza viene infatti spesso *assunta come ipotesi* nella costruzione del modello di un problema e successivamente utilizzata per il calcolo delle probabilità).
Utilizzando la [10] è facile dimostrare (prova a farlo per esercizio) che, se A e B sono due eventi indipendenti con probabilità diverse da 0 e 1, anche le coppie di eventi:

$$\overline{A} \text{ e } B, \quad A \text{ e } \overline{B}, \quad \overline{A} \text{ e } \overline{B}$$

sono indipendenti.

Approfondimento
Correlazione in probabilità e statistica

ESEMPI Stabilire se due eventi sono indipendenti

a. Lanciamo per due volte un dado e consideriamo i due eventi A: «il primo numero uscito è 2», B: «il secondo numero uscito è 5». I due eventi sono indipendenti?
b. Lanciamo per due volte un dado; siano A l'evento: «il primo numero uscito è 2» e B l'evento: «la somma dei due numeri usciti è 5». I due eventi sono indipendenti?

a. È intuitivo che il numero uscito nel primo lancio del dado non influenza il numero che uscirà nel secondo lancio, quindi ci aspettiamo che i due eventi siano indipendenti. Verifichiamolo in base alla [10]. È immediato che:

$$p(A) = \frac{1}{6} \quad \text{e} \quad p(B) = \frac{1}{6}$$

Per calcolare $p(A \cap B)$ notiamo che, assumendo come spazio campionario $\Omega = X \times X$, con $X = \{1, 2, 3, 4, 5, 6\}$, abbiamo 36 casi possibili (ed equiprobabili) di cui 1 solo favorevole, quindi $p(A \cap B) = \frac{1}{36}$.

È facile a questo punto riconoscere che $p(A) \cdot p(B) = p(A \cap B)$.

b. La possibilità di ottenere 5 come somma dei due numeri usciti dipende dal numero uscito nel primo lancio; per esempio, se il primo numero uscito fosse 6, sarebbe impossibile ottenere 5 come somma dei due numeri. Ci aspettiamo quindi che i due eventi **non** siano indipendenti. Verifichiamolo in base alla [10].

È semplice calcolare la probabilità dell'evento A: $p(A) = \frac{1}{6}$.

Per il calcolo della probabilità di B assumiamo come spazio campionario $\Omega = X \times X$, essendo $X = \{1, 2, 3, 4, 5, 6\}$. Allora $|\Omega| = 36$ e l'evento B è rappresentato dal seguente sottoinsieme di Ω:

$$\{(1, 4); (2, 3); (3, 2); (4, 1)\}$$

Dunque $p(B) = \frac{4}{36} = \frac{1}{9}$.

L'evento $A \cap B$, ossia «il primo numero uscito è 2 e la somma dei due numeri usciti è 5» è rappresentato da $\{(2, 3)\}$, quindi: $p(A \cap B) = \frac{1}{36}$.

È immediato a questo punto riconoscere che $p(A) \cdot p(B) \neq p(A \cap B)$.

ATTENZIONE!

Presta attenzione a non confondere i due concetti di eventi *indipendenti* e di eventi *incompatibili*. In particolare, è un errore frequente ritenere che due eventi incompatibili siano anche indipendenti; in realtà, è vero esattamente il contrario, cioè che due eventi *incompatibili* (aventi probabilità non nulla) non sono mai *indipendenti*. Infatti, in queste ipotesi la [10] non può essere verificata, essendo il primo membro *uguale a zero* (poiché $A \cap B = \emptyset$ e $p(A \cap B) = 0$) e il secondo membro *diverso da zero* (perché stiamo supponendo $p(A) \neq 0$ e $p(B) \neq 0$).

ESEMPIO Utilizzo della proprietà di indipendenza: i numeri ritardatari

Molti degli scommettitori del lotto giocano sui numeri ritardatari. Per esempio, nel 2005 il numero 53 non uscì sulla ruota di Venezia per ben 182 estrazioni; credendo che l'uscita del 53 fosse tanto più probabile quanto maggiore era il ritardo accumulato, ci fu un incremento notevole del numero di giocate. L'aspettativa degli scommettitori era fondata?

Per rispondere al problema occorre confrontare la probabilità di uscita del 53 in una estrazione qualsiasi con la probabilità di uscita del 53 *sapendo* che non è uscito nelle 182 estrazioni precedenti.

- **Calcoliamo la probabilità p^* di uscita del 53 in un'estrazione qualsiasi**

La probabilità che, in una estrazione di cinque numeri del lotto, il numero 53 esca per primo è $\frac{1}{90}$; la probabilità che su cinque numeri estratti il 53 esca per secondo è uguale alla precedente:

$$\underbrace{\frac{1}{89}}_{\substack{\text{probabilità di}\\\text{estrarre il 53}\\\text{nella seconda}\\\text{estrazione}}} \cdot \underbrace{\frac{89}{90}}_{\substack{\text{probabilità di}\\\text{non avere}\\\text{estratto il 53 nella}\\\text{prima estrazione}}} = \frac{1}{90} \quad \text{e così via}$$

Dunque la probabilità che il 53 esca tra i cinque numeri estratti è $p^* = \frac{5}{90} = \frac{1}{18}$.

- **Calcoliamo la probabilità di uscita del 53 alla 183-esima estrazione, sapendo che non è uscito nelle 182 estrazioni precedenti**

Assumiamo, come è del tutto ragionevole, che le estrazioni siano *indipendenti* (il che ci consentirà di utilizzare la regola del prodotto) e indichiamo con E_i l'evento «il numero 53 esce alla i-esima estrazione», con i variabile tra 1 e 183; dobbiamo calcolare:

$$p(\underbrace{E_{183}}_{\substack{\text{il 53 esce}\\\text{alla 183-esima}\\\text{estrazione}}} \underbrace{|}_{\substack{\text{sapendo}\\\text{che}}} \underbrace{\overline{E_1} \cap \overline{E_2} \cap \dots\dots \cap \overline{E_{182}}}_{\substack{\text{non è uscito nelle 182}\\\text{estrazioni precedenti}}}) =$$

$$= \frac{p(E_{183} \cap \overline{E_1} \cap \overline{E_2} \cap \dots\dots \cap \overline{E_{182}})}{p(\overline{E_1} \cap \overline{E_2} \cap \dots\dots \cap \overline{E_{182}})} =$$

Definizione di probabilità condizionata

$$= \frac{p^*(1-p^*)^{182}}{(1-p^*)^{182}} = p^*$$

La probabilità che il numero 53 esca alla 183-esima estrazione è p^*, mentre la probabilità che non esca nelle 182 estrazioni precedenti, per la regola del prodotto, è uguale a $(1-p^*)^{182}$

Abbiamo dunque trovato che la probabilità di uscita del 53 *sapendo* che non è uscito nelle 182 estrazioni precedenti è *uguale* alla probabilità di uscita in un'estrazione *qualsiasi*. Puoi quindi comprendere il significato della frase, spesso citata, «il caso non ha memoria».

- **Commentiamo il risultato**

Il risultato trovato appare a prima vista sorprendente, poiché ci si aspetta che il ritardo nell'uscita debba in qualche modo ridurre l'attesa di uscita successiva. In realtà, riflettendo più a fondo, è del tutto comprensibile: infatti le estrazioni sono *indipendenti*, quindi l'esito di nessuna di esse può influenzare gli esiti delle estrazioni successive.

Sarebbe certamente irragionevole scommettere sul fatto che il 53 non esca per 182 estrazioni, essendo estremamente bassa la probabilità di tale evento (infatti $(1-p^*)^{182} = \left(\frac{17}{18}\right)^{182} \simeq 0{,}00003$).

Tuttavia, se questo evento estremamente improbabile è *già accaduto* (come nel caso del ritardo descritto), **non** sarebbe irragionevole scommettere che anche nella 183-esima estrazione il 53 non uscirà. Infatti, in base a quanto abbiamo visto, la probabilità di uscita del 53 alla 183-esima estrazione è ancora $\frac{1}{18} \simeq 5{,}6\%$ e, di conseguenza, quella che non esca è $\frac{17}{18} \simeq 94{,}4\%$: di molto superiore!

Il problema delle prove ripetute

Consideriamo il seguente problema.

◆ **PROBLEMA**
In un compito in classe Paolo deve rispondere a 5 quesiti a risposta multipla: ogni quesito è costituito da 4 risposte, di cui una sola è quella esatta. Paolo risponde a caso a tutti i quesiti, in modo indipendente uno dall'altro. Qual è la probabilità che dia esattamente tre risposte esatte?

Osserviamo che:

- rispondere a caso ai 5 quesiti equivale a ripetere 5 volte l'«esperimento» che consiste nella scelta a caso di una risposta tra 4 possibili;
- ciascun «esperimento» può avere *due* possibili esiti: la scelta della risposta *esatta* (diciamo che si è avuto un «successo»), con probabilità $p = \frac{1}{4}$, oppure la scelta della risposta *sbagliata* (diciamo che si è avuto un «insuccesso») con probabilità $1 - p$, cioè $\frac{3}{4}$.

Con questa interpretazione, il problema chiede di determinare *la probabilità di avere 3 successi, nell'esecuzione di 5 prove indipendenti, in ciascuna delle quali la probabilità di successo* è $\frac{1}{4}$. Il calcolo di tale probabilità si ottiene in base alle seguenti quattro osservazioni.

1. Esaminiamo preliminarmente un caso *particolare*: supponiamo per esempio che Paolo risponda correttamente ai primi 3 quesiti e che sbagli le risposte agli ultimi 2 quesiti; la probabilità di questo singolo evento, dal momento che stiamo supponendo le risposte *indipendenti*, è uguale a:

$$\underbrace{\frac{1}{4} \cdot \frac{1}{4} \cdot \frac{1}{4}}_{\text{risposte esatte nei primi 3 quesiti}} \cdot \underbrace{\frac{3}{4} \cdot \frac{3}{4}}_{\text{risposte sbagliate negli ultimi due}} = \left(\frac{1}{4}\right)^3 \left(\frac{3}{4}\right)^2$$

2. I 3 successi possono presentarsi con *ordine diverso*; per esempio Paolo potrebbe rispondere correttamente al primo, al terzo e all'ultimo quesito invece che ai primi tre, ma la probabilità dell'evento corrispondente *resterebbe uguale* a quella calcolata in 1. Precisamente, in quest'ultimo caso la probabilità sarebbe:

$$\underbrace{\frac{1}{4} \cdot \frac{3}{4} \cdot \frac{1}{4} \cdot \frac{3}{4} \cdot \frac{1}{4}}_{\text{Le risposte esatte sono la prima, la terza e l'ultima}} = \left(\frac{1}{4}\right)^3 \left(\frac{3}{4}\right)^2$$

3. I possibili ordini diversi con cui possono presentarsi 3 successi e 2 insuccessi è uguale a $\binom{5}{3}$: infatti una sequenza costituita da 3 successi e 2 insuccessi è individuata univocamente una volta che sono note le 3 posizioni dei successi e i modi in cui si possono scegliere 3 posizioni tra 5 è uguale a $\binom{5}{3}$.

4. Poiché i $\binom{5}{3}$ modi in cui si possono realizzare i 3 successi sono eventi *incompatibili*, ciascuno di probabilità $\left(\frac{1}{4}\right)^3\left(\frac{3}{4}\right)^2$, la probabilità cercata risulta uguale alla *somma* delle probabilità di tutti questi eventi incompatibili, quindi è uguale a:

$$\binom{5}{3}\left(\frac{1}{4}\right)^3\left(\frac{3}{4}\right)^2 = 10 \cdot \frac{1}{64} \cdot \frac{9}{16} = \frac{45}{512} \simeq 0{,}088 = 8{,}8\%$$

Ragionamenti analoghi possono ripetersi più in generale e conducono al seguente teorema.

TEOREMA 4 | Prove ripetute

La probabilità che si realizzino k successi nell'esecuzione di n prove identiche e indipendenti, in ciascuna delle quali la probabilità di successo è p, è uguale a:

$$\binom{n}{k} p^k (1-p)^{n-k}$$

Esercizi p. 182

5. Il teorema di disintegrazione e la formula di Bayes

In questo paragrafo presentiamo alcuni importanti teoremi che derivano dall'applicazione del concetto di probabilità condizionata.

Il teorema di disintegrazione

Consideriamo lo spazio campionario Ω e supponiamo che $H_1, H_2, ..., H_n$ siano una **partizione** di Ω, cioè che $H_1, H_2, ..., H_n$ siano eventi di probabilità non nulla di Ω, a due a due incompatibili, la cui unione è Ω:

$$\Omega = H_1 \cup H_2 \cup ..... \cup H_n$$

Considerato un qualsiasi evento $A \subseteq \Omega$, gli insiemi $A \cap H_1, A \cap H_2, ..., A \cap H_n$ risultano a loro volta a due a due incompatibili e la loro unione è A (**Fig. 4**, in cui $n = 4$):

$$A = (A \cap H_1) \cup (A \cap H_2) \cup ... \cup (A \cap H_n)$$

Poiché gli eventi $A \cap H_1, A \cap H_2, ..., A \cap H_n$ sono disgiunti, abbiamo l'uguaglianza:

$$p(A) = p(A \cap H_1) + p(A \cap H_2) + ... + p(A \cap H_n)$$

che, in base alla formula delle probabilità composte, si può scrivere nella forma:

$$p(A) = p(A|H_1) \cdot p(H_1) + p(A|H_2) \cdot p(H_2) + ... + p(A|H_n) \cdot p(H_n)$$

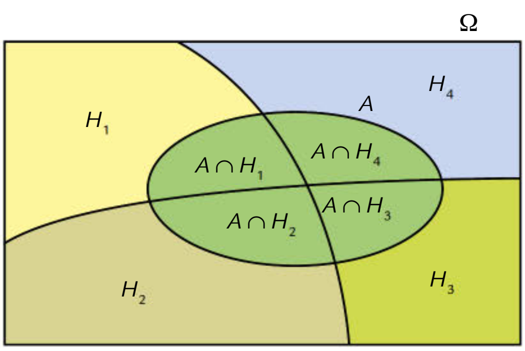


Figura 4

TEOREMA 5 | Teorema di disintegrazione (o della probabilità totale)

Sia $H_1, H_2, \ldots, H_n$ una collezione di eventi che forma una **partizione** dello spazio campionario. Allora, per ogni evento A, risulta:

$$p(A) = p(A|H_1) \cdot p(H_1) + p(A|H_2) \cdot p(H_2) + \ldots + p(A|H_n) \cdot p(H_n) \quad [11]$$

ATTENZIONE!

1. Il fatto che $H_1, H_2, \ldots, H_n$ definiscano una *partizione* dello spazio campionario significa che $H_1, H_2, \ldots, H_n$:
- hanno probabilità non nulle;
- sono a due a due incompatibili;
- la loro unione è lo spazio campionario.

2. La formula [11] è valida anche nel caso in cui gli eventi H_i, con $i = 1, \ldots, n$, non formano una partizione di Ω, purché abbiano probabilità non nulle, siano a due a due incompatibili e la loro unione contenga l'evento A.

È importante fare alcune osservazioni.

a. La più semplice applicazione del **Teorema 5** riguarda il caso in cui si considerano, come partizione dello spazio campionario, i due insiemi:

$$H_1 = B \quad \text{e} \quad H_2 = \overline{B}$$

dove B è un qualunque evento di probabilità non nulla; in tal caso la formula [11] diventa più semplicemente:

$$p(A) = p(A|B) \cdot p(B) + p(A|\overline{B}) \cdot p(\overline{B})$$

b. Gli eventi $H_1, H_2, \ldots, H_n$ che costituiscono una partizione dello spazio campionario prendono anche il nome di **alternative** perché se ne può verificare uno e uno solo. L'utilità del **Teorema 5** dipende dal fatto che spesso è difficile calcolare $p(A)$ mentre è più facile calcolare $p(A|H_i)$, perché ciò significa calcolare $p(A)$ con l'informazione *aggiuntiva* proveniente dal sapere che si è verificato H_i.

ESEMPIO Applicazione del teorema di disintegrazione

Un esperto di cavalli ritiene che il purosangue Furia sia più forte se corre con la pioggia. In particolare, l'esperto stima che Furia possa vincere con probabilità 10% in caso di tempo asciutto e con probabilità 25% in caso di pioggia. Il servizio meteorologico prevede, per l'ora della gara, tempo asciutto con una probabilità del 30%. Qual è la probabilità che Furia vinca la sua gara?

- **Formalizziamo il problema**

Indichiamo con A l'evento «il giorno della gara il tempo è asciutto» e con $\overline{A}$ l'evento «il giorno della gara piove»; sia inoltre V l'evento «Furia vince la gara».

Sappiamo che:

$p(V|A) = 0{,}1$ Furia vince con probabilità 10% se il tempo è asciutto

$p(V|\overline{A}) = 0{,}25$ Furia vince con probabilità 25% se piove

$p(A) = 0{,}3$ Il tempo è asciutto con probabilità uguale al 30%

Dobbiamo calcolare $p(V)$.

- **Calcoliamo le probabilità**

I due eventi A e $\overline{A}$ costituiscono un insieme di *alternative* (perché sono uno il complementare dell'altro) quindi possiamo applicare il **Teorema 5** per calcolare la probabilità richiesta:

$$p(V) = p(V|A) \cdot p(A) + p(V|\overline{A}) \cdot p(\overline{A}) =$$
$$= p(V|A) \cdot p(A) + p(V|\overline{A}) \cdot (1 - p(A)) =$$
$$= 0{,}1 \cdot 0{,}3 + 0{,}25 \cdot (1 - 0{,}3) = 0{,}205 = 20{,}5\%$$

VISUALIZZIAMO I CONCETTI **Problemi di probabilità e diagrammi ad albero**

Possiamo rappresentare il problema risolto nel precedente esempio con il diagramma ad albero in **Fig. 5**.

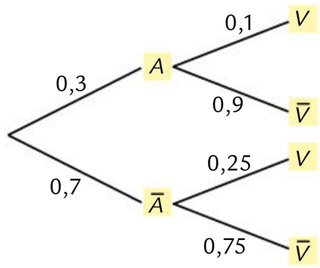


Figura 5

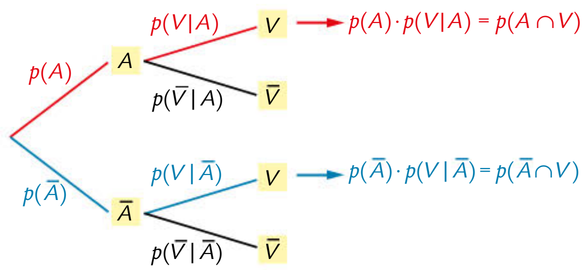


Figura 6

Nel diagramma in **Fig. 6** abbiamo esplicitato il significato delle probabilità rappresentate sui rami. È importante osservare quanto segue.

1. *La somma delle probabilità corrispondenti ai rami che escono da un medesimo nodo è sempre uguale a 1.* Ciò dipende dal fatto che i rami che escono da un nodo rappresentano eventi complementari.
2. *La probabilità dell'evento rappresentato da un cammino (cioè da una successione di rami) è il prodotto delle probabilità segnate sui rami da cui è costituito.* Questa proprietà non è altro che la «trasposizione» sul diagramma ad albero della regola delle probabilità composte; per esempio, il cammino rappresentato in rosso in **Fig. 6** rappresenta l'evento $A \cap V$ e risulta:

$$\underbrace{p(A \cap V)}_{\text{probabilità dell'evento rappresentato dal cammino in rosso}} = \underbrace{p(A)}_{\text{probabilità segnata sul primo ramo del cammino}} \cdot \underbrace{p(V|A)}_{\text{probabilità segnata sul secondo ramo del cammino}}$$

3. *La probabilità di un evento è la somma delle probabilità di tutti i cammini che conducono a esso.* Questa proprietà è la «trasposizione» sul diagramma ad albero della formula che esprime il teorema di disintegrazione. Per esempio, nel diagramma in **Fig. 6**, ci sono due cammini (quello colorato in rosso e quello colorato in blu) che conducono all'evento V; la probabilità di V è la somma delle probabilità dei due eventi rappresentati da questi due cammini; infatti:

$$p(V) = \underbrace{p(V|A) \cdot p(A)}_{\text{è la probabilità dell'evento } A \cap V\text{, cioè dell'evento rappresentato dal cammino in rosso}} + \underbrace{p(V|\overline{A}) \cdot p(\overline{A})}_{\text{è la probabilità dell'evento } \overline{A} \cap V\text{, cioè dell'evento rappresentato dal cammino in blu}}$$

La risoluzione di molti problemi di probabilità può essere facilitata rappresentandoli con opportuni diagrammi ad albero e tenendo presente le regole **1**, **2** e **3**.

La formula di Bayes

Abbiamo visto nel paragrafo precedente la formula delle probabilità composte:

$$p(A \cap B) = p(A) \cdot p(B|A) \qquad [12]$$

e la formula simmetrica che si ottiene invertendo i ruoli di A e B:

$$p(B \cap A) = p(B) \cdot p(A|B) \qquad [13]$$

Poiché $A \cap B = B \cap A$, dal confronto tra la [12] e la [13] si ottiene che:

$$p(A) \cdot p(B|A) = p(B) \cdot p(A|B)$$

Dividendo i due membri di quest'ultima uguaglianza per $p(A)$ (supposta diversa da zero), otteniamo la «formula di Bayes» che esprime un legame tra $p(B|A)$ e $p(A|B)$.

TEOREMA 6 | Formula di Bayes

Dati due eventi A e B, tali che $p(A) \neq 0$, risulta:

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

DALLA STORIA

La formula di Bayes è così chiamata in omaggio al matematico **Thomas Bayes** (1702-1761).

La formula di Bayes viene spesso applicata congiuntamente al teorema di disintegrazione, che si rivela utile per il calcolo di $p(A)$.

ESEMPIO Controllo qualità

Una fabbrica di sacchetti di carta ha due linee di produzione: la prima linea produce 500 pezzi al giorno, di cui il 2% (in media) difettosi; la seconda linea produce 300 pezzi al giorno di cui l'1% (in media) difettoso. Facendo un controllo a caso su tutta la produzione giornaliera si trova un sacchetto difettoso; qual è la probabilità che esso provenga dalla prima linea?

- **Formalizziamo il problema**

Consideriamo gli eventi:

L_1: «il sacchetto scelto proviene dalla linea 1»;
L_2: «il sacchetto scelto proviene dalla linea 2»;
D: «il sacchetto scelto è difettoso».

In base ai dati sappiamo che:

$$p(L_1) = \frac{500}{800} = \frac{5}{8} \qquad p(L_2) = \frac{300}{800} = \frac{3}{8}$$

$$p(D|L_1) = 0{,}02 \qquad p(D|L_2) = 0{,}01$$

Dobbiamo calcolare: $p(L_1|D)$.

- **Calcoliamo le probabilità**

In base alla formula di Bayes:

$$p(L_1|D) = \frac{p(D|L_1) \cdot p(L_1)}{p(D)}$$

L'unica probabilità che ci manca è $p(D)$, che possiamo calcolare applicando il teorema di disintegrazione:

$$p(D) = p(D|L_1) \cdot p(L_1) + p(D|L_2) \cdot p(L_2) =$$

$$= 0{,}02 \cdot \frac{5}{8} + 0{,}01 \cdot \frac{3}{8} = \frac{13}{800}$$

Pertanto:

$$p(L_1|D) = \frac{p(D|L_1) \cdot p(L_1)}{p(D)} = \frac{0{,}02 \cdot \frac{5}{8}}{\frac{13}{800}} = \frac{10}{13}$$

- **Commentiamo il risultato**

In questo caso possiamo ritrovare il risultato ottenuto calcolando direttamente la frequenza relativa corrispondente alla probabilità. I sacchetti difettosi prodotti (in media) dalla prima linea sono il 2% di 500, cioè 10, mentre i sacchetti difettosi prodotti dalla seconda linea sono (in media) l'1% di 300, cioè 3. Quindi i sacchetti difettosi prodotti (in media) dalla fabbrica ogni giorno sono in tutto $10 + 3 = 13$ e di questi 10 provengono dalla prima linea di produzione; rispetto al totale dei pezzi difettosi quelli provenienti dalla prima linea sono dunque $\frac{10}{13}$.

ESEMPIO Test clinico

Un test clinico è efficace nel 98% dei casi nell'individuare una data malattia nelle persone effettivamente malate (vale a dire: se una persona malata si sottopone al test, nel 98% dei casi il test risulta positivo). Il test tuttavia può generare anche dei «falsi positivi», cioè dare esito positivo nell'1% delle persone sane che si sottopongono al test. È noto inoltre che la malattia colpisce lo 0,5% della popolazione. Se una persona risulta positiva al test, qual è la probabilità che sia effettivamente malata?

- **Formalizziamo il problema**

Consideriamo i due eventi:

M: «la persona è malata»

T^+: «il test risulta positivo»

I dati forniti dal problema si possono formalizzare così:

$$p(T^+|M) = 0{,}98 \qquad p(T^+|\overline{M}) = 0{,}01 \qquad p(M) = 0{,}005$$

L'obiettivo è calcolare $p(M|T^+)$.

- **Calcoliamo le probabilità**

Per la formula di Bayes:

$$p(M|T^+) = \frac{p(T^+|M) \cdot p(M)}{p(T^+)}$$

L'unica probabilità che ci manca è $p(T^+)$, che possiamo calcolare applicando il teorema di disintegrazione:

$$\begin{aligned} p(T^+) &= p(T^+|M) \cdot p(M) + p(T^+|\overline{M}) \cdot p(\overline{M}) = \\ &= p(T^+|M) \cdot p(M) + p(T^+|\overline{M}) \cdot (1 - p(M)) = \\ &= 0{,}98 \cdot 0{,}005 + 0{,}01 \cdot (1 - 0{,}005) = \\ &= 0{,}01485 \end{aligned}$$

In definitiva:

$$\begin{aligned} p(M|T^+) &= \frac{p(T^+|M) \cdot p(M)}{p(T^+)} = \\ &= \frac{0{,}98 \cdot 0{,}005}{0{,}01485} \simeq 0{,}33 \end{aligned}$$

- **Commentiamo il risultato**

Il risultato ci dice quindi che una persona risultata positiva al test ha soltanto il 33% circa di probabilità di risultare effettivamente malata! E questo nonostante il test abbia alta sensibilità: infatti individua la malattia nel 98% dei malati. Questo apparente paradosso dipende dal fatto che la malattia considerata è a bassa incidenza: la probabilità che un individuo della popolazione sia malato è infatti solo lo 0,5%, cioè solo 1 individuo su 200 risulta malato. La maggior parte dei positivi risultano falsi positivi proprio per questo motivo: perché la quasi totalità della popolazione è sana.

Per esempio, supponiamo che 20 000 persone si sottopongano al test; ci aspettiamo allora che i malati siano lo 0,5%, cioè 100, e di questi il 98% risulteranno *veri positivi* al test, cioè 98. Nei restanti 19 900 sani, i *falsi positivi* saranno l'1%, cioè 199: quindi in totale, tra i $98 + 199 = 297$ *positivi* al test, soltanto 98 sono veramente malati, appunto circa il 33%.

COLLEGHIAMO I CONCETTI — Vari metodi risolutivi per problemi di probabilità

Fin qui abbiamo visto principalmente due metodi per il calcolo delle probabilità:

1. l'utilizzo della definizione classica e le regole del calcolo combinatorio;
2. l'utilizzo dei teoremi studiati, in particolare la regola delle probabilità composte, il teorema di disintegrazione e la formula di Bayes; quest'ultimo metodo può essere applicato implicitamente costruendo opportuni diagrammi ad albero e utilizzando le regole sui diagrammi ad albero che equivalgono ai teoremi.

A seconda dello specifico problema, può rivelarsi più opportuno l'utilizzo dell'uno o dell'altro metodo; in altri casi la risoluzione può avvenire secondo vari approcci, come per esempio nel seguente problema.

◆ PROBLEMA
Un'urna contiene 10 palline: 2 verdi, 3 gialle e 5 rosse. Qual è la probabilità che estraendo successivamente due palline, senza reimmissione, almeno una sia gialla?

Siano E l'evento «almeno una delle due palline estratte è gialla» e G_i l'evento «esce una pallina gialla alla i-esima estrazione», con $i = 1, 2$.

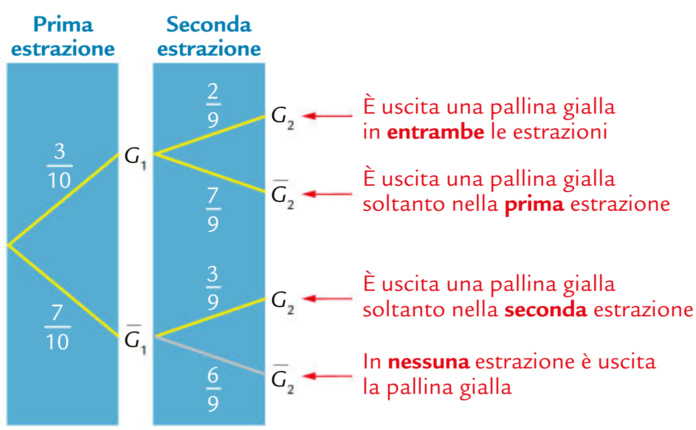
▶ 1° modo: passaggio all'evento contrario e calcolo combinatorio
Sia E l'evento «almeno una delle due palline estratte sia gialla»; abbiamo che:

$$p(E) \underbrace{=}_{\substack{\text{passaggio}\\ \text{all'evento}\\ \text{contrario}}} 1 - p(\underbrace{\overline{E}}_{\substack{\text{nessuna delle}\\ \text{palline estratte}\\ \text{è gialla}}}) = 1 - \frac{\overbrace{7\cdot 6}^{\substack{\text{casi favorevoli:}\\ \text{disposizioni senza ripetizione}\\ \text{di 2 palline scelte tra quelle non gialle}}}}{\underbrace{10\cdot 9}_{\substack{\text{casi possibili:}\\ \text{disposizioni senza ripetizione}\\ \text{di 2 palline scelte tra le 10}}}} = 1 - \frac{7}{15} = \frac{8}{15}$$

▶ 2° modo: passaggio all'evento contrario e regola delle probabilità composte

$$p(E) = 1 - p(\overline{E}) = 1 - p(\overline{G_1} \cap \overline{G_2}) = 1 - \underbrace{p(\overline{G_1})p(\overline{G_2}|\overline{G_1})}_{\substack{\text{formula delle}\\ \text{probabilità composte}}} = 1 - \frac{7}{10}\cdot\frac{6}{9} = \frac{8}{15}$$

▶ 3° modo: regola delle probabilità composte, senza calcolo combinatorio
Costruiamo un opportuno diagramma ad albero.



L'evento E si realizza nei primi tre casi indicati con la freccia; la probabilità di E è perciò la somma delle probabilità dei tre cammini colorati in giallo:

$$p(E) = \frac{3}{10}\cdot\frac{2}{9} + \frac{3}{10}\cdot\frac{7}{9} + \frac{7}{10}\cdot\frac{3}{9} = \frac{8}{15}$$

Esercizi p. 190

6. Le varie definizioni di probabilità e l'approccio assiomatico

L'approccio frequentista e l'approccio soggettivo

La definizione classica di probabilità risulta di fatto inapplicabile in molti casi concreti. Discutiamo insieme alcuni problemi.

◆ **PROBLEMA 1**

Immagina di essere il titolare di una società di assicurazioni. Devi stabilire quanto far pagare il rischio grandine ai viticoltori che ne richiedono la copertura, nella zona di Franciacorta, area del bresciano nota per la qualità dei vini che vi si producono. Il premio assicurativo annuo dovrà comprendere, ovviamente, i costi di agenzia, le tasse, il profitto della società, ma sarà principalmente determinato da un numero: *la probabilità che grandini prima della vendemmia in quella particolare area geografica*. Come si può calcolare un tale valore di probabilità?

Le previsioni del tempo sono attendibili, in vario grado, per periodi non superiori ai cinque giorni; l'unica possibilità è dunque raccogliere i dati storici disponibili, verificare quante volte ha grandinato nel periodo oggetto di osservazione e, ragionevolmente, assumere come *probabilità il rapporto fra il numero degli anni con almeno una grandinata disastrosa e il numero degli anni presi in esame*.

◆ **PROBLEMA 2**

In certe epoche storiche o in certe nazioni si trova una maggiore propensione a scommettere su avvenimenti di ogni tipo, tanto da far diventare nota la figura del *bookmaker*. Immagina di essere un bookmaker e di avere una proposta di scommessa di questo tipo: «il prossimo papa sarà un cardinale della Chiesa sudamericana». Come puoi valutare la probabilità dell'evento (senza la cui valutazione non sai quali quote assegnare in caso di vincita)?

In questo caso non è possibile né l'approccio classico né quello statistico dell'esempio precedente. Puoi solo consultare un vero esperto della materia che, in base alle sue informazioni, potrà darti una sua previsione in termini percentuali; tale previsione, in mancanza d'altro, diventerà la valutazione della probabilità dell'evento.
Riflettiamo sui due problemi che abbiamo esaminato.

1. Nel primo caso (problema dell'assicuratore) non è possibile l'impostazione con strumenti matematici classici, ma si può ricorrere al metodo statistico per dare una valutazione della probabilità cercata.
2. Nel secondo caso (problema del bookmaker) non è possibile applicare né il metodo classico né quello statistico: il grado di fiducia può essere assegnato solo da una valutazione che possiamo definire *soggettiva* (senza dare alcuna connotazione negativa al termine).

Per descrivere questi ultimi due tipi di approccio alla valutazione della probabilità di un evento sono state formulate le seguenti definizioni alternative di probabilità.

Con GeoGebra
Simulazione del gioco «Gira la ruota!»

DEFINIZIONE | Definizione frequentista di probabilità

Consideriamo un evento, relativo a un esperimento ripetibile nelle medesime condizioni quante volte si vuole. Si definisce **probabilità** dell'evento la sua **frequenza relativa**, osservata in un numero sufficientemente grande di esperimenti.

PER LA PRECISIONE

La definizione frequentista data a fianco è in realtà una versione «semplificata» della definizione introdotta da **Richard von Mises** nel 1928 e che qui non abbiamo potuto proporre perché non abbiamo ancora studiato il concetto di limite (che vedrai nel proseguimento dei tuoi studi).

DEFINIZIONE | Definizione soggettiva di probabilità

Dato un evento E, si chiama **probabilità** di E un numero compreso tra 0 e 1, che un individuo assegna a E come misura del **grado di fiducia** che egli ripone nel verificarsi di E. La probabilità che l'individuo assegna a E può pensarsi come il prezzo p che è disposto a pagare per ricevere l'importo 1 se l'evento E si verifica e l'importo 0 se l'evento E non si verifica.

L'assegnazione di una probabilità secondo l'approccio soggettivo è vincolata dal cosiddetto **principio di coerenza**, in base al quale l'individuo che è disposto a scommettere la somma *p* sul realizzarsi di un dato evento deve essere disposto a scambiare il proprio ruolo con quello del banco. Per esempio, se Paolo ritiene uguale al 20% la probabilità che una data squadra vinca, significa che ritiene equo pagare 1 euro per vincerne 5 nel caso in cui l'evento si verifichi; in base al principio di coerenza, però, Paolo deve anche essere disposto a pagare 5 euro a un'altra persona che scommetta sulla vittoria della squadra, versandogli 1 euro.

Le critiche alle varie definizioni di probabilità e l'approccio assiomatico

Riflettendo sulle tre definizioni di probabilità che abbiamo presentato (classica, frequentista e soggettiva), è facile rendersi conto che ciascuna presenta degli inconvenienti.

- La **definizione classica** richiede che lo spazio degli eventi sia *finito* e che tutti gli eventi elementari siano *equipossibili* (ovvero *equiprobabili*), il che porta a una notevole *restrizione* delle situazioni cui è applicabile; inoltre parlare di eventi equiprobabili nella definizione stessa di probabilità genera una circolarità nella definizione che pone problemi logici.
- La **definizione frequentista** richiede la *ripetibilità* di un esperimento *nelle stesse condizioni* un gran numero di volte, ma ciò non sempre è possibile (pensa per esempio a una partita di calcio); inoltre non è precisato esattamente quante volte deve essere ripetuto l'esperimento per avere una valutazione attendibile della probabilità, quindi la valutazione potrà differire in relazione al numero di esperimenti che vengono effettuati.
- La **definizione soggettiva**, pur non consentendo assegnazioni di probabilità arbitrarie (perché vincolate dal *principio di coerenza*), può però portare, da parte di individui diversi, ad assegnare a uno stesso evento probabilità diverse, in base alle diverse informazioni in possesso degli individui medesimi.

Queste considerazioni fanno emergere chiaramente che il problema di formulare una *definizione* unificante di probabilità di un evento, che indichi *esplicitamente* come calcolarla e che possa adattarsi a tutte le variegate situazioni in cui interviene il concetto di probabilità, appare di difficile soluzione.
Le difficoltà sono insite nella questione e sono testimoniate, anche storicamente, dal fatto che il problema ha occupato a lungo i matematici ed è stato risolto solo in tempi relativamente recenti, nel 1933, dal matematico russo **Kolmogorov**. Ma qual è stata la chiave che ha consentito di superare le difficoltà? L'idea vincente di Kolmogorov è stata di affrontare il calcolo della probabilità secondo un approccio *assiomatico* che, come la geometria euclidea, si sviluppa a partire da un esiguo numero di assiomi. In questa impostazione il concetto di probabilità diventa un *concetto primitivo* (come punto, retta, piano in geometria), che non viene definito *esplicitamente*, ma che è *implicitamente* descritto dagli assiomi assunti a fondamento della teoria stessa. L'impostazione assiomatica permette cioè a Kolmogorov di non esplicitare esattamente *come* valutare la probabilità (lasciando quindi la libertà di seguire l'approccio più adatto al caso in esame), limitandosi a indicare quali sono le regole formali che una misura di probabilità deve soddisfare per poter essere dichiarata tale. Queste regole, assunte appunto come *assiomi*, sono quelle caratteristiche che intuitivamente siamo portati ad attribuire a una misura di probabilità di un evento (e che abbiamo visto in particolare essere soddisfatte nella definizione classica):

- la probabilità di un evento deve essere un numero compreso tra 0 e 1;
- la probabilità dell'evento certo deve essere la massima possibile, cioè 1;
- la probabilità dell'unione di due eventi incompatibili deve essere la somma delle probabilità dei due singoli eventi.

Figura 7 Kolmogorov risolve nel 1933 il problema dei fondamenti del calcolo della probabilità, con un approccio assiomatico che rivoluziona il modo di intendere la materia.

ATTENZIONE!
I teoremi sul calcolo delle probabilità che abbiamo dimostrato relativamente alla definizione classica possono essere dimostrati anche soltanto sulla base degli assiomi. Perciò tutti i teoremi sono applicabili, indipendentemente dall'approccio (classico frequentista, soggettivo) utilizzato per valutare la probabilità dei singoli eventi.

ASSIOMI | Assiomi di una misura di probabilità

Dato uno spazio campionario Ω, la probabilità $p(A)$ di un evento $A \subseteq \Omega$ è un numero reale tale che:

a. $0 \leq p(A) \leq 1$;

b. se $A = \Omega$, allora $p(A) = 1$;

c. se A e B sono due eventi incompatibili, allora:

$$p(A \cup B) = p(A) + p(B)$$

Nell'approccio assiomatico, dunque, quando si deve attribuire una probabilità nei singoli casi concreti, cioè stabilire il *valore numerico* che la esprime, si è liberi di seguire l'approccio più consono al caso in esame (classico, frequentista o soggettivo), con l'unico obbligo di non violare gli assiomi.

ESEMPIO Dado truccato

Consideriamo un dado cubico con le facce numerate da 1 a 6. Il dado è truccato in modo tale che la probabilità di ottenere ciascun numero è proporzionale al numero stesso. Si lancia il dado una volta; qual è la probabilità dell'evento: «esce un numero pari»?

- **Analisi preliminare**

Lo spazio campionario dell'evento è $\Omega = \{1, 2, 3, 4, 5, 6\}$, come nel caso di un dado regolare, ma ora gli eventi elementari **non** sono *equiprobabili*. Come possiamo assegnare una probabilità a ciascuno degli eventi elementari di Ω? Dobbiamo tenere conto *sia* delle informazioni fornite dal testo del problema, *sia* degli assiomi introdotti poc'anzi!
Iniziamo a concentrarci sulle informazioni fornite dal testo: sappiamo che la probabilità di ottenere ciascun numero è *proporzionale* al numero stesso; quindi le probabilità degli eventi elementari possono essere espresse in funzione della costante di proporzionalità k come indicato nella seguente tabella.

Numero	1	2	3	4	5	6
Probabilità di ottenerlo	k	$2k$	$3k$	$4k$	$5k$	$6k$

Ora ricordiamoci che, in base agli assiomi, la somma delle probabilità di tutti gli eventi elementari deve essere 1, quindi:

$$k + 2k + 3k + 4k + 5k + 6k = 1$$

Risolvendo questa equazione, otteniamo:

$$k = \frac{1}{21}$$

Quindi le probabilità degli eventi elementari devono essere le seguenti:

Numero	1	2	3	4	5	6
Probabilità di ottenerlo	$\frac{1}{21}$	$\frac{2}{21}$	$\frac{3}{21}$	$\frac{4}{21}$	$\frac{5}{21}$	$\frac{6}{21}$

- **Determiniamo la probabilità dell'evento cercato**

Osserviamo che l'evento E: «esce un numero pari» è l'unione dei tre eventi elementari E_1: «esce il numero 2», E_2: «esce il numero 4» ed E_3: «esce il numero 6».
Poiché i tre eventi elementari sono *incompatibili*, la probabilità dell'evento E è la somma delle probabilità di E_1, E_2 e E_3, dunque è uguale a:

$$\frac{2}{21} + \frac{4}{21} + \frac{6}{21} = \frac{12}{21} = \frac{4}{7}$$

La legge dei grandi numeri

Mediante l'approccio assiomatico è anche possibile *dimostrare* un importante teorema che getta un ponte tra la definizione classica e la definizione frequentista, noto come *legge dei grandi numeri*. A grandi linee, questo teorema esprime quanto segue.

TEOREMA | Legge dei grandi numeri

Supponiamo che a un evento E sia attribuibile una probabilità sia secondo la definizione classica, sia secondo quella frequentista. Allora, in una serie di prove ripetute, l'evento E si manifesta con una frequenza relativa che tende, al crescere del numero delle prove, a coincidere con il valore teorico della sua probabilità.

Con GeoGebra
La legge dei
grandi numeri

Questo teorema è alla base di un importante metodo per valutare la probabilità di un evento, utilizzato in tutti i casi in cui l'approccio puramente matematico non è conosciuto o risulta troppo complesso, il cosiddetto **metodo Monte Carlo**. Esso consiste sostanzialmente nell'eseguire al computer un gran numero di simulazioni del fenomeno in esame e di calcolare la frequenza relativa dell'evento di cui si vuole calcolare la probabilità: in forza della legge dei grandi numeri, all'aumentare del numero di simulazioni si ottengono approssimazioni sempre migliori della probabilità cercata. Vedremo alcuni semplici esempi di applicazione di questo metodo nella scheda di approfondimento.

ESEMPIO

Alla fine del Settecento George Ledere lanciò 4040 volte una moneta, ottenendo «testa» 2048 volte, con una frequenza uguale a $\dfrac{2048}{4040} = 0{,}5069\ldots$

Un esperimento simile fu ripetuto un secolo dopo da Karl Pearson, che lanciò una moneta 24 000 volte ottenendo «testa» 12 012 volte, con una frequenza uguale a $\dfrac{12\,012}{24\,000} = 0{,}5005$, valore ancora più vicino alla probabilità teorica di ottenere «testa», che è uguale a $\dfrac{1}{2}$.

Esercizi p. 197

Spazio campionario ed evento

Si chiama **spazio campionario** (e si indica usualmente con Ω) l'insieme di tutti i possibili esiti di un esperimento aleatorio, cioè di un esperimento il cui risultato non può essere previsto con certezza. Si chiama **evento** ogni sottoinsieme dello spazio campionario.

ESEMPIO

Nell'esperimento che consiste nel lancio di un dado $\Omega = \{1, 2, 3, 4, 5, 6\}$. L'insieme $E = \{1, 3, 5\} \subset \Omega$ è l'evento che corrisponde all'uscita di un numero dispari.

Definizione classica

Dato uno spazio campionario Ω formato da n elementi (detti «casi possibili») che hanno tutti la stessa possibilità di verificarsi e un evento $A \subseteq \Omega$ formato da k elementi (detti «casi favorevoli»), la probabilità di A è il rapporto tra il numero di casi favorevoli e il numero di casi possibili:

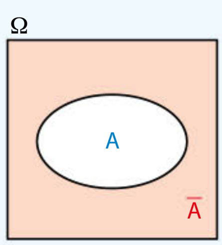
$p(A) = \dfrac{k}{n}$ Risulta sempre: $0 \leq p(A) \leq 1$

Operazioni tra eventi

Operazioni che è possibile eseguire tra eventi (analoghe alle operazioni insiemistiche).

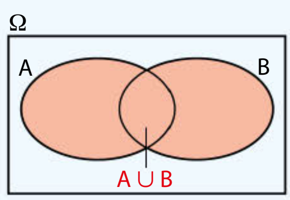
Evento contrario

L'evento $\overline{A}$, **contrario** di A, è l'evento che si verifica quando **non** si verifica A.



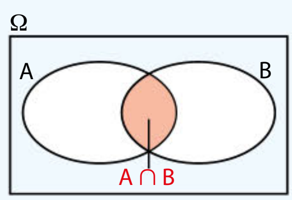
Evento unione

L'evento $A \cup B$, **unione** di A e B, è l'evento che si verifica quando si verifica A **oppure** B.



Evento intersezione

L'evento $A \cap B$, **intersezione** di A e B, è l'evento che si verifica quando si verificano **contemporaneamente** A e B.



Probabilità dell'evento contrario

Risulta:

$p(\overline{A}) = 1 - p(A)$

Legame tra la probabilità dell'evento unione e la probabilità dell'evento intersezione

Risulta:

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

Probabilità condizionata

La probabilità condizionata dell'evento A dato l'evento B è:

$$p(A|B)=\frac{p(A\cap B)}{p(B)} \text{ con } p(B)\neq 0$$

Analogamente:

$$p(B|A)=\frac{p(A\cap B)}{p(A)} \text{ con } p(A)\neq 0$$

Eventi indipendenti

Due eventi A e B per cui $p(A|B)=p(A)$ oppure, in modo equivalente, per cui $p(B|A)=p(B)$.

Probabilità composte

Dati due eventi A e B valgono le relazioni:

$p(A\cap B)=p(B)\cdot p(A|B)$

$p(A\cap B)=p(A)\cdot p(B|A)$

In particolare quindi:

$p(A)\cdot p(B|A)=p(B)\cdot p(A|B)$

Regola del prodotto

Due eventi A e B sono *indipendenti* se e solo se risulta:

$p(A\cap B)=p(A)\cdot p(B)$

Teorema di Bayes

Dati due eventi A e B tali che $p(A)\neq 0$ risulta:

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

Prove ripetute

Consideriamo n prove ripetute e indipendenti di uno stesso esperimento. Per ogni prova, sia p la probabilità che l'esperimento dia esito positivo (si usa dire che si è verificato un «successo). La probabilità di avere esattamente k successi in n prove è uguale a:

$$\binom{n}{k} p^k (1-p)^{n-k}$$

Teorema di disintegrazione

Se H_1, H_2, …, H_n è una collezione di eventi che forma una partizione dello spazio campionario Ω, allora per ogni evento $A\subseteq\Omega$ risulta:

$$p(A) = p(A\cap H_1)+\ldots+p(A\cap H_n) = p(A|H_1)p(H_1)+\ldots+p(A|H_n)p(H_n)$$

Unità 4

Esercizi

1. Richiami di calcolo delle probabilità

Teoria p. 138

Il linguaggio degli eventi

Test

Nei test 1-6 fai riferimento alla seguente situazione. In un giardino ci sono rose e tulipani; sia le rose sia i tulipani sono di due tipi: o rossi o gialli. Si sceglie a caso un fiore e si indica con S l'evento «il fiore è giallo» e con R l'evento «il fiore è una rosa».

1 L'evento «il fiore è una rosa gialla» è rappresentato da:

A $R \cap S$ B $R \cup S$ C $\overline{R}$ D $\overline{S}$

2 L'evento «il fiore è rosso» è rappresentato da:

A $R \cap S$ B $R \cup S$ C $\overline{R}$ D $\overline{S}$

3 L'evento «il fiore è un tulipano rosso» è rappresentato da:

A $\overline{R} \cap S$ B $R \cup \overline{S}$ C $\overline{R} \cap \overline{S}$ D $\overline{R} \cup \overline{S}$

4 L'evento «il fiore è una rosa rossa» è rappresentato da:

A $R \cap \overline{S}$ B $R \cup \overline{S}$ C $\overline{R} \cap S$ D $\overline{R} \cup S$

5 L'evento «il fiore è giallo o una rosa» è rappresentato da:

A $R \cap \overline{S}$ B $R \cup \overline{S}$ C $R \cap S$ D $R \cup S$

6 Quale delle seguenti è una coppia di eventi incompatibili?

A «Il fiore è rosso» e «il fiore è un tulipano»
B «Il fiore è una rosa» e «il fiore è un tulipano»
C «il fiore è giallo» e «il fiore è una rosa»
D «il fiore è rosso» e «il fiore è una rosa»

7 **Vero o falso?**

Fai riferimento all'esperimento che consiste nel lancio di un dado regolare a sei facce.

a. l'evento «esce un numero dispari» è elementare V F
b. l'evento «esce un numero maggiore di 5» è elementare V F
c. l'evento «esce un numero minore di 5» è elementare V F
d. l'evento «esce un numero maggiore di 8» è impossibile V F
e. l'evento «esce un numero minore di 8» è certo V F
f. l'evento «esce un numero primo» è impossibile V F
g. l'evento «esce un numero primo o divisibile per 2» è certo V F

[3 affermazioni vere e 4 false]

Per ciascuno dei seguenti esperimenti aleatori, individua uno spazio campionario.

8 L'estrazione, per due volte consecutive con reimmissione, di una pallina da un'urna che ne contiene 2 bianche e 1 rossa.

9 Il lancio di una moneta, ripetuto successivamente per due volte.

10 La scelta, per due volte successive, di una carta fra i quattro assi (senza reimmissione del primo asso estratto).

11 La scelta a caso di un numero reale compreso tra 0 e 1, inclusi 0 e 1.

12 Il lancio di un dado a sei facce.

13 Il lancio di due dadi a sei facce e la registrazione della somma dei due numeri ottenuti.

14 L'estrazione di un numero al lotto.

15 L'uscita di un numero alla roulette.

16 Il lancio successivo di una moneta e la registrazione del numero di volte necessario per ottenere per la prima volta «croce».

17 Un'urna contiene quattro biglie numerate da 1 a 4. Le biglie di numero pari sono rosse, quelle di numero dispari sono verdi. Per ciascuno dei seguenti esperimenti aleatori, individua uno spazio campionario.

a. Si estrae una biglia e si annota il numero ottenuto.
b. Si estrae una biglia e si annota il colore della biglia.
c. Si estraggono, contemporaneamente, due biglie e si annotano i due numeri ottenuti.
d. Si estraggono due biglie, una dopo l'altra, dopo aver rimesso la prima nell'urna, e si annotano i due numeri ottenuti.
e. Si estraggono due biglie, una dopo l'altra, senza rimettere la prima nell'urna, e si annotano i due numeri ottenuti.

18 Una moneta regolare viene lanciata due volte.

a. Determina lo spazio campionario.
b. Barbara vince se esce «croce» al primo lancio, mentre Stefano vince se esce «testa» al secondo lancio. Determina i sottoinsiemi dello spazio campionario che rappresentano i seguenti eventi:

i. Barbara vince
ii. Stefano perde
iii. Barbara e Stefano vincono entrambi
iv. almeno uno dei due vince
v. non vincono né Barbara né Stefano
vi. Barbara vince e Stefano perde

19 ESERCIZIO GUIDATO

Considera l'esperimento che consiste nei lanci successivi, per cinque volte, di una moneta.

a. Descrivi qual è lo spazio campionario (non è richiesto di elencare tutti gli esiti esplicitamente).
b. Considera l'evento «si ottengono almeno tre 'testa'» ed esprimilo, a parole, come unione di eventi incompatibili.

---

a. Lo spazio campionario Ω è l'insieme di tutte le 5-ple della forma $(x_1, x_2, x_3, x_4, x_5)$ dove $x_1, x_2, x_3, x_4, x_5 \in \{..., ...\}$.
b. L'evento «si ottengono almeno tre 'testa'» equivale a «si ottengono esattamente tre 'testa' o esattamente .... o esattamente ....», quindi è l'unione dei seguenti tre eventi a due a due incompatibili: «si ottengono esattamente ..... 'testa'», «si ottengono esattamente ..... 'testa'» e «si ottengono esattamente ..... 'testa'».

20 Considera l'esperimento che consiste nell'estrarre successivamente, per quattro volte, una carta da un mazzo di 52 (eseguendo ogni estrazione senza rimettere nel mazzo la carta estratta).

a. Descrivi qual è lo spazio campionario (non è richiesto di elencare tutti gli esiti esplicitamente).
b. Considera l'evento «si estraggono al massimo due assi» ed esprimilo, a parole, come unione di eventi incompatibili.

21 Da un'urna contenente 6 biglie rosse, 4 verdi e 2 gialle se ne estrae una a caso. Sia A l'evento «la biglia estratta è rossa» e B l'evento «la biglia estratta è verde». Esprimi a parole i seguenti eventi:

a. $A \cap B$ **b.** $A \cup B$ **c.** $\overline{A} \cap \overline{B}$ **d.** $A \cup \overline{B}$

22 Siano A, B e C tre eventi. Esprimi in notazione insiemistica i seguenti eventi descritti a parole.

a. Si verifica l'evento A o l'evento B.
b. Si verificano sia l'evento B sia l'evento C.
c. Non si verifica l'evento B.
d. Non si verifica né l'evento A né l'evento C.
e. Non si verifica nessuno dei tre eventi.
f. Si verifica almeno uno dei tre eventi.
g. Si verifica soltanto l'evento A.
h. Si verificano esattamente due dei tre eventi.

Proprietà della probabilità

23 Per ciascuno dei seguenti numeri, stabilisci se può rappresentare la probabilità di un evento, giustificando la risposta: $\frac{3}{4}$; 0; -1; $\frac{7}{6}$; 1; 15%; 0,25; $\frac{8}{9}$; 1,3.

24 Determina per quali valori di k l'espressione $\frac{k+5}{3}$ può rappresentare la probabilità di un evento. $[-5 \leq k \leq -2]$

25 Determina per quali valori di k l'espressione $\frac{3-2k}{4}$ può rappresentare la probabilità di un evento. $\left[-\frac{1}{2} \leq k \leq \frac{3}{2}\right]$

26 Determina per quali valori di k l'espressione $\frac{k+2}{k+1}$ può rappresentare la probabilità di un evento. $[k \leq -2]$

27 Determina per quali valori di k l'espressione $\frac{k}{4-k}$ può rappresentare la probabilità di un evento. $[0 \leq k \leq 2]$

2. Valutazione della probabilità secondo la definizione classica

Teoria p. 141

Esercizi introduttivi

Test

28 Si estrae a caso una carta da un mazzo di 52; qual è la probabilità di estrarre una carta di picche?

A $\frac{3}{8}$ B $\frac{3}{52}$ C $\frac{1}{4}$ D $\frac{5}{26}$

29 La probabilità di estrarre uno dei quattro Assi da un mazzo di carte è $\frac{1}{13}$. Quante carte ha quel mazzo?

A 32 B 40 C 50 D 52

30 Si lanciano due dadi regolari a sei facce a caso; qual è la probabilità di ottenere un doppio 6?

A $\frac{1}{2}$ B $\frac{1}{3}$ C $\frac{1}{18}$ D $\frac{1}{36}$

31 **Caccia all'errore.** Barbara è incinta di due gemelli e sa che sono dizigoti. Barbara ritiene che la probabilità che almeno uno dei due sia femmina sia di $\frac{2}{3}$, in base al seguente ragionamento: il numero di femmine che può nascere può essere 0 (se entrambi i gemelli sono maschi), oppure 1 oppure 2. Abbiamo dunque 3 casi possibili, di cui 2 favorevoli (la nascita di 1 femmina o di 2 femmine).
Ritieni che il ragionamento di Barbara sia corretto?

Applicazioni della definizione classica

32 ESERCIZIO SVOLTO

Qual è la probabilità di estrarre a caso un numero naturale di due cifre e avere un quadrato perfetto?

I numeri naturali di due cifre sono 90. I quadrati di due cifre sono: 16, 25, 36, 49, 64, 81. Perciò si hanno 6 casi favorevoli e 90 casi possibili. La probabilità richiesta è: $\frac{6}{90} = \frac{1}{15}$.

33 Dal 1973 al 2004 sono stati assegnati dal Ministero delle finanze 80 milioni di codici fiscali. Di questi, 14 000 non sono unici, cioè sono stati dati a più di un individuo. Qual è la probabilità di essere lo sfortunato cui viene assegnato un codice doppione? [0,0175%]

34 Uno scaffale contiene libri di 5 diverse case editrici: 20 della De Agostini, 15 della Mondadori, 10 della Rizzoli, 8 della Zanichelli e 5 della Garzanti. Si prende a caso un libro, senza guardare. Qual è la probabilità che il libro non sia né della Mondadori né della Garzanti? $\left[\frac{19}{29}\right]$

35 Dal sacchetto della tombola si estrae un numero. Calcola la probabilità di estrarre un numero divisibile per 10 ma non divisibile per 30. $\left[\frac{1}{15}\right]$

36 I nati dall'1 al 23 settembre sono del segno della Vergine; chi nasce fra il 24 settembre e il 23 ottobre è del segno della Bilancia. Presa a caso una persona nata a settembre, qual è la probabilità che sia del segno Bilancia? $\left[\frac{7}{30}\right]$

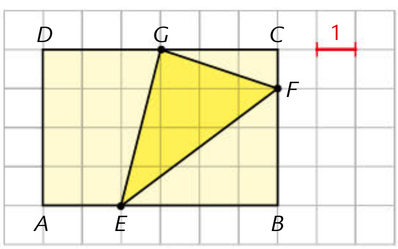
37 A una cena fra medici partecipano tre chirurghi, 2 pediatri e 4 internisti. Il cameriere sceglie a caso uno dei medici e ipotizza che sia un chirurgo. Qual è la probabilità che si sbagli? $\left[\frac{2}{3}\right]$

38 Un barista sorteggia fra i suoi clienti un panettone del valore di 50 euro. Prepara 90 tagliandi numerati da 1 a 90; vince il tagliando che riporta lo stesso numero del primo estratto al gioco del lotto sulla ruota di Milano del sabato prima di Natale. Il barista vende solo 30 biglietti. Qual è la probabilità che il panettone rimanga non assegnato? $\left[\frac{2}{3}\right]$

39 Un sacchetto contiene dei biglietti, su ciascuno dei quali è trascritto un numero primo minore di 100. Si estrae un biglietto a caso; qual è la probabilità che esso sia pari? $\left[\frac{1}{25}\right]$

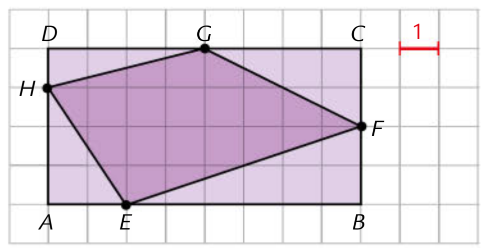
Collegamenti **Probabilità e geometria**

40 Scelto a caso un punto P del rettangolo $ABCD$ in figura, qual è la probabilità che esso appartenga al triangolo EFG? Supponi che i quadrati che formano la griglia abbiano lati di misura 1.



$\left[\frac{13}{48}\right]$

41 Scelto a caso un punto P del rettangolo $ABCD$ in figura, qual è la probabilità che esso appartenga al quadrilatero $EFGH$? Supponi che i quadrati che formano la griglia abbiano lati di misura 1.



$\left[\frac{17}{32}\right]$

---

42 Considera un cerchio di raggio r e un quadrato $ABCD$ inscritto nel cerchio. Scelto a caso un punto all'interno del cerchio, qual è la probabilità che esso cada all'interno del quadrato $ABCD$? $\left[\frac{2}{\pi}\right]$

43 Considera un cerchio di raggio 2. Scelto a caso un punto all'interno del cerchio, qual è la probabilità che disti più di 1 dal centro del cerchio? $\left[\frac{3}{4}\right]$

44 Considera nel piano cartesiano l'insieme dei punti di coordinate (x, y), essendo x e y due numeri interi con $0 \le x \le 5$ e $0 \le y \le 5$. Scegliendo a caso un punto appartenente a questo insieme, qual è la probabilità che appartenga alla bisettrice del primo e del terzo quadrante? $\left[\frac{1}{6}\right]$

45 Considera nel piano cartesiano l'insieme dei punti di coordinate (x, y), essendo x e y due numeri interi con $-2 \le x \le 2$ e $-2 \le y \le 2$. Scegliendo a caso un punto appartenente a questo insieme, qual è la probabilità che appartenga alla retta di equazione $y = 2x$? $\left[\frac{3}{25}\right]$

Problemi con equazioni

46 **ESERCIZIO GUIDATO**

In un'urna ci sono 12 biglie rosse e n biglie nere. Estraendo a caso una biglia dalla scatola, la probabilità che essa sia nera è 0,4. Quante biglie nere ci sono nell'urna?

- Dalle informazioni date si ottiene l'equazione nell'incognita n:

$$\frac{n}{n+12} = \ldots\ldots\ldots\ldots$$

- Risolvi questa equazione rispetto a n e concludi. [8]

47 Una scatola contiene palline di vario colore. In particolare, si sa che ci sono 8 palline rosse e che la probabilità di estrarre una pallina rossa è 0,25. Quante palline ci sono nell'urna? [32]

48 In un'urna ci sono complessivamente 50 biglie: alcune rosse e alcune nere. Estraendo a caso una biglia dalla scatola, la probabilità che essa sia rossa è 0,32. Determina:

a. la probabilità di estrarre una pallina nera;
b. il numero di palline nere e il numero di palline rosse contenute nell'urna. [**a.** 0,68; **b.** rosse = 16, nere = 34]

49 Una scatola contiene palline di tre colori: nere, bianche e rosse.

a. Sapendo che ci sono 12 palline bianche e che la probabilità di pescare dalla scatola una pallina bianca è $\frac{6}{13}$, determina quante palline ci sono complessivamente nella scatola.
b. Sapendo che la probabilità di pescare dalla scatola una pallina nera è $\frac{3}{4}$ della probabilità di pescare una pallina rossa, determina il numero di palline nere e il numero di palline rosse che sono contenute nella scatola.

[**a.** 26; **b.** nere: 6, rosse: 8]

50 Due urne A e B contengono rispettivamente 15 e 20 palline. Le palline contenute in ciascuna urna sono in parte bianche e in parte nere. Le palline bianche contenute nell'urna B sono il doppio delle palline bianche contenute nell'urna A. La somma tra la probabilità di pescare una pallina bianca dall'urna A e la probabilità di pescare una pallina nera dall'urna B è uguale a $\frac{4}{5}$. Quante palline bianche sono contenute nell'urna A? [6]

51 Due urne A e B contengono rispettivamente 25 e 40 palline. Le palline contenute in ciascuna urna sono in parte bianche e in parte nere. Le palline bianche contenute nell'urna B sono 2 in più delle palline bianche contenute nell'urna A. La somma tra la probabilità di pescare una pallina bianca dall'urna A e la probabilità di pescare una pallina nera dall'urna B è uguale a $\frac{11}{10}$. Quante palline bianche sono contenute nell'urna A? [10]

52 Un'urna contiene 1 pallina bianca e n palline nere. Viene estratta una pallina dall'urna; poi, senza rimettere la pallina estratta nell'urna, ne viene estratta una seconda. Quante palline nere devono essere contenute nell'urna per fare sì che la probabilità che le due palline estratte siano entrambe nere sia il 90% della probabilità che la prima pallina estratta sia nera? [10]

53 Un'urna contiene n palline bianche e n palline nere. Viene estratta una pallina dall'urna; poi, senza rimettere la pallina estratta nell'urna, ne viene estratta una seconda.

a. Determina la probabilità che le due palline estratte siano entrambe bianche.
b. Determina la probabilità che le due palline estratte siano entrambe nere.
c. Determina la probabilità che le due palline estratte siano entrambe dello stesso colore.
d. Sapendo che la probabilità che le due palline estratte siano dello stesso colore è $\frac{11}{23}$, qual è il valore di n?

$\left[\textbf{a. } \frac{n-1}{2(2n-1)}; \textbf{ b. } \frac{n-1}{2(2n-1)}; \textbf{ c. } \frac{n-1}{2n-1}; \textbf{ d. } n = 12\right]$

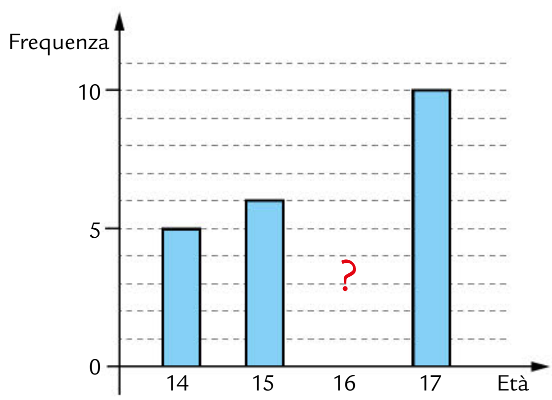
Interpretazione di grafici

54 Un gruppo di ragazzi partecipa a una gara sportiva. Il diagramma a barre qui sotto mostra la distribuzione delle frequenze delle età dei ragazzi del gruppo, ma è incompleto perché manca la frequenza dei ragazzi di 16 anni.

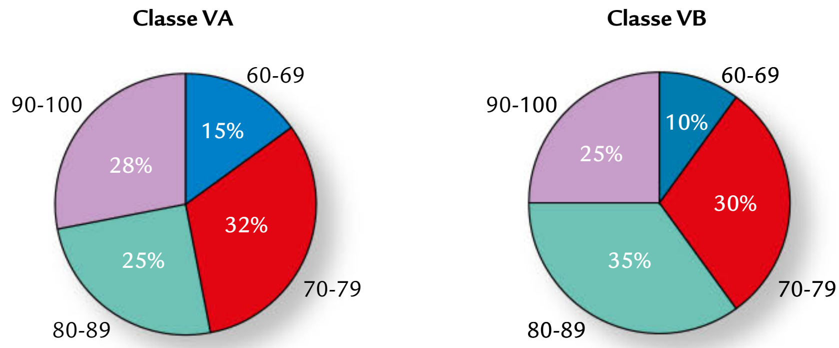
Sapendo che l'età media dei ragazzi del gruppo è 15,8 anni, determina:

a. da quanti ragazzi è composto il gruppo;
b. qual è la probabilità che, scegliendo a caso un ragazzo del gruppo, il ragazzo scelto abbia 16 anni;
c. qual è la probabilità che, scegliendo a caso un ragazzo del gruppo, il ragazzo scelto abbia 16 o 17 anni;
d. qual è la probabilità che, scegliendo a caso un ragazzo del gruppo, il ragazzo scelto abbia al massimo 16 anni.

$\left[\textbf{a. } 30; \textbf{ b. } \frac{3}{10}; \textbf{ c. } \frac{19}{30}; \textbf{ d. } \frac{2}{3}\right]$



55 I due diagrammi a torta qui sotto rappresentano la distribuzione dei voti conseguiti alla maturità dagli studenti delle due classi VA e VB. È noto che, scegliendo a caso uno studente nell'insieme costituito dall'unione degli insiemi degli studenti della VA e della VB, la probabilità che egli abbia conseguito un voto superiore o uguale a 80 è del 57%. Quale delle due classi ha un maggior numero di studenti? Giustifica adeguatamente la risposta.



Utilizzo di diagrammi ad albero

56 Costruisci il diagramma ad albero che rappresenta tutti i possibili esiti che si possono ottenere lanciando due volte una moneta regolare; quindi calcola la probabilità che esca:

a. «croce» entrambe le volte;
b. «croce» la prima volta e «testa» la seconda;
c. almeno una volta «testa».

$\left[\text{a. } \frac{1}{4}; \text{ b. } \frac{1}{4}; \text{ c. } \frac{3}{4}\right]$

57 Si lancia una moneta regolare tre volte, consecutivamente.

a. Rappresenta tramite un diagramma ad albero tutti i possibili esiti.
b. Determina la probabilità che si ottenga «testa» per la prima volta al primo lancio.
c. Determina la probabilità che si ottenga «testa» per la prima volta al secondo lancio.
d. Determina la probabilità che si ottenga «testa» per la prima volta al terzo lancio.
e. Determina la probabilità che non si ottenga mai «testa».

$\left[\text{b. } \frac{1}{2}; \text{ c. } \frac{1}{4}; \text{ d. } \frac{1}{8}; \text{ e. } \frac{1}{8}\right]$

58 Le tre bandiere italiana, francese e spagnola vengono disposte a caso, una a fianco all'altra.

a. Rappresenta, con l'aiuto di un diagramma ad albero, tutti i possibili modi in cui possono essere disposte le tre bandiere.
b. Calcola la probabilità che la bandiera italiana si trovi in mezzo alle altre due.
c. Calcola la probabilità che la bandiera francese si trovi a uno dei due estremi.

$\left[\text{b. } \frac{1}{3}; \text{ c. } \frac{2}{3}\right]$

59 Alberto, Barbara, Carlo e Donatella devono essere interrogati dal loro professore di matematica. Ciascuno di loro viene interrogato singolarmente e il loro professore sceglie a caso l'ordine in cui interrogarli.

a. Rappresenta con un diagramma ad albero tutti i possibili ordini di interrogazione.
b. Determina la probabilità che i quattro ragazzi siano interrogati in ordine alfabetico.
c. Determina la probabilità che Alberto venga interrogato per primo.
d. Determina la probabilità che Donatella sia interrogata prima di Barbara.
e. Determina la probabilità che sia Donatella sia Carlo siano interrogati prima di Barbara.

$\left[\text{b. } \frac{1}{24}; \text{ c. } \frac{1}{4}; \text{ d. } \frac{1}{2}; \text{ e. } \frac{1}{3}\right]$

60 Alberto vuole acquistare un nuovo televisore; può scegliere fra tre modelli, i cui schermi hanno dimensioni diverse, di prezzi 540 euro, 608 euro e 654 euro. Decide inoltre di comprare un decoder per il suo vecchio televisore; per il decoder ha due possibilità di scelta: uno dal costo di 32 euro e uno dal costo di 48 euro. Alberto sceglie a caso sia il modello del televisore, sia quello del decoder.
Determina:

a. la probabilità che Alberto spenda complessivamente meno di 700 euro;
b. la probabilità che Alberto spenda complessivamente più di 650 euro.

[**a.** $\frac{5}{6}$; **b.** $\frac{1}{2}$]

61 Una pulce si trova sull'asse delle ascisse, precisamente nell'origine degli assi. A ogni passo la pulce compie a caso un salto di una unità in avanti (cioè nella direzione e nel verso delle ascisse positive) o indietro (cioè nella direzione e nel verso delle ascisse negative). Supponi che la pulce compia quattro salti.

a. Rappresenta la situazione tramite un diagramma ad albero e deduci quali sono le ascisse dei punti in cui la pulce può trovarsi dopo i quattro salti.
b. Per ciascuno dei punti in cui la pulce può trovarsi dopo i quattro salti, calcola la probabilità che la pulce si trovi in quel punto.

[**a.** Punti possibili: $(\pm 4, 0)$; $(\pm 2, 0)$; $(0, 0)$; **b.** la probabilità che la pulce si trovi nel punto $(-4, 0)$ è $\frac{1}{16}$ ed è uguale alla probabilità che si trovi in $(4, 0)$; la probabilità che la pulce si trovi in $(-2, 0)$ è $\frac{1}{4}$ ed è uguale alla probabilità che si trovi in $(2, 0)$; la probabilità che la pulce si trovi nell'origine è $\frac{3}{8}$]

Utilizzo di tabelle a doppia entrata

62 Si lancia per due volte consecutivamente un dado regolare a sei facce e si considera il prodotto dei due numeri ottenuti. Completa la seguente tabella, in modo da rappresentare il prodotto dei numeri ottenuti in corrispondenza di tutti i possibili esiti dei due lanci del dado.

Lancio 2 / Lancio 1	1	2	3	4	5	6
1	1	2	3	4	5	6
2						
3						
4						
5						
6						

Rispondi quindi alle seguenti domande.

a. Qual è la probabilità che il prodotto dei due numeri ottenuti sia 6?
b. Qual è la probabilità che il prodotto dei due numeri ottenuti sia 12?
c. Qual è la probabilità che il prodotto dei due numeri ottenuti sia maggiore di 12?
d. Qual è la probabilità che il prodotto dei due numeri ottenuti sia minore di 12?
e. Qual è la probabilità che il prodotto dei due numeri ottenuti sia minore di 36?
f. Qual è la probabilità che il prodotto dei due numeri ottenuti sia maggiore di 36?
g. Qual è la probabilità che il prodotto dei due numeri ottenuti sia minore di 37?

[**a.** $\frac{1}{9}$; **b.** $\frac{1}{9}$; **c.** $\frac{13}{36}$; **d.** $\frac{19}{36}$; **e.** $\frac{35}{36}$; **f.** 0; **g.** 1]

63 Si lancia per due volte consecutivamente un dado regolare a sei facce e si considera lo *scarto* tra i due numeri ottenuti, ossia il valore assoluto della loro differenza.

- Costruisci una tabella (simile a quella dell'esercizio precedente) che fornisca lo scarto dei due numeri ottenuti in corrispondenza di tutti i possibili esiti dei due lanci del dado.
- Tenendo conto della tabella, rispondi alle seguenti domande.

a. Qual è la probabilità che lo scarto tra i due numeri sia 1?
b. Qual è la probabilità che lo scarto tra i due numeri sia 2?
c. Qual è la probabilità che lo scarto tra i due numeri sia 3?

[**a.** $\frac{5}{18}$; **b.** $\frac{2}{9}$; **c.** $\frac{1}{6}$]

64 Un sacchetto contiene cinque palline, numerate da 1 a 5. Si estrae a caso una pallina, quindi la si rimette nel sacchetto e si estrae una seconda pallina.

a. Costruisci una tabella in cui rappresenterai la somma dei due numeri ottenuti in corrispondenza di tutti i possibili esiti delle due estrazioni.
b. Qual è la probabilità che i due numeri estratti abbiano somma uguale a 8?
c. Qual è la probabilità che i due numeri estratti abbiano somma minore di 5?
d. Qual è la probabilità che i due numeri estratti abbiano somma divisibile per 3?

$\left[\mathbf{b.}\ \frac{3}{25};\ \mathbf{c.}\ \frac{6}{25};\ \mathbf{d.}\ \frac{9}{25}\right]$

65 Si è interrogato un gruppo di 100 persone, chiedendo loro se sono favorevoli all'energia nucleare. Il sondaggio ha prodotto i risultati riportati (in parte) nella seguente tabella.

Età / Opinione espressa	Meno di 35 anni	Più di 35 anni	Totale
Favorevoli		12	44
Contrari			
Totale	52		100

a. Completa la tabella.
b. Tra le persone interrogate favorevoli al nucleare se ne sceglie una a caso: qual è la probabilità che abbia meno di 35 anni?
c. Tra le persone interrogate che hanno meno di 35 anni se ne sceglie una a caso: qual è la probabilità che sia contraria all'energia nucleare?
d. Tra le persone che hanno partecipato al sondaggio se ne sceglie una a caso; calcola la probabilità:
- che abbia più di 35 anni e sia contraria al nucleare;
- che sia favorevole al nucleare;
- che abbia meno di 35 anni e sia favorevole al nucleare;
- che abbia più di 35 anni.

$\left[\mathbf{b.}\ \frac{8}{11};\ \mathbf{c.}\ \frac{5}{13};\ \mathbf{d.}\ \frac{9}{25},\ \frac{11}{25},\ \frac{8}{25},\ \frac{12}{25}\right]$

66 Un fiorista mette in svendita 200 fiori, in parte rose e in parte tulipani. Sia le rose sia i tulipani sono di due tipi: o di colore rosso o di colore giallo. Il 60% dei fiori in offerta sono rossi e il 35% sono tulipani. Inoltre le rose rosse sono 70. Scegliendo a caso un fiore tra quelli in svendita, calcola qual è la probabilità:

a. che sia giallo;
b. che sia una rosa;
c. che sia un tulipano rosso;
d. che sia una rosa gialla.

(*Suggerimento*: può essere utile costruire una tabella tipo quella dell'esercizio precedente per rappresentare i dati)

$\left[\mathbf{a.}\ \frac{2}{5};\ \mathbf{b.}\ \frac{13}{20};\ \mathbf{c.}\ \frac{1}{4};\ \mathbf{d.}\ \frac{3}{10}\right]$

Utilizzo delle regole del calcolo combinatorio

67 ESERCIZIO SVOLTO

Considera i numeri di cinque cifre aventi come cifre soltanto 1 o 2; per esempio: 21 211, 11 222 e così via... Si sceglie a caso uno di questi numeri; qual è la probabilità che le prime due cifre siano uguali a 1?

1° modo

- Calcoliamo anzitutto quanti sono i numeri composti da cinque cifre, uguali a 1 o 2. Poiché per ogni cifra abbiamo 2 possibilità di scelta, abbiamo in tutto: $2 \cdot 2 \cdot 2 \cdot 2 \cdot 2 = 2^5$ numeri siffatti.
- Tra questi numeri, quanti sono quelli che hanno le prime due cifre uguali a 1? Pensiamo a come possiamo fare a costruirli: per le prime due cifre abbiamo 1 sola possibilità di scelta (1); per le altre tre cifre abbiamo 2 possibilità di scelta (1 o 2), quindi in definitiva tali numeri sono complessivamente: $1 \cdot 1 \cdot 2 \cdot 2 \cdot 2 = 2^3$.
- La probabilità richiesta può essere calcolata come rapporto tra casi favorevoli e casi possibili (perché?), perciò è uguale a $\frac{2^3}{2^5}$, cioè a $\frac{1}{4}$.

2° modo
In alternativa, si può procedere più rapidamente considerando come spazio campionario l'insieme delle coppie ordinate formate soltanto dalle *prime due* cifre (le restanti tre cifre sono infatti irrilevanti): in questo modo l'insieme dei casi possibili diviene 4 e c'è 1 solo caso favorevole. Quindi la probabilità è $\frac{1}{4}$.

68 ESERCIZIO GUIDATO

Qual è la probabilità che 5 stazioni ferroviarie consecutive su una stessa linea, siano disposte nell'ordine alfabetico crescente e decrescente del nome della città?

- I «casi possibili», tutti equiprobabili, sono i modi in cui le 5 stazioni possono essere disposte lungo la linea ferroviaria. Essi sono tanti quanti le permutazioni di 5 oggetti, cioè 5!
- I «casi favorevoli» sono solo 2: quello in cui le cinque stazioni sono ordinate in senso crescente e quello in cui le cinque stazioni sono ordinate in senso decrescente.
- Ora puoi concludere, costruendo il rapporto tra i casi favorevoli e quelli possibili. Troverai che la probabilità richiesta è uguale a $\frac{1}{60}$.

69 ESERCIZIO GUIDATO

Qual è la probabilità di fare una «ambata» al lotto, cioè che si giochi un singolo numero e che esso faccia parte dei cinque estratti?

1° modo

- I casi possibili, tutti equiprobabili, sono le cinquine *non ordinate* costruibili con i 90 numeri del lotto (1, 2, ..., 90), cioè $\binom{90}{5}$.
- I casi favorevoli sono tanti quanti le cinquine (non ordinate) che contengono il numero giocato. Queste ultime sono tante quante le quaterne (non ordinate) costruibili utilizzando gli 89 numeri diversi da quello giocato, cioè $\binom{89}{4}$.
- Ora hai tutti gli elementi per calcolare la probabilità cercata; se svolgi i calcoli correttamente, troverai che essa è uguale a $\frac{1}{18}$.

2° modo
La probabilità che, in un'estrazione di cinque numeri del lotto, il numero giocato esca per primo è ovviamente $\frac{1}{90}$; la probabilità che su cinque numeri estratti il numero giocato esca per secondo è uguale alla precedente, infatti è data dal prodotto:

$$\underbrace{\frac{1}{89}}_{\substack{\text{probabilità}\\ \text{di estrarre il}\\ \text{numero giocato}\\ \text{nella seconda}\\ \text{estrazione}}} \cdot \underbrace{\frac{89}{90}}_{\substack{\text{probabilità}\\ \text{che non sia stato}\\ \text{estratto il numero}\\ \text{giocato nella}\\ \text{prima estrazione}}} = \frac{1}{90} \quad \text{e così via.}$$

Dunque la probabilità che il numero giocato esca tra i cinque numeri estratti è:

$$\frac{5}{90} = \frac{1}{18}.$$

70 Stefano digita a caso dieci cifre sulla tastiera del telefono, senza tenere conto del fatto che il numero telefonico deve iniziare per 0 e ammettendo la possibilità di digitare più volte la stessa cifra. Qual è la probabilità che abbia composto il numero di sua zia Maria? $\left[\frac{1}{10^{10}}\right]$

ESERCIZI

71 Quanti sono i possibili anagrammi (anche senza significato) della parola «scuola»? Si sceglie a caso uno di questi anagrammi; qual è la probabilità che la terza lettera dell'anagramma sia una vocale e l'ultima una consonante? $\left[720; \frac{3}{10}\right]$

72 Il numero di serie di una banconota è formato da una lettera e da 11 numeri.

a. Quanti numeri di serie distinti si possono generare, supponendo che la lettera possa essere una qualsiasi dell'alfabeto a 26 lettere e ogni numero possa essere una cifra qualsiasi tra 0 e 9?

b. Scelto a caso un numero di serie fra tutti quelli possibili, qual è la probabilità che esso inizi con una vocale e termini con un numero pari diverso da zero? $\left[\text{a. } 26 \cdot 10^{11}; \text{ b. } \frac{1}{13}\right]$

73 Un'agenzia matrimoniale ha 70 clienti maschi, di cui 50 hanno 40 anni o più, e 80 clienti donne, di cui 20 hanno 35 anni o più. Scelta a caso una coppia formata da un uomo e una donna clienti dell'agenzia, qual è la probabilità che lui abbia 40 anni o più e lei abbia meno di 35 anni? $\left[\frac{15}{28}\right]$

74 **E se?** Una password è formata da 6 caratteri, che possono essere cifre (0, 1, ..., 9) o lettere minuscole dell'alfabeto italiano composto da 5 vocali e 16 consonanti. Inoltre, cifre e lettere possono essere ripetuti.

a. Quante diverse password possono essere generate?

b. Generando a caso una password, qual è la probabilità di comporne una che comincia con una vocale e finisce con una consonante?

c. Generando a caso una password, qual è la probabilità di comporne una che ha due vocali nelle prime due posizioni, due cifre nelle successive e due consonanti nelle ultime due posizioni?

▸ Come cambierebbero le risposte lasciando cadere l'ipotesi che cifre e lettere possano essere ripetuti?

$\left[\text{a. } 31^6; \text{ b. } \frac{80}{31^2}; \text{ c. } \frac{5^2 \cdot 10^2 \cdot 16^2}{31^6}; \text{ a. } 530\,122\,320; \text{ b. } \frac{8}{93}; \text{ c. } \frac{200}{245\,427}\right]$

75 Anna e Barbara si recano al ristorante. Entrambe ordinano un antipasto, un primo e un dessert. È possibile scegliere fra 6 antipasti, 8 primi e 4 dessert. Anna sceglie il suo menu, quindi lo sceglie Barbara, selezionando a caso fra gli antipasti, i primi e i dessert disponibili. Calcola le seguenti probabilità:

a. che Barbara abbia scelto lo stesso antipasto di Anna;

b. che Barbara abbia scelto lo stesso antipasto e lo stesso primo di Anna;

c. che Barbara abbia scelto lo stesso menu di Anna. $\left[\text{a. } \frac{1}{6}; \text{ b. } \frac{1}{48}; \text{ c. } \frac{1}{192}\right]$

76 Tre persone salgono sull'ascensore al piano terreno. L'ascensore si ferma a sei diversi piani. Supponi che ciascuna delle tre persone che hanno preso l'ascensore scenda a caso a uno di questi sei piani.

a. Qual è la probabilità che le tre persone scendano tutte al primo piano?

b. Qual è la probabilità che le tre persone scendano tutte allo stesso piano?

c. Qual è la probabilità che le tre persone scendano al quarto o al quinto piano (ammettendo sia la possibilità che scendano tutte e tre allo stesso piano, sia la possibilità che alcune scendano al quarto piano e le rimanenti al quinto)?

d. Qual è la probabilità che le tre persone scendano o tutte e tre al quarto piano o tutte e tre al quinto piano?

$\left[\text{a. } \frac{1}{216}; \text{ b. } \frac{1}{36}; \text{ c. } \frac{1}{27}; \text{ d. } \frac{1}{108}\right]$

77 Si scelgono simultaneamente sei carte da un mazzo di 40 (cioè da un mazzo ottenuto da quello di 52 eliminando gli 8, i 9 e i 10). Qual è la probabilità che tra le sei carte estratte ci sia l'asso di cuori? $\left[\frac{3}{20}\right]$

Zoom sull'enunciato

78 Si scelgono simultaneamente cinque carte da un mazzo di 40 (cioè da un mazzo ottenuto da quello di 52 eliminando gli 8, i 9 e i 10). Qual è la probabilità che tra le cinque carte estratte ci siano l'asso di cuori *e* quello di fiori?
Qual è la probabilità che tra le cinque carte estratte ci sia l'asso di cuori *o* l'asso di fiori? Qual è la probabilità che tra le cinque carte estratte ci sia *o* l'asso di cuori *o* l'asso di fiori (ma non entrambi)?

Guida alla lettura

Il problema pone tre domande che differiscono solo per un connettivo: «e» , «o» (usato in senso *inclusivo*), «o... o...» (cioè «o» usato in senso *esclusivo*).

$\left[\frac{1}{78}; \frac{37}{156}; \frac{35}{156}\right]$

79 **Videolezione** Considera un mazzo di 52 carte.

a. Se si estrae a caso una carta dal mazzo, qual è la probabilità di estrarre una carta che non sia né una figura, né una carta di cuori?

b. Se si estraggono a caso e simultaneamente due carte dal mazzo, qual è la probabilità che tra le due carte estratte non ci sia né una figura né una carta di cuori?

$\left[\text{a. } \frac{15}{26}; \text{ b. } \frac{145}{442}\right]$

80 Un'urna contiene 4 palline nere e 3 palline bianche. Si estraggono contemporaneamente due palline dall'urna. Calcola la probabilità:

a. di estrarre due palline nere;

b. di estrarre due palline bianche;

c. di estrarre due palline dello stesso colore;

d. di estrarre due palline di colori differenti.

$\left[\text{a. } \frac{2}{7}; \text{ b. } \frac{1}{7}; \text{ c. } \frac{3}{7}; \text{ d. } \frac{4}{7}\right]$

81 Qual è la probabilità di fare un «ambo» al lotto, cioè che si giochino 2 numeri ed essi facciano parte dei cinque estratti?

$\left[\frac{2}{801}\right]$

---

82 Qual è la probabilità di fare un «terno» al lotto, cioè che si giochino 3 numeri ed essi facciano parte dei cinque estratti?

$\left[\frac{1}{11\,748}\right]$

83 Una classe è costituita da 12 maschi e 8 femmine. Si scelgono tre alunni a caso nella classe; qual è la probabilità che siano tutti dello stesso sesso?

$\left[\frac{23}{95}\right]$

84 In una città ci sono sei hotel. Un dato giorno, tre persone prenotano ciascuna una camera in uno degli hotel, in modo del tutto casuale. Qual è la probabilità che le tre persone si trovino in tre hotel differenti?

$\left[\frac{5}{9}\right]$

85 Si estraggono a caso tre carte da un mazzo di 40. Qual è la probabilità che siano tutte dello stesso seme?

$\left[\frac{12}{247}\right]$

---

86 **Videolezione** Sette amici, quattro ragazzi e tre ragazze, si recano al cinema e si siedono a caso, tutti vicini, sulle poltrone di una stessa fila. Calcola la probabilità:

a. che i ragazzi siano tutti vicini tra loro;

b. che le ragazze siano tutte vicine tra loro;

c. che i ragazzi siano tutti vicino tra loro e le ragazze tutte vicine tra loro.

$\left[\text{a. } \frac{4}{35}; \text{ b. } \frac{1}{7}; \text{ c. } \frac{2}{35}\right]$

87 Un'urna contiene 3 palline rosse, 4 palline blu e 5 palline verdi. Si estraggono dall'urna simultaneamente tre palline; determina la probabilità:

a. di estrarre tre palline dello stesso colore;

b. di estrarre tre palline di colore differente.

Risolvi poi lo stesso problema supponendo che le tre palline, anziché essere estratte simultaneamente, siano estratte consecutivamente, rimettendo nell'urna dopo ciascuna estrazione la pallina estratta, prima dell'estrazione successiva.

$\left[\text{a. } \frac{3}{44}; \text{ b. } \frac{3}{11}; \text{ nel caso di estrazioni consecutive con reimmissione: a. } \frac{1}{8}; \text{ b. } \frac{5}{24}\right]$

88 Un'urna contiene 21 gettoni, su ciascuno dei quali è impressa una delle 21 lettere dell'alfabeto italiano. Si estrae dall'urna, senza reimmissione, una terna ordinata di gettoni. Calcola la probabilità:

a. che la terna estratta sia formata soltanto da vocali;

b. che la terna estratta sia formata soltanto da consonanti;

c. che la terna estratta corrisponda a un anagramma dell'aggettivo «MIO».

$\left[\text{a. } \frac{1}{133}; \text{ b. } \frac{8}{19}; \text{ c. } \frac{1}{1330}\right]$

89 Una riserva naturale è popolata da 25 cervi. Cinque di essi vengono catturati, marchiati e rimessi in libertà. Successivamente vengono catturati a caso quattro cervi. Qual è la probabilità che esattamente due di essi siano marchiati?

$\left[\frac{38}{253}\right]$

90 Un'urna contiene otto palline, numerate da 1 a 8. Determina la probabilità che:

a. estraendo consecutivamente tre palline, senza reimmissione, si ottengano tre numeri pari;

b. estraendo consecutivamente tre palline, con reimmissione, si ottengano tre numeri pari;

c. estraendo simultaneamente tre palline, non si ottenga alcun multiplo di 3.

$\left[\text{a. } \frac{1}{14}; \text{ b. } \frac{1}{8}; \text{ c. } \frac{5}{14}\right]$

91 Considera le lettere da cui è formata a parola «missississippi» raffigurata qui sotto:

miSsISsiPpi

Supponiamo di considerare distinte due lettere che differiscono per il carattere (maiuscolo o minuscolo) o per il colore. Permutando a caso le lettere, qual è la probabilità di ottenere ancora la parola «missississippi»? $\left[\frac{1}{34\,650}\right]$

92 In base ad alcune analisi statistiche, si è stabilito che il 5% dei componenti elettronici prodotti da un'azienda presenta qualche difetto. Da un lotto di 40 componenti elettronici se ne estraggono simultaneamente 5. Determina la probabilità che nel campione estratto:

a. non ci sia alcun componente difettoso;

b. ci sia esattamente un componente difettoso.

$\left[\text{a. } \frac{119}{156}; \text{ b. } \frac{35}{156}\right]$

93 Un'urna contiene n palline. Se si estrae una pallina, la probabilità che essa sia gialla è $\frac{4}{7}$. Se invece si estraggono contemporaneamente due palline, la probabilità che esse siano gialle è $\frac{38}{119}$. Quanto vale n? Quante sono le palline gialle contenute nell'urna?

[$n = 35$; 20 palline gialle]

94 Il numero 3240 è uguale a $2^3 \cdot 3^4 \cdot 5$.

a. Tenendo conto della scomposizione in fattori primi di 3240 e del principio fondamentale del calcolo combinatorio, determina quanti sono i divisori di 3240.

b. Scelto a caso un divisore di 3240, qual è la probabilità che sia una potenza di 2 (con esponente positivo o nullo)?

$\left[\text{a. } 40; \text{ b. } \frac{1}{10}\right]$

3. I primi teoremi sul calcolo delle probabilità

Teoria p. 145

Esercizi introduttivi

Test

95 L'evento A ha probabilità $\frac{2}{5}$; allora la probabilità dell'evento $\overline{A}$ è uguale a:

A $\frac{1}{5}$ B $\frac{2}{5}$ C $\frac{3}{5}$ D $\frac{4}{5}$

96 Due eventi A e B sono incompatibili e tali che $p(A) = 0{,}15$ e $p(B) = 0{,}35$; allora la probabilità dell'evento $A \cup B$:

A è uguale a 0

B è uguale a 0,5

C è uguale a 0,75

D i dati sono insufficienti per determinarla

97 Due eventi A e B sono tali che $p(A) = 0,5$, $p(B) = 0,3$ e $p(A \cap B) = 0,1$; allora la probabilità dell'evento $A \cup B$:

A è uguale a 0,5

B è uguale a 0,7

C è uguale a 0,75

D i dati sono insufficienti per determinarla

98 Siano A e B due eventi incompatibili; quale delle seguenti affermazioni è *falsa*?

A $A \cap B$ è vuoto

B La probabilità di $A \cap B$ è 0

C La probabilità di $A \cup B$ è 1

D $p(A) = p(A \cup B) - p(B)$

99 **Vero o falso?**

a. la probabilità di un evento e quella del suo contrario sono sempre diverse tra loro V F

b. la probabilità di $A \cup B$ è sempre maggiore della somma delle probabilità di A e B V F

c. due eventi contrari sono incompatibili V F

d. due eventi incompatibili sono contrari V F

e. $p(A \cap B) + p(A \cup B) = p(A) + p(B)$ V F

[2 affermazioni vere e 3 false]

Collegamenti **Probabilità e logica**

100 **La negazione di una proposizione / 1**

a. Si estrae una carta da un mazzo di 40. Esprimi a parole gli eventi *contrari* dei seguenti, ossia le *negazioni* delle proposizioni che esprimono gli eventi, ricordando le leggi di de Morgan (vedi ***Volume 1*** Unità 3):

A: «è stato estratto un Fante o un Re» B: «è stato estratto un Fante di cuori»

b. Si estraggono cinque carte da un mazzo di 40. Esprimi a parole gli eventi *contrari* dei seguenti, ossia le *negazioni* delle proposizioni che esprimono gli eventi, ricordando come si negano le proposizioni che contengono dei quantificatori:

A: «tutte le carte estratte sono di cuori» B: «almeno una delle carte estratte è di picche»

101 **La negazione di una proposizione / 2**

a. Un'urna contiene 3 biglie verdi e 5 gialle. Si estraggono 4 biglie dall'urna e si considera l'evento A: «le quattro biglie estratte sono gialle». L'evento *contrario* di A è «nessuna biglia estratta è gialla» oppure «almeno una delle biglie estratte è verde»?

b. Sempre in relazione alla situazione descritta al punto precedente, considera l'evento B: «al massimo due delle biglie estratte sono gialle». L'evento *contrario* di B è «tra le biglie estratte ce ne sono esattamente tre o quattro gialle» oppure «tra le biglie estratte ce ne sono due o più gialle»?

c. Estraiamo un numero a caso dal sacchetto della tombola. Considera l'evento C: «il numero estratto è dispari e multiplo di 10». L'evento *contrario* di C è «il numero estratto è pari o non è multiplo di 10» oppure «il numero estratto è pari e non è multiplo di 10»?

Applicazioni dei teoremi sulle probabilità dell'unione, dell'intersezione e dell'evento contrario

Negli Esercizi 102-106 A e B sono due eventi.

102 Sapendo che $p(A) = 0{,}5$, $p(B) = 0{,}6$ e $p(A \cap B) = 0{,}2$, quanto vale $p(A \cup B)$? [0,9]

103 Sapendo che A e B sono incompatibili, $p(A) = 0{,}36$ e $p(B) = 0{,}21$, quanto vale $p(A \cup B)$? [0,57]

104 Sapendo che $p(A) = 0{,}7$, $p(\overline{B}) = 0{,}6$ e $p(A \cup B) = 0{,}9$, quanto vale $p(A \cap B)$? [0,2]

105 Sapendo che $p(A) = 0{,}2$, $p(A \cap B) = 0{,}1$ e $p(A \cup B) = 0{,}7$, quanto vale $p(B)$? [0,6]

106 Siano A e B due eventi incompatibili tali che $p(A) = 0{,}4$ e $p(B) = 0{,}3$.

a. Calcola la probabilità dell'evento $A \cup B$.

b. Stabilisci se gli eventi $\overline{A}$ e $\overline{B}$ sono incompatibili.

[**a.** 0,7; **b.** non sono incompatibili: infatti la probabilità delle loro intersezione è 0,3]

Problemi sulla probabilità dell'evento contrario

107 **ESERCIZIO GUIDATO**

Si lancia una moneta regolare per quattro volte. Calcola la probabilità:

a. che non esca mai «testa»;

b. che esca almeno una volta «testa»;

c. che esca al massimo una volta «testa»;

d. che esca «testa» più di una volta.

a. La probabilità che in quattro lanci non esca mai «testa» equivale alla probabilità che in quattro lanci esca sempre «croce»: pertanto la probabilità richiesta è uguale a ....

b. L'evento «esce 'testa' almeno una volta» è l'evento contrario di «non esce mai 'testa'», pertanto la sua probabilità è data da 1 meno la probabilità calcolata al punto **a**.

c. L'evento E: «esce 'testa' al massimo una volta» equivale a «non esce mai 'testa' o esce 'testa' esattamente una volta» ossia all'unione due eventi incompatibili A: «non esce mai 'testa'» e B: «esce 'testa' esattamente una volta», quindi la probabilità di E è uguale alla somma delle probabilità di A e B.

d. L'evento «esce 'testa' più di una volta» è l'evento contrario di «esce 'testa' al massimo una volta», pertanto la sua probabilità è uguale a ....

$\left[\textbf{a. } \frac{1}{16}; \textbf{ b. } \frac{15}{16}; \textbf{ c. } \frac{5}{16}; \textbf{ d. } \frac{11}{16}\right]$

108 Scegli a caso un numero intero compreso tra 1 e 10, inclusi 1 e 10.

a. Calcola la probabilità dei seguenti eventi:

A: «il numero scelto è dispari»

B: «il numero scelto è maggiore o uguale a 8»

C: «il numero scelto è primo»

b. Esprimi a parole gli eventi contrari e calcola le loro probabilità.

$\left[\text{a. } \frac{1}{2}, \frac{3}{10}, \frac{2}{5}; \text{ b. } \frac{1}{2}, \frac{7}{10}, \frac{3}{5}\right]$

109 Ciascuna delle sei facce di un cubo viene colorata a caso in bianco o in nero.

a. In quanti modi diversi può essere colorato il cubo?

b. Qual è la probabilità che almeno due facce del cubo siano colorate con colori differenti? $\left[\text{a. } 64; \text{ b. } \frac{31}{32}\right]$

110 Si lanciano due dadi regolari a sei facce. Calcola:

a. la probabilità che i due numeri usciti siano uguali;

b. la probabilità che i due numeri usciti siano distinti.

$\left[\text{a. } \frac{1}{6}, \text{ b. } \frac{5}{6}\right]$

111 Un dado regolare a sei facce viene lanciato per tre volte. Qual è la probabilità che i numeri usciti non siano tutti e tre uguali? $\left[\frac{35}{36}\right]$

112 Un'urna contiene 20 palline, numerate da 1 a 20. Viene estratta una pallina dall'urna, quindi, senza rimettere la prima pallina estratta nell'urna, si estrae una seconda pallina. Qual è la probabilità che il prodotto dei due numeri impressi sulle palline estratte sia un numero pari? $\left[\frac{29}{38}\right]$

113 **ESERCIZIO GUIDATO**

A una gara sportiva partecipano cinque atleti, tra cui Paolo, Alessandro e Alberto.

a. Qual è la probabilità che il primo classificato sia Paolo, Alessandro o Alberto?

b. Qual è la probabilità che almeno uno tra Paolo, Alessandro e Alberto sia tra i primi due classificati?

c. Qual è la probabilità che almeno uno tra Paolo, Alessandro e Alberto sia tra i primi tre classificati?

---

a. Vi sono 5 possibili vincitori (tanti quanti i partecipanti) e 3 casi favorevoli (corrispondenti alla vincita di Paolo, Alessandro o Alberto). Quindi la probabilità cercata è $\frac{3}{5}$.

b. Utilizza il passaggio all'evento contrario. La probabilità richiesta è uguale a 1 meno la probabilità dell'evento *E*: «nessuno dei tre amici rientra tra i primi due classificati». Osserva che *E* equivale a sua volta all'evento «i primi due classificati sono i restanti due atleti partecipanti»: quest'ultimo evento ha probabilità $\frac{1}{\ldots}$, poiché le possibili scelte di due atleti tra i cinque partecipanti sono $\binom{5}{2} = \ldots\ldots$ Pertanto la probabilità richiesta è uguale a .....

c. Osserva che, essendo i partecipanti alla gara in tutto 5, necessariamente uno tra Paolo, Alessandro e Alberto deve comparire tra i primi tre: si tratta quindi di calcolare la probabilità dell'evento ....

$\left[\text{a. } \frac{3}{5}; \text{ b. } \frac{9}{10}; \text{ c. } 1\right]$

114 Una password è costituita da cinque numeri, ciascuno dei quali è una delle dieci cifre (0, 1, 2, ..., 9). È consentito che le cifre siano ripetute.

a. Quante password di questo tipo si possono generare?

b. Scelta a caso una di tali password, qual è la probabilità che almeno due delle cifre della password siano uguali? $\left[\text{a. } 10^5; \text{ b. } \frac{436}{625}\right]$

115 Quattro amici discutono delle loro date di nascita.

a. Qual è la probabilità che almeno due di essi siano nati lo stesso mese?

b. Sapendo che nessuno dei quattro amici è nato il 29 febbraio, qual è la probabilità che almeno due di essi siano nati nello stesso giorno dell'anno?

$\left[\text{a. } \frac{41}{96} \simeq 43\%; \text{ b. } \frac{795\,341}{48\,627\,125} \simeq 1{,}6\%\right]$

Problemi sulla probabilità dell'unione e dell'intersezione di eventi

116 Si estrae a caso una carta da un mazzo di 40. Calcola la probabilità:

a. di estrarre una figura;

b. di estrarre una carta di colore nero;

c. di estrarre una figura di colore nero;

d. di estrarre una figura o una carta di colore nero.

$\left[\text{a. } \frac{3}{10}; \text{ b. } \frac{1}{2}; \text{ c. } \frac{3}{20}; \text{ d. } \frac{13}{20}\right]$

117 Da un sacchetto che contiene 12 biglie numerate da 1 a 12 se ne estrae una a caso. Calcola la probabilità che il numero sulla biglia estratta:

a. sia pari;

b. sia divisibile per 3;

c. sia pari e divisibile per 3;

d. sia pari o divisibile per 3. $\left[\text{a. } \frac{1}{2}; \text{ b. } \frac{1}{3}; \text{ c. } \frac{1}{6}; \text{ d. } \frac{2}{3}\right]$

118 Un'urna contiene 8 palline: 5 sono verdi e sono numerate da 1 a 5, mentre le altre 3 sono rosse e sono numerate da 1 a 3. Si estrae a caso una pallina dall'urna; considera i seguenti eventi:

V: «la pallina estratta è verde»

R: «la pallina estratta è rossa»

P: «la pallina estratta ha impresso un numero pari»

a. Determina la probabilità dei tre eventi V, R, P.

b. Determina la probabilità degli eventi $V \cap R$, $V \cap P$, $R \cap P$.

c. Determina la probabilità degli eventi $V \cup R$, $V \cup P$, $R \cup P$. $\left[\mathbf{a.}\ \frac{5}{8}, \frac{3}{8}, \frac{3}{8};\ \mathbf{b.}\ 0, \frac{1}{4}, \frac{1}{8};\ \mathbf{c.}\ 1, \frac{3}{4}, \frac{5}{8}\right]$

119 Un ufficio bancario ha due sportelli, diciamo A e B, di cui almeno uno sempre aperto. La probabilità che sia aperto lo sportello A è 0,7 mentre la probabilità che sia aperto lo sportello B è 0,6. Qual è la probabilità che siano aperti entrambi gli sportelli? [0,3]

120 In un liceo, $\frac{1}{4}$ degli allievi frequentano la prima classe, il 60% degli allievi sono maschi e gli allievi maschi che frequentano la prima sono $\frac{1}{10}$ del numero complessivo di allievi del liceo. Scelto a caso un allievo del liceo, qual è la probabilità che sia maschio o frequenti la prima? $\left[\frac{3}{4}\right]$

121 Un circolo sportivo, avente 125 soci, organizza due tornei, uno di tennis e uno di calcetto; 60 soci partecipano al torneo di tennis, 45 al torneo di calcetto e 25 a entrambi i tornei. Si sceglie a caso uno dei soci del circolo; determina la probabilità che tale socio partecipi:

a. a entrambi i tornei;
b. soltanto al torneo di tennis;
c. soltanto al torneo di calcetto;
d. ad almeno uno dei due tornei;
e. a nessuno dei due tornei.

$\left[\mathbf{a.}\ \frac{1}{5};\ \mathbf{b.}\ \frac{7}{25};\ \mathbf{c.}\ \frac{4}{25};\ \mathbf{d.}\ \frac{16}{25};\ \mathbf{e.}\ \frac{9}{25}\right]$

122 ESERCIZIO GUIDATO

Un ospedale ha due sale operatorie, diciamo X e Y, che hanno la stessa probabilità di essere occupate. La probabilità che almeno una delle due sale risulti occupata è 0,8; la probabilità che entrambe le sale operatorie siano occupate è 0,4.

a. Qual è la probabilità che la sala operatoria X sia libera?
b. Qual è la probabilità che entrambe le sale operatorie siano libere?
c. Qual è la probabilità che almeno una delle due sale operatorie sia libera?
d. Qual è la probabilità che una sola delle due sale operatorie sia libera?

a. Sia A l'evento «la sala operatoria X è occupata» e B l'evento «la sala operatoria Y è occupata». In base ai dati sai che:

$$p(A \cup B) = 0{,}8 \qquad p(A \cap B) = 0{,}4 \qquad p(A) = p(B)$$

Poni $p(A) = p(B) = x$ e sostituisci le informazioni note nella relazione:

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

Ottieni un'equazione in x che, risolta, fornisce la probabilità che la sala operatoria X sia occupata. La probabilità che X sia libera è la probabilità dell'evento contrario di A, quindi è uguale a ....

b. Devi calcolare la probabilità dell'evento $\overline{A} \cap \overline{B}$. Ricorda che, per le leggi di de Morgan, risulta $\overline{A} \cap \overline{B} = \overline{(A \cup B)}$, quindi $p(\overline{A} \cap \overline{B}) = 1 - p(A \cup B)$.

c. Devi calcolare la probabilità dell'evento $\overline{A} \cup \overline{B}$. Puoi procedere in modo simile al caso **b**.

d. Devi calcolare la probabilità dell'evento $(\overline{A} \cap B) \cup (A \cap \overline{B})$. Osserva che si tratta dell'unione di due eventi incompatibili, quindi la probabilità dell'evento $(\overline{A} \cap B) \cup (A \cap \overline{B})$ è uguale a:

$$p(\overline{A} \cap B) + p(A \cap \overline{B})$$

Per calcolare, per esempio, $p(\overline{A} \cap B)$, osserva che $\overline{A} \cap B = B - (A \cap B)$. In modo analogo puoi calcolare $p(A \cap \overline{B})$.

[**a.** 0,4; **b.** 0,2; **c.** 0,6; **d.** 0,4]

123 Il 70% degli studenti di una classe andrà in vacanza al mare, il 40% andrà in montagna e il 30% andrà sia al mare sia in montagna. Qual è la probabilità che uno studente scelto a caso in quella classe non andrà in vacanza né al mare né in montagna? [0,2]

124 In una data popolazione, la probabilità che un individuo presenti il carattere genetico A è il doppio di quella che presenti il carattere genetico B; inoltre la probabilità che un individuo presenti entrambi i caratteri genetici è 0,2 e quella che presenti almeno uno dei due caratteri è 0,7. Scelto a caso un individuo in quella popolazione, determina:

a. qual è la probabilità che presenti il carattere A;
b. qual è la probabilità che non presenti il carattere B;
c. qual è la probabilità che presenti il carattere A ma non il carattere B. [**a.** 0,6; **b.** 0,7; **c.** 0,4]

Matematica e controllo della qualità

125 Un oggetto prodotto da una macchina può presentare due tipi di difetti, diciamo A e B.
Scelto a caso un oggetto prodotto dalla macchina, la probabilità che presenti il difetto A è 0,2; la probabilità che presenti il difetto B è 0,3 e la probabilità che non presenti alcun difetto è 0,6.
Determina la probabilità che l'oggetto:

a. presenti almeno uno dei due difetti;
b. presenti entrambi i difetti;
c. non presenti il difetto A ma presenti il difetto B. [**a.** 0,4; **b.** 0,1; **c.** 0,2]

126 Un oggetto prodotto da una macchina può presentare due tipi di difetti, diciamo A e B. Scelto a caso un oggetto prodotto dalla macchina, la probabilità che presenti il difetto A è 0,1; la probabilità che presenti il difetto B è 0,2 e la probabilità che non presenti alcun difetto è 0,75. Determina la probabilità che l'oggetto:

a. presenti almeno uno dei due difetti;
b. presenti entrambi i difetti;
c. non presenti il difetto A ma presenti il difetto B. [**a.** 0,25; **b.** 0,05; **c.** 0,15]

127 Paola, il pomeriggio, è solita chiamare due sue amiche, Maria e Luisa, che non sempre però le rispondono. La probabilità che Maria le risponda è del 25% e la probabilità che Luisa le risponda è del 50%. Inoltre, la probabilità che Maria o Luisa le rispondano è una volta e mezza la probabilità che non le risponda nessuna delle due. Qual è la probabilità che sia Maria sia Luisa rispondano alla telefonata di Paola? [0,15]

128 Un'urna contiene 4 biglie, che hanno impressi i quattro numeri: $-4, -2, 3, 5$. Viene estratta una biglia dall'urna e si prende nota del numero impresso sulla biglia, che indichiamo con x. Dopo avere rimesso la biglia nell'urna, viene estratta una seconda biglia: diciamo y il numero impresso sulla seconda biglia estratta. Sia A l'evento «la somma di x e di y è un numero positivo» e B l'evento «il prodotto di x e di y è un numero positivo». Calcola la probabilità degli eventi: A, B, $A \cap B$ e $A \cup B$. $\left[\frac{5}{8}, \frac{1}{2}, \frac{1}{4}, \frac{7}{8}\right]$

129 Un'urna contiene sei palline numerate da 1 a 6. Si estraggono successivamente dall'urna tre palline, senza rimettere, a ogni estrazione, la pallina precedentemente estratta nell'urna. Considera i tre eventi:

A: «è uscita la pallina con il numero 2 nella prima estrazione»
B: «è uscita la pallina con il numero 3 nella seconda estrazione»
C: «non è uscita la pallina con il numero 4 in nessuna estrazione»

a. Calcola la probabilità degli eventi A, B e C.
b. Calcola la probabilità degli eventi $A \cap B$, $B \cap C$ e $A \cap C$.
c. Calcola la probabilità degli eventi $A \cup B$, $B \cup C$ e $A \cup C$. $\left[\textbf{a. } \frac{1}{6}, \frac{1}{6}, \frac{1}{2}; \textbf{ b. } \frac{1}{30}, \frac{1}{10}, \frac{1}{10}; \textbf{ c. } \frac{3}{10}, \frac{17}{30}, \frac{17}{30}\right]$

130 ESERCIZIO GUIDATO

Una password è costituita da due cifre, da 0 a 9, seguite da due lettere minuscole dell'alfabeto italiano. È ammessa la possibilità che le cifre o le lettere siano ripetute.

a. Quante password distinte di questo tipo si possono costruire?
b. Considera gli eventi A: «le due cifre della password sono diverse tra loro» e B: «le due lettere della password sono diverse tra loro» e determina la probabilità degli eventi:

$A, B, A \cap B, A \cup B$

a. Per il principio fondamentale del calcolo combinatorio, le password possibili sono: $10 \cdot ... \cdot 21 \cdot ...$

b. Per calcolare le probabilità dei tre eventi A, B e $A \cap B$, utilizza la definizione osservando che i casi *possibili* sono quelli determinati al punto **a.** mentre i casi *favorevoli*, per il principio fondamentale del calcolo combinatorio, sono:

- nel caso dell'evento A: $10 \cdot 9 \cdot ... \cdot ...$
- nel caso dell'evento B: $10 \cdot ... \cdot 21 \cdot ...$
- nel caso dell'evento $A \cap B$: $10 \cdot ... \cdot 21 \cdot ...$

Infine, puoi calcolare la probabilità di $A \cup B$, ricordando che $p(A \cup B) = p(A) + p(B) - p(A \cap B)$.

$\left[\textbf{a. } 44\,100; \textbf{ b. } \frac{9}{10}, \frac{20}{21}, \frac{6}{7}, \frac{209}{210}\right]$

131 Una password è formata da 6 numeri, ciascuno dei quali è una cifra tra 0 e 9 (0 e 9 inclusi), con la possibilità che i numeri siano ripetuti.

a. Quante password di questo tipo si possono generare?

La password del computer di Marco è del tipo descritto all'inizio. Scrivendo una password a caso, tra tutte quelle di cui al punto **a.**, qual è la probabilità che:

b. le prime due cifre della password scritta siano corrette;
c. le ultime tre cifre della password scritta siano corrette;
d. le prime due cifre o le ultime tre cifre della password scritta siano corrette.

Esprimi tutte le probabilità sotto forma di percentuale. [**a.** 10^6; **b.** 1%; **c.** 0,1%; **d.** 1,099%]

132 Un club ha 25 iscritti, di cui 9 donne e 16 uomini. Fra gli iscritti vengono scelte a caso tre persone che ricopriranno i ruoli di presidente, vicepresidente e segretario. Qual è la probabilità che il comitato direttivo scelto contenga almeno un uomo e almeno una donna? $\left[\frac{18}{25}\right]$

4. Probabilità composte ed eventi indipendenti

Teoria p. 147

Esercizi introduttivi

133 **Vero o falso?**

a. $p(A \cap B) = p(B) \cdot p(A|B)$ V F
b. $p(A \cap B) = p(B) \cdot p(B|A)$ V F
c. $p(A) \cdot p(B|A) = p(B) \cdot p(A|B)$ V F
d. due eventi incompatibili sono sempre indipendenti V F
e. due eventi indipendenti sono sempre incompatibili V F

[2 affermazioni vere e 3 false]

Test

134 Se $p(A) = \frac{1}{2}$ e $p(B|A) = \frac{1}{3}$, allora $p(A \cap B)$ è uguale a:

A $\frac{1}{4}$ B $\frac{1}{5}$ C $\frac{1}{6}$

D le informazioni non sono sufficienti per stabilirne il valore

135 Siano A e B due eventi indipendenti tali che:

$$p(A) = \frac{1}{2} \quad \text{e} \quad p(B) = \frac{2}{3}$$

Allora $p(A \cap B)$ è uguale a:

A $\frac{1}{2}$ B $\frac{1}{3}$ C $\frac{1}{6}$

D le informazioni non sono sufficienti per stabilirne il valore

136 Siano A e B due eventi indipendenti tali che:

$$p(A) = \frac{1}{2} \text{ e } p(\overline{B}) = \frac{1}{3}$$

Allora $p(A \cap B)$ è uguale a:

A $\frac{1}{2}$ B $\frac{1}{3}$ C $\frac{1}{6}$

D le informazioni non sono sufficienti per stabilirne il valore

137 Siano A e B due eventi. Si sa che $p(A) = \frac{1}{2}$, $p(B) = \frac{1}{3}$ e $p(A \cap B) = \frac{1}{5}$. Allora i due eventi A e B:

A sono incompatibili
B sono indipendenti
C non sono né incompatibili né indipendenti
D le informazioni date non sono sufficienti per stabilirlo

Probabilità condizionata e formula delle probabilità composte

138 ESERCIZIO SVOLTO

I coniugi Rossi hanno due figli. Sapendo che almeno uno dei due è maschio, qual è la probabilità che siano entrambi maschi?

1° modo

Per una coppia che ha due figli lo spazio campionario di riferimento è $\Omega = \{(F, F);\ (F, M);\ (M, F);\ (M, M)\}$. Avendo la coppia descritta nell'esercizio almeno un figlio maschio, bisogna passare a considerare come spazio campionario effettivo il sottoinsieme $\Omega' = \{(F, M);\ (M, F);\ (M, M)\}$. In Ω' vi è un solo evento che realizza il fatto che entrambi i figli siano maschi, quindi la probabilità richiesta è: $\dfrac{1}{3}$.

2° modo

Sia A l'evento «i due figli della coppia sono entrambi maschi» e B l'evento «almeno uno dei due figli della coppia è maschio». Si tratta di determinare $p(A|B)$. In base alla definizione:

$$p(A|B) = \frac{p(A \cap B)}{p(B)} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3}$$

139 Si è lanciato un dado regolare a sei facce e si sa che si è ottenuto un numero pari. Qual è la probabilità che il numero ottenuto sia maggiore di 2? $\left[\dfrac{2}{3}\right]$

140 Si è estratto un numero dal sacchetto della tombola. Sapendo che il numero estratto ha due cifre ed è pari, qual è la probabilità che esso sia il 50? $\left[\dfrac{1}{41}\right]$

141 Fra tutti i numeri naturali minori o uguali a 20 se ne estrae uno casualmente. Si sa che l'estratto è un numero primo. Qual è la probabilità che sia minore o uguale a 11? $\left[\dfrac{5}{8}\right]$

142 Si lanciano due dadi regolari a sei facce e si ottiene una somma dei due punteggi maggiore di 7. Qual è la probabilità di avere ottenuto due facce uguali? $\left[\dfrac{1}{5}\right]$

143 ESERCIZIO GUIDATO

La probabilità che Marta sia scelta per la recita della sua scuola è del 55%; se non viene scelta per la recita, Marta andrà al cinema con una probabilità del 40%. Qual è la probabilità che Marta non venga scelta per la recita e vada al cinema?

- Sia A l'evento «Marta viene scelta per la recita» e B l'evento «Marta va al cinema».
  In base ai dati del problema: $p(A) = 0{,}55$ e $p(B|\overline{A}) = 0{,}4$
- Devi calcolare $p(\overline{A} \cap B)$. In base alla regola delle probabilità composte: $p(\overline{A} \cap B) = p(\overline{A}) \cdot p(B|\overline{A})$
- Tenendo conto che: $p(\overline{A}) = 1 - p(A) = \dots$
  e che $p(B|\overline{A})$ è data, puoi concludere. [18%]

144 Maria lancia una moneta e, se esce «testa», l'indomani si farà interrogare in Italiano, altrimenti si farà interrogare in Matematica. Maria stima che la probabilità di prendere più di 7 nell'interrogazione è $\dfrac{2}{3}$ in Italiano e $\dfrac{1}{3}$ in Matematica.

a. Qual è la probabilità che l'indomani Maria si faccia interrogare in Italiano e prenda più di 7?

b. Qual è la probabilità che l'indomani Maria si faccia interrogare in Matematica e prenda più di 7? $\left[\text{a. } \dfrac{1}{3};\ \text{b. } \dfrac{1}{6}\right]$

145 Il 5% delle lampadine prodotte in una fabbrica sono difettose. La probabilità che una lampadina difettosa venga scartata è del 90%. Scelta a caso una lampadina, qual è la probabilità che sia difettosa ma non venga scartata? [0,5%]

146 Il 98% delle lampadine prodotte in una fabbrica sono prive di difetti. La probabilità che una lampadina difettosa venga scartata è del 40%. Scelta a caso una lampadina, qual è la probabilità che sia difettosa e venga scartata? [0,8%]

147 **ESERCIZIO GUIDATO**

Da un'urna contenente 6 palline rosse e 5 palline bianche si estraggono a caso successivamente 2 palline, senza rimpiazzo. Calcola la probabilità di estrarre due palline rosse, in due modi diversi:

a. utilizzando la formula delle probabilità composte;

b. utilizzando la definizione classica e il principio fondamentale del calcolo combinatorio.

a. Indichiamo con R_1 l'evento «la prima pallina estratta è rossa» e con R_2 l'evento «la seconda pallina estratta è rossa». Devi calcolare $p(R_1 \cap R_2)$. In base alla formula delle probabilità composte:

$$p(R_1 \cap R_2) = p(R_1) \cdot p(R_2|R_1) \qquad [*]$$

Poiché $p(R_1) = \frac{\dots}{11}$ e $p(R_2|R_1) = \frac{5}{\dots}$, sostituendo nella **[*]** puoi concludere facilmente.

b. Calcola la probabilità richiesta come rapporto tra casi favorevoli e casi possibili, ricordando il principio fondamentale del calcolo combinatorio per il calcolo dei casi favorevoli e di quelli possibili:

$$p(R_1 \cap R_2) = \frac{5 \cdot \dots}{11 \cdot \dots} = \frac{\dots}{\dots}$$

Metodi a confronto

148 Da un'urna contenente 8 palline rosse e 6 palline bianche si estraggono a caso successivamente 2 palline, senza rimpiazzo. Calcola la probabilità di estrarre due palline bianche, in due modi diversi:
a. utilizzando la formula delle probabilità composte;
b. utilizzando la definizione classica e il principio fondamentale del calcolo combinatorio. $\left[\frac{15}{91}\right]$

149 Un'urna contiene 12 biglie bianche e 8 nere. Calcola la probabilità che, estraendo tre biglie, una per volta, senza reimmissione, esse risultino tutte bianche, in due modi diversi:
a. utilizzando la formula delle probabilità composte;
b. utilizzando la definizione classica e il teorema fondamentale del calcolo combinatorio. $\left[\frac{11}{57}\right]$

150 Determina la probabilità che alla prossima estrazione del Lotto sulla ruota di Milano escano 5 numeri pari, in due modi diversi:
a. utilizzando la formula delle probabilità composte;
b. utilizzando la definizione classica e il teorema fondamentale del calcolo combinatorio. $\left[\frac{287}{10324}\right]$

Indipendenza e regola del prodotto

151 A e B sono due eventi tali che $p(A) = 20\%$ e $p(B) = 30\%$. Quanto deve valere $p(A \cap B)$ perché A e B siano indipendenti? [6%]

152 A e B sono due eventi indipendenti. Se $p(A) = 50\%$ e $p(A \cap B) = 40\%$, quanto vale la probabilità di B? [80%]

153 A e B sono due eventi. Si sa che $p(A) = \frac{1}{2}$, $p(B) = \frac{1}{5}$, $p(A \cup B) = \frac{3}{5}$. Gli eventi A e B sono indipendenti? [Sì; perché?]

154 Si estrae una carta da un mazzo di quaranta, si rimette la carta nel mazzo e si estrae una seconda carta. I due eventi: «la prima carta è un Asso» e «la seconda carta è un Asso» sono indipendenti? Dopo aver risposto intuitivamente, verifica la tua risposta in base alla regola del prodotto. [Sì]

155 Si estrae una carta da un mazzo di quaranta, quindi, senza rimettere la carta nel mazzo, se ne estrae una seconda. I due eventi: «la prima carta è una Regina» e «la seconda carta è un 2» sono indipendenti? [No]

156 Si estrae a caso una carta da un mazzo di 52. Gli eventi «la carta estratta è un fante» e «la carta estratta è di cuori» sono indipendenti? [Sì]

157 Si lancia un dado regolare a sei facce. Gli eventi: «è uscito un numero dispari» ed «è uscito un numero primo» sono indipendenti? [No]

158 Un dado regolare a sei facce viene lanciato una volta. Considera gli eventi:

A: «il numero uscito è dispari»

B: «il numero uscito è maggiore di n»

Per quali valori di $n \in \{1, 2, 3, 4, 5, 6\}$ gli eventi A e B sono indipendenti? [2, 4, 6]

159 Un dado regolare a sei facce viene lanciato una volta. Considera gli eventi:

A: «il numero uscito è pari»
B: «il numero uscito è minore di n»

Per quali valori di $n \in \{1, 2, 3, 4, 5, 6\}$ gli eventi A e B sono indipendenti? [1, 3, 5]

160 Una famiglia possiede due automobili. Ciascuna delle due auto, indipendentemente dall'altra, può trovarsi nella condizione di dover andare in officina con probabilità $\frac{5}{12}$. Qual è la probabilità che quella famiglia abbia entrambe le macchine in officina? $\left[\frac{25}{144}\right]$

161 Una famiglia possiede due televisori, il cui funzionamento è indipendente uno dall'altro. Per ciascuno dei due televisori, la probabilità di un guasto è $\frac{1}{1000}$. Qual è la probabilità che entrambi i televisori funzionino? [0,998001]

162 Per una persona bloccata sotto una valanga la probabilità di sopravvivere dopo un'ora è del 15%. Una squadra di soccorso raggiunge il luogo dove due escursionisti sono coperti da una valanga caduta un'ora prima. Supponendo che la sopravvivenza di ciascun escursionista sia indipendente dalla sopravvivenza dell'altro, qual è la probabilità di trovare in vita:

a. entrambi gli escursionisti;
b. almeno uno degli escursionisti. $\left[\text{a. } \frac{9}{400}; \text{ b. } \frac{111}{400}\right]$

163 Barbara il sabato sera va a mangiare in pizzeria con probabilità uguale al 95% e sceglie casualmente fra la pizzeria A e la pizzeria B. Paolo va a mangiare in pizzeria tutti i sabati sera, nella stessa ora di Barbara, scegliendo anche lui casualmente tra la pizzeria A e la pizzeria B. Qual è la probabilità che in un dato sabato Barbara e Paolo si incontrino nella pizzeria A? Qual è la probabilità che Paolo e Barbara si incontrino il prossimo sabato sera? $\left[\frac{19}{80}; \frac{19}{40}\right]$

Matematica e sport

164 **E se?** Alex e Bob, appassionati cestisti, hanno percentuali di realizzazione al tiro libero rispettivamente del 50% e del 60%. È più probabile che Alex centri due canestri su due tentativi o che Bob ne centri tre su tre? Assumi l'indipendenza dei tiri a canestro.

▸ Cambierebbe la risposta se la percentuale di realizzazione di Bob fosse del 65% invece che del 60%? [Alex; sì]

165 Una città ha una squadra di basket e una di calcio. La probabilità che la prima squadra vinca il suo campionato è del 25%, mentre la probabilità che la seconda vinca il suo è del 30%. Calcola la probabilità che:

a. entrambe le squadre vincano il campionato;
b. nessuna delle due squadre vinca il campionato;
c. almeno una delle due squadre vinca il campionato;
d. solo una delle due squadre vinca il campionato.

$\left[\text{a. } \frac{3}{40}; \text{ b. } \frac{21}{40}; \text{ c. } \frac{19}{40}; \text{ d. } \frac{2}{5}\right]$

166 Un'urna contiene 10 palline, di cui 6 bianche, numerate da 1 a 6 e 4 nere, numerate da 1 a 4. Si estraggono dall'urna due palline, simultaneamente.

a. Determina la probabilità dell'evento A: estrarre due palline bianche.
b. Determina la probabilità dell'evento B: estrarre due palline dello stesso colore.
c. Determina la probabilità dell'evento C: estrarre due palline che recano un numero dispari.
d. Gli eventi A e B sono indipendenti? Gli eventi A e C? Gli eventi B e C?

[a. $\frac{1}{3}$; b. $\frac{7}{15}$; c. $\frac{2}{9}$; d. nessuna delle tre coppie i eventi è costituita da eventi indipendenti]

167 Siano A e B due eventi indipendenti, tali che $p(A \cup B) = \frac{4}{5}$ e $p(\overline{B}) = \frac{3}{5}$. Calcola la probabilità dell'evento A. $\left[\frac{2}{3}\right]$

168 Paolo si reca in biblioteca per prendere in prestito due libri di suo interesse. La probabilità che il primo libro che desidera Paolo sia già in prestito è 0,4; la probabilità che il secondo libro che desidera Paolo sia già in prestito è 0,3. La probabilità che almeno uno dei due libri sia già in prestito è p.

a. Esprimi in funzione di p la probabilità che entrambi i libri siano già in prestito.

b. Determina per quale valore di p i due eventi «il primo libro è già in prestito» e «il secondo libro è già in prestito» sono indipendenti.

c. In corrispondenza del valore di p determinato al punto **b.**, determina la probabilità che esattamente uno dei due libri sia già in prestito. [**a.** $0{,}7 - p$; **b.** $p = 0{,}58$; **c.** 0,46]

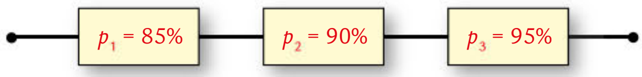
169 Un'urna contiene 20 palline bianche ed n palline nere, con $n \in \mathbf{N}$ ed $n \geq 2$. Un giocatore estrae per 10 volte, consecutivamente, una pallina dall'urna, rimettendo ogni volta la pallina estratta nell'urna prima dell'estrazione successiva. Determina il minimo valore di n per cui la probabilità di estrarre, nelle dieci estrazioni, almeno una pallina nera sia superiore al 99%. [$n = 12$]

Matematica e affidabilità di un sistema

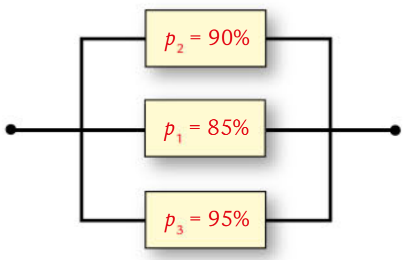
Nella risoluzione degli esercizi 170-176 tieni presente che si dice «affidabilità» di un sistema la probabilità che il sistema funzioni correttamente.

170 Un sistema è formato da tre componenti indipendenti, le cui probabilità di funzionamento sono $p_1 = 85\%$, $p_2 = 90\%$, $p_3 = 95\%$. Determina l'affidabilità del sistema in ciascuno dei seguenti casi:

a. se i singoli componenti sono tutti collegati in serie:



b. se i singoli componenti sono tutti collegati in parallelo:



Arrotonda i risultati alla terza cifra e decimale e scrivili sotto forma di percentuale.

(*Suggerimento*: se i componenti sono collegati in serie, il sistema risulta funzionante se e solo se lo sono *tutti* i componenti, mentre se sono collegati in parallelo il sistema risulta funzionante se e solo se *almeno uno* di essi lo è)

[**a.** 72,7%; **b.** 99,9%]

171 Un sistema è formato da n componenti identici e indipendenti, ciascuno dei quali ha una probabilità di funzionamento uguale a p.

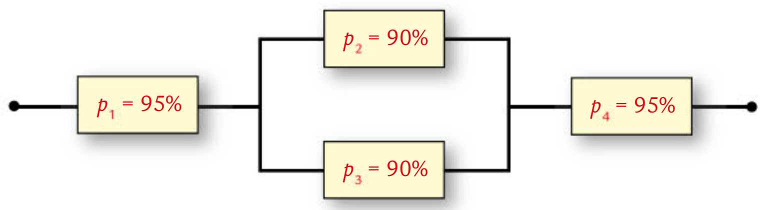
a. Esprimi in funzione di p ed n l'affidabilità del sistema, se i singoli componenti sono tutti collegati in parallelo.

b. Supposto $p = 80\%$, determina il minimo numero di componenti da cui deve essere formato il sistema affinché la sua affidabilità sia maggiore del 99%. [**a.** $1 - (1 - p)^n$; **b.** $n = 3$]

172 Due componenti indipendenti, se posti in serie, formano un sistema la cui affidabilità è uguale al 72%; gli stessi componenti, se posti in parallelo, formano un sistema la cui probabilità di funzionamento è uguale al 98%. Determina la probabilità di funzionamento dei singoli componenti. [80%, 90%]

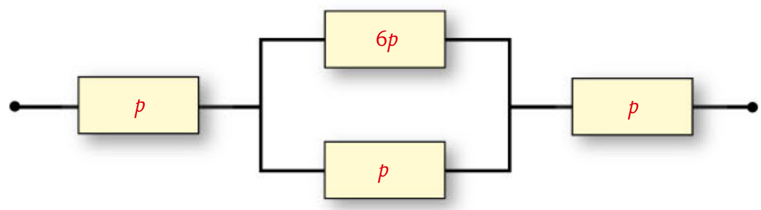
173 Due componenti indipendenti, se posti in serie, formano un sistema la cui probabilità di guasto è uguale al 32%; gli stessi componenti, se posti in parallelo, formano un sistema la cui probabilità di guasto è uguale al 3%. Determina l'affidabilità dei singoli componenti. [80%, 85%]

174 In figura è rappresentato un sistema formato da quattro componenti indipendenti; sono indicate le affidabilità dei singoli componenti.



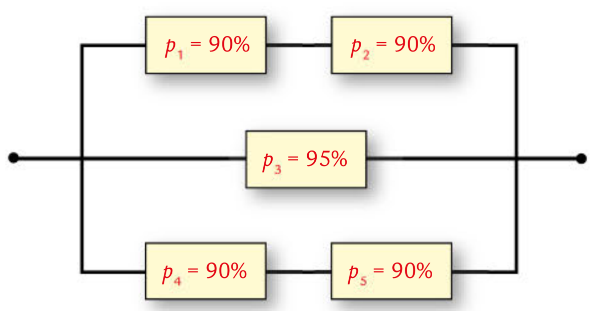
Determina la probabilità di funzionamento del sistema. Scrivi il risultato sotto forma di percentuale, arrotondato con due cifre decimali. [89,35%]

175 In figura è rappresentato un sistema formato da quattro componenti indipendenti; sono indicate le affidabilità dei singoli componenti.



a. Determina, in funzione di p, la probabilità di funzionamento del sistema.
b. Determina per quale valore di p la probabilità di funzionamento del sistema è massima. $\left[\textbf{a. } 7p^3 - 6p^4; \textbf{ b. } p = \frac{7}{8}\right]$

176 In figura è rappresentato un sistema formato da cinque componenti indipendenti; sono indicate le affidabilità dei singoli componenti.



Il sistema risulta funzionante se lo è almeno uno dei tre blocchi che sono stati posti in parallelo. Determina la probabilità di funzionamento del sistema. Scrivi il risultato sotto forma di percentuale, arrotondato con due cifre decimali. [99,82%]

Prove ripetute

177 ESERCIZIO SVOLTO

Barbara tutte le mattine al sabato prende l'autobus per recarsi a scuola. La probabilità che l'autobus sia in ritardo è uguale a 0,1. Calcoliamo la probabilità che durante una settimana (escludendo la domenica in cui Barbara non prende l'autobus):

a. l'autobus sia in ritardo esattamente due volte;

b. l'autobus sia in ritardo al massimo due volte;

c. l'autobus sia in ritardo almeno una volta.

Possiamo utilizzare il modello delle prove ripetute dove l'evento «l'autobus è in ritardo» corrisponde a un «successo» con probabilità $p = 0{,}1$ (supponendo indipendenti gli eventuali ritardi nei vari giorni della settimana).

a. Dobbiamo calcolare la probabilità di avere 2 successi:

$$\binom{6}{2}(0{,}1)^2(1-0{,}1)^{6-2} = \binom{6}{2}(0{,}1)^2(0{,}9)^4 \simeq 0{,}098 = 9{,}8\%$$

b. Dobbiamo calcolare la probabilità di avere *al massimo* due successi, cioè di averne 0 oppure 1 oppure 2; la probabilità cercata è quindi:

$$\underbrace{\binom{6}{0}(0{,}1)^0(0{,}9)^6}_{\text{probabilità di 0 successi}} + \underbrace{\binom{6}{1}(0{,}1)^1(0{,}9)^5}_{\text{probabilità di 1 successo}} + \underbrace{\binom{6}{2}(0{,}1)^2(0{,}9)^4}_{\text{probabilità di 2 successi}} \simeq 0{,}984 = 98{,}4\%$$

c. Conviene utilizzare il passaggio all'evento contrario; la probabilità cercata è 1 meno la probabilità che l'autobus sia sempre in orario, dunque è uguale a:

$$1-(0{,}9)^6 \simeq 0{,}469 = 46{,}9\%$$

178 **ESERCIZIO GUIDATO**

Si lancia una moneta regolare per 6 volte, successivamente. Qual è la probabilità che esca «testa» esattamente 4 volte?

Utilizza il modello delle prove ripetute dove l'evento «è uscita testa» corrisponde a un «successo» con probabilità $p = \frac{1}{2}$. La probabilità di avere 4 successi in 6 prove ripetute è uguale a:

$$\binom{6}{4}\left(\frac{1}{2}\right)^{\dots}\left(1-\frac{1}{2}\right)^{\dots} = \dots$$

179 Si lancia un dado regolare a sei facce e si vince se esce il 3 o il 6.

a. Qual è la probabilità di vincere?

b. Qual è la probabilità che, lanciando il dado 4 volte, si vinca esattamente due volte?

c. Qual è la probabilità che, lanciando il dado 4 volte, si vinca esattamente 3 volte?

$\left[\text{a. } \frac{1}{3}; \text{ b. } \frac{8}{27}; \text{ c. } \frac{8}{81}\right]$

180 Da un'urna contenente 20 palline, di cui 5 bianche e 15 nere, si effettuano quattro estrazioni successive, con reimmissione. Qual è la probabilità di estrarre esattamente 2 palline bianche?

$\left[\frac{27}{128}\right]$

181 Da un mazzo di 32 carte (contenente 8 carte di ogni seme) vengono estratte successivamente cinque carte, con reimmissione. Qual è la probabilità di estrarre esattamente tre carte di cuori?

$\left[\frac{45}{512}\right]$

Realtà e modelli

182 **Menu al ristorante.** Quattro persone sono sedute al tavolo di un ristorante, in procinto di ordinare un primo piatto. Per ciascuna delle quattro persone, la probabilità di scegliere il risotto ai funghi è uguale a 0,2. Qual è la probabilità che esattamente due delle quattro persone scelgano il risotto ai funghi, supponendo che ciascuno scelga indipendentemente dagli altri?

$\left[\frac{96}{625}\right]$

183 **Videogiochi.** Paolo sta giocando a un videogioco e deve colpire il nemico. Ogni volta che spara un colpo, la probabilità di colpirlo è uguale a 0,4. Se spara cinque colpi, qual è la probabilità di colpire il nemico esattamente tre volte?

$\left[\frac{144}{625}\right]$

184 **Numero di partecipanti a una riunione.** Il comitato di un'associazione è composto da 8 persone. Ciascun membro del comitato partecipa alle riunioni, indipendentemente dagli altri, con una probabilità del 50%. Calcola la probabilità che alla prossima riunione:

a. tutti i membri del comitato siano presenti;

b. siano presenti al massimo sei membri del comitato;

c. sia presente almeno un membro del comitato.

$\left[\mathbf{a.}\ \frac{1}{256};\ \mathbf{b.}\ \frac{247}{256};\ \mathbf{c.}\ \frac{255}{256}\right]$

185 **Guasti su una linea ferroviaria.** La probabilità che in un mese si verifichi un guasto su una certa linea ferroviaria è uguale a 0,1. Supponendo i guasti indipendenti gli uni dagli altri, calcola la probabilità che in un anno:

a. non si verifichi alcun guasto; **b.** si verifichi almeno un guasto; **c.** si verifichino esattamente due guasti.

Per ciascuna probabilità, scrivi l'espressione che la esprime, quindi, con l'aiuto di una calcolatrice, fornisci il risultato arrotondato a meno di un centesimo.

$\left[\mathbf{a.}\ \left(\frac{9}{10}\right)^{12} \simeq 0{,}28;\ \mathbf{b.}\ 1-\left(\frac{9}{10}\right)^{12} \simeq 0{,}72;\ \mathbf{c.}\ \frac{22 \cdot 3^{21}}{10^{12}} \simeq 0{,}23\right]$

186 **Test a risposta multipla.** In un compito in classe Paolo deve rispondere a 5 quesiti a risposta multipla: ogni quesito è costituito da 4 risposte, di cui una sola è quella esatta. Il test è superato rispondendo correttamente ad almeno 3 domande. Essendo del tutto impreparato, Paolo risponde a caso a tutti i quesiti. Qual è la probabilità che Paolo superi il test? [Circa 10,4%]

187 **Overbooking.** La «sovraprenotazione», meglio nota come «overbooking», è una strategia adottata dalle compagnie aeree: consiste nel vendere in prenotazione più posti a sedere di quelli effettivamente disponibili, perché è improbabile che tutti i passeggeri si presentino al check in. Consideriamo un volo che possa ospitare 150 passeggeri e supponiamo che vengano venduti 154 biglietti, cioè 4 posti in più di quelli disponibili. Supponiamo inoltre che la probabilità che un passeggero non si presenti all'imbarco sia del 5% e che i passeggeri si presentino all'aeroporto indipendentemente gli uni dagli altri. Qual è la probabilità che nessun passeggero sia in sovrannumero, al momento dell'imbarco? [Circa 95,22%]

188 **Telecomunicazioni.** In un canale di trasmissione binario un messaggio viene codificato in una «parola» costituita da una sequenza di 10 bit. La probabilità che un singolo bit venga trasmesso correttamente è del 90%. Il codice è in grado di correggere al massimo 2 errori di trasmissione. Assumendo che eventuali errori di trasmissione dei singoli bit siano indipendenti, calcola la probabilità che la trasmissione della parola codificata, cioè della sequenza di 10 bit, avvenga correttamente. [Circa 93%]

Matematica e informatica

189 Il server web di una azienda può avere dei malfunzionamenti dovuti al sovraccarico, nel caso di un numero troppo elevato di richieste. La probabilità che in una data giornata si verifichi un malfunzionamento dovuto al sovraccarico è 0,02. Supponendo che malfunzionamenti che avvengono in giorni distinti siano indipendenti tra loro, qual è la probabilità che nel corso del mese di giugno, si verifichino più di 2 malfunzionamenti? [Circa 2,2%]

190 Un calcolatore è collegato a una rete che consente l'accesso a un massimo di 50 utenti. Gli operatori che possono richiedere l'accesso alla rete sono 53 e la probabilità che, a un dato istante, un operatore richieda l'accesso è 0,8. Supponendo che le richieste di accesso dei singoli operatori siano indipendenti l'una dall'altra, determina la probabilità che, a un dato istante, la rete sia satura (ossia tutti gli accessi disponibili siano in uso).

[Circa 0,34%]

5. Il teorema di disintegrazione e la formula di Bayes

Teoria p. 154

Esercizi introduttivi

Test

191 Sapendo che $p(A) = \frac{1}{3}$, $p(X|A) = \frac{3}{4}$ e $p(X|\overline{A}) = \frac{1}{4}$, quanto vale $p(X)$?

A $\frac{5}{12}$ B $\frac{3}{4}$ C $\frac{7}{12}$ D I dati sono insufficienti per determinarla.

192 Sapendo che $p(A) = \frac{1}{2}$, $p(B) = \frac{1}{4}$, quale delle seguenti uguaglianze è certamente vera?

A $p(A|B) = 2p(B|A)$ B $p(A|B) = 3p(B|A)$ C $p(A|B) = 4p(B|A)$ D Nessuna delle precedenti

193 Se $p(B|A) = \frac{1}{3}$, $p(A) = \frac{1}{2}$ e $p(B) = \frac{1}{4}$, allora $p(A|B)$ è uguale a:

A $\frac{1}{3}$ B $\frac{2}{3}$ C $\frac{1}{4}$ D $\frac{3}{4}$

194 Se $p(X) = \frac{13}{24}$, $p(A) = \frac{1}{6}$ e $p(X|\overline{A}) = \frac{1}{2}$, allora $p(X|A)$ è uguale a:

A $\frac{1}{4}$ B $\frac{3}{8}$ C $\frac{3}{4}$ D i dati sono insufficienti per determinarla

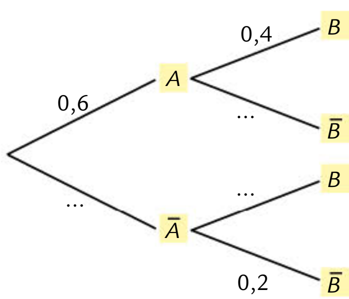
195 Se $p(A|B) = \frac{2}{3}$, $p(B|A) = \frac{1}{3}$ e $p(A) = \frac{1}{2}$, allora $p(B)$ è uguale a:

A $\frac{1}{4}$ B $\frac{3}{8}$ C $\frac{3}{4}$ D i dati sono insufficienti per determinarla

Interpretazione di grafici

196 Completa il diagramma ad albero qui sotto, scrivendo al posto dei puntini le probabilità mancanti. Utilizzando il diagramma ad albero, determina poi la probabilità degli eventi seguenti:

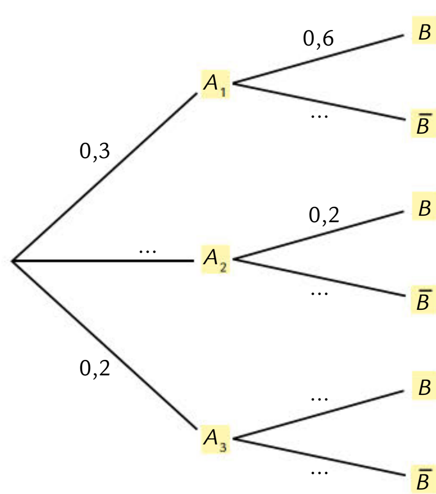
a. $A \cap \overline{B}$
b. $\overline{A} \cap B$
c. B



[**a.** 0,36; **b.** 0,32; **c.** 0,56]

197 Sapendo che A_1, A_2, A_3 costituiscono una partizione dello spazio campionario e $p(A_3 \cap B) = 0{,}06$, completa il diagramma ad albero qui sotto, scrivendo al posto dei puntini le probabilità mancanti. Utilizzando il diagramma ad albero, calcola poi la probabilità degli eventi seguenti:

a. $A_1 \cap \overline{B}$
b. $A_2 \cap \overline{B}$
c. $A_3 \cap B$
d. B



[**a.** 0,12; **b.** 0,4; **c.** 0,06; **d.** 0,34]

Problemi sul teorema di disintegrazione

198 ESERCIZIO GUIDATO

Si hanno a disposizione due monete. Una delle due è regolare, mentre l'altra è truccata in modo che la probabilità che esca «testa» sia $\frac{1}{3}$. Si sceglie a caso una delle due monete e si lancia. Qual è la probabilità che esca «testa»?

- Considera gli eventi:

  A: «la moneta è regolare» B: «la moneta è truccata» T: «esce testa»

Devi calcolare $p(T)$

- Per il teorema di disintegrazione:

  $p(T) = p(T|A) \cdot p(A) + p(T|B) \cdot p(B) = \dots\dots$ $\left[\frac{5}{12}\right]$

199 Si hanno a disposizione due monete. Una delle due è regolare, mentre l'altra è truccata in modo che la probabilità che esca «croce» sia $\frac{2}{5}$. Si sceglie a caso una delle due monete e si lancia. Qual è la probabilità che esca «testa»? $\left[\frac{11}{20}\right]$

200 In un'urna A sono contenute 5 palline bianche e 5 palline nere, mentre in un'urna B, sono contenute 4 palline bianche e 6 palline nere. Si sceglie a caso un'urna e si pesca una pallina. Qual è la probabilità che sia nera? $\left[\frac{11}{20}\right]$

201 In un'urna A sono contenute 9 palline, numerate da 1 a 9. In un'urna B sono contenute 6 palline, numerate da 1 a 6. Scelta a caso un'urna, qual è la probabilità di estrarre una pallina che reca un numero multiplo di 3? $\left[\frac{1}{3}\right]$

202 Due tenenti di polizia, Colombo e Sheridan, si alternano in maniera casuale nel servizio alla centrale di polizia: il primo è di turno tre giorni su sette, il secondo quattro giorni su sette. Vale la regola che un poliziotto si occupa solo dei reati che si verificano durante il suo turno di servizio. Sapendo che Colombo risolve 8 casi su 10 e Sheridan 6 su 10, qual è la probabilità, per un malvivente che commette un reato, di restare impunito? $\left[\frac{11}{35} \simeq 31{,}43\%\right]$

203 Il pilota di formula 1 Alfonso partecipa a una gara. Secondo gli esperti, la probabilità che Alfonso vinca la gara è dell'80% in caso di pioggia e del 60% nel caso che non piova. Il servizio meteorologico prevede che, sul circuito, per il periodo della corsa, la probabilità di avere pioggia è del 65%. Qual è la probabilità che ha Alfonso di vincere la gara? [73%]

204 Si hanno due urne; la prima contiene 10 palline bianche e 8 verdi, la seconda 6 bianche e 4 gialle. Si estrae un numero dal sacchetto della tombola: se questo è minore o uguale a 50 si estrae una pallina dalla prima urna, in caso contrario si estrae una pallina dal secondo contenitore.

a. Qual è la probabilità che la pallina estratta sia bianca?
b. Qual è la probabilità che la pallina estratta sia gialla?
c. Qual è la probabilità che la pallina estratta sia verde? $\left[\mathbf{a.}\ \frac{233}{405};\ \mathbf{b.}\ \frac{8}{45};\ \mathbf{c.}\ \frac{20}{81}\right]$

205 Si dispone di 3 scatole identiche A, B, C. La scatola A contiene 10 lampadine, 4 delle quali sono difettose. La scatola B contiene 6 lampadine, 2 delle quali sono difettose. La scatola C contiene 8 lampadine, 3 delle quali sono difettose. Le lampadine difettose sono indistinguibili dalle altre.

a. Da una delle scatole, scelta a caso, si estrae, anch'essa a caso, 1 lampadina. Qual è la probabilità che la lampadina estratta sia difettosa?
b. Se invece di procedere come indicato in **a.** si estrae a caso una lampadina da ciascuna delle scatole, qual è la probabilità che due delle lampadine estratte siano funzionanti e l'altra difettosa?

$\left[\mathbf{a.}\ \frac{133}{360};\ \mathbf{b.}\ \frac{53}{120}\right]$

Matematica e telecomunicazioni

206 Un segnale binario (cioè un segnale che può assumere solo i valori 0 e 1) è trasmesso in un canale di trasmissione in cui la probabilità che un segnale trasmesso come 0 sia ricevuto correttamente è del 90% mentre la probabilità che un segnale trasmesso come 1 sia ricevuto correttamente è del 95%. Se si verifica un errore nella trasmissione, il segnale viene invertito (cioè il segnale trasmesso come 0 viene ricevuto come 1 e il segnale trasmesso come 1 viene ricevuto come 0). La probabilità di trasmettere il segnale 0 è del 60% e quella di trasmettere il segnale 1 è del 40%. Calcola:

a. la probabilità totale di errore (cioè dell'evento «trasmetto 0 e ricevo 1 oppure trasmetto 1 e ricevo 0»);

b. la probabilità che un segnale trasmesso venga ricevuto come 1. [**a.** 8%; **b.** 44%]

207 Un segnale binario (cioè un segnale che può assumere solo i valori 0 e 1) è trasmesso in un canale di trasmissione in cui la probabilità che venga commesso un errore nella trasmissione di un segnale emesso come 0 è e_0, mentre la probabilità che venga commesso un errore nella trasmissione di un segnale emesso come 1 è e_1. Assumi che eventuali errori di trasmissione siano indipendenti. La probabilità che un segnale sia emesso con valore 0 è p. Scrivi le espressioni che esprimono:

a. la probabilità che la sequenza 1010 sia trasmessa correttamente;

b. la probabilità che un singolo segnale binario sia trasmesso correttamente.

[**a.** $(1-e_0)^2(1-e_1)^2$; **b.** $(1-e_0)p+(1-e_1)(1-p)$]

208 Si hanno a disposizione:

– due urne A e B, tali che l'urna A contiene 3 palline bianche e 5 nere mentre l'urna B contiene 6 palline bianche e 4 nere;

– una moneta truccata, per cui la probabilità di ottenere testa in un lancio è uguale a p.

Un gioco consiste nel lanciare la moneta truccata, quindi estrarre una pallina dall'urna A se è uscita testa, dall'urna B se è uscita croce. Il giocatore vince se la pallina estratta è nera.

a. Determina la probabilità che il giocatore vinca.

b. In corrispondenza di quale valore di p la probabilità di vincere è uguale a quella di perdere?

$$\left[\textbf{a. } \frac{9}{40}p+\frac{2}{5};\ \textbf{b. } p=\frac{4}{9}\right]$$

209 Si hanno due urne, U_1 e U_2, tali che:

– l'urna U_1 contiene n palline bianche e 3 nere;

– l'urna U_2 contiene 2 palline bianche e 1 pallina nera.

Un gioco consiste nell'estrarre a caso una pallina da U_1, porre la pallina estratta in U_2 ed estrarre una pallina da U_2. Il giocatore vince se la pallina estratta da U_2 è bianca.

a. Calcola la probabilità che il giocatore vinca.

b. Determina il minimo valore di n per cui la probabilità del giocatore di vincere è superiore al 70%.

$$\left[\textbf{a. } \frac{3n+6}{4n+12};\ \textbf{b. } n=13\right]$$

210 Abbiamo due dadi cubici: il dado A presenta quattro facce bianche e due facce nere. Il dado B presenta una faccia bianca, due facce nere e tre facce rosse. Si lancia il dado A: se la faccia ottenuta è bianca, si lancia ancora il dado A, se la faccia ottenuta è nera, si lancia il dado B.

Calcola la probabilità:

a. di ottenere una faccia nera nel secondo lancio, sapendo che è stata ottenuta una faccia nera nel primo;

b. di ottenere due facce nere;

c. di ottenere una faccia bianca al secondo lancio.

$$\left[\textbf{a. } \frac{1}{3};\ \textbf{b. } \frac{1}{9};\ \textbf{c. } \frac{1}{2}\right]$$

211 ESERCIZIO GUIDATO

Un'azienda acquista componenti elettronici da tre fornitori, diciamo A, B, C. La metà dei componenti elettronici viene acquistata dal fornitore A, il 20% dal fornitore B e i restanti dal fornitore C. Si stima inoltre che:

- **l'1% dei componenti acquistati dal fornitore A è difettoso;**
- **il 4% dei componenti acquistati dal fornitore B è difettoso;**
- **il 3% dei componenti complessivamente acquistati dai tre fornitori sono difettosi.**

Scelto a caso un componente tra quelli complessivamente acquistati, calcola la probabilità che:

a. **sia stato acquistato dal fornitore A e sia difettoso;**
b. **sia stato acquistato dal fornitore B e sia difettoso;**
c. **sia stato acquistato dal fornitore C e sia difettoso;**
d. **sia stato acquistato dal fornitore A, sapendo che è difettoso.**
e. **sia stato acquistato dal fornitore A o dal fornitore B, sapendo che è difettoso.**

Considera gli eventi:

A: «il componente è stato acquistato dal fornitore A»
B: «il componente è stato acquistato dal fornitore B»
C: «il componente è stato acquistato dal fornitore C»
D: «il componente è difettoso»

In base ai dati è noto che:

$$p(A) = 0{,}5 \qquad p(B) = 0{,}2 \qquad p(C) = \dots\dots \qquad p(D|A) = 0{,}01 \qquad p(D|B) = 0{,}04 \qquad p(D) = 0{,}03$$

a. Devi calcolare:

$$p(A \cap D) = p(A) \cdot p(D|A) = \dots\dots$$

b. Puoi procedere analogamente al punto **a.**

c. Per il teorema di disintegrazione:

$$p(D) = p(A \cap D) + p(B \cap D) + p(C \cap D)$$

Da questa relazione puoi ricavare il valore di $p(C \cap D)$, tenendo conto che $p(D)$ è noto per ipotesi, mentre $p(A \cap D)$ e $p(B \cap D)$ sono state ricavate nei punti precedenti.

d. Devi calcolare $p(A|D) = \dfrac{p(A \cap D)}{p(D)} = \dots\dots$

e. Devi calcolare:

$$p(A \cup B|D) = \frac{p((A \cup B) \cap D)}{p(D)} \underset{\substack{\uparrow \\ \text{Proprietà distributiva} \\ \text{dell'intersezione} \\ \text{rispetto all'unione}}}{=} \frac{p((A \cap D) \cup (B \cap D))}{p(D)} \underset{\substack{\uparrow \\ \text{Gli eventi} \\ A \cap D \text{ e } B \cap D \\ \text{sono incompatibili}}}{=} \frac{p(A \cap D) + p(B \cap D)}{p(D)} = \dots$$

$\left[\textbf{a. } \dfrac{1}{200}; \textbf{ b. } \dfrac{1}{125}; \textbf{ c. } \dfrac{17}{1000}; \textbf{ d. } \dfrac{1}{6}; \textbf{ e. } \dfrac{13}{30}\right]$

212 Il personale relativo al settore di produzione di un'azienda è composto da ingegneri, tecnici e addetti alla manutenzione. Gli ingegneri e i tecnici costituiscono rispettivamente il 15% e il 75% del personale del settore di produzione. Tra gli ingegneri le donne sono il 50%, tra i tecnici le donne sono il 20% e tra gli addetti alla manutenzione le donne sono il 40%. Scelto a caso un dipendente dell'azienda del settore di produzione, qual è la probabilità:

a. che sia un addetto alla manutenzione;
b. che sia una donna addetta alla manutenzione;
c. che sia un tecnico di sesso maschile;
d. che sia una donna;
e. che sia un uomo.

$\left[\textbf{a. } \dfrac{1}{10}; \textbf{ b. } \dfrac{1}{25}; \textbf{ c. } \dfrac{3}{5}; \textbf{ d. } \dfrac{53}{200}; \textbf{ e. } \dfrac{147}{200}\right]$

213 ESERCIZIO SVOLTO

Alex e Bob lanciano due dadi, in un gioco in cui vince chi per primo ottiene un doppio 6. Se Alex lancia per primo, qual è la probabilità che sia lui a vincere la sfida?

Sia A l'evento «Vince Alex» e B l'evento «al primo lancio si ottiene un doppio 6».
Per il teorema di disintegrazione possiamo scrivere la seguente relazione:

$$\underbrace{p(A)}_{\text{probabilità che Alex vinca}} = \underbrace{p(A\,|\,B)}_{\text{probabilità che Alex vinca, se al primo lancio esce un doppio 6}} \cdot \underbrace{p(B)}_{\text{probabilità che nel primo lancio esca un doppio 6}} + \underbrace{p(A\,|\,\overline{B})}_{\text{probabilità che Alex vinca, se al primo lancio non esce un doppio 6}} \cdot \underbrace{p(\overline{B})}_{\text{probabilità che nel primo lancio non esca un doppio 6}} \qquad [*]$$

Poniamo ora $p(A) = x$ (la probabilità che ci interessa calcolare) e osserviamo che:

- ovviamente $p(A|B) = 1$, $p(B) = \dfrac{1}{36}$ e $p(\overline{B}) = 1 - \dfrac{1}{36} = \dfrac{35}{36}$;
- $p(A|\overline{B})$, dal momento che stiamo assumendo che il doppio 6 non esca nel primo lancio, equivale alla probabilità di vittoria di Alex in un gioco in cui è Bob a iniziare per primo, ed è perciò uguale a $1 - x$ (infatti, se Bob inizia per primo, si trova nella identica situazione di Alex, quindi la sua probabilità di vittoria è x, mentre quella di sconfitta, ovvero di vittoria di Alex, è $1 - x$).

La **[*]** si traduce quindi nell'equazione:

$$x = 1 \cdot \frac{1}{36} + (1 - x)\frac{35}{36}$$

che risolta fornisce $x = \dfrac{36}{71}$. Risulta $\dfrac{36}{71} > 50\%$: il fatto che Alex inizi per primo, come è intuitivo aspettarsi, lo rende favorito.

214 Anna e Francesco si disputano l'ultimo trancio di pizza a tavola. Francesco propone un onesto «pari e dispari», mentre Anna furbescamente lo raggira: «Lanciamo a turno la moneta in attesa della prima 'testa' a indicare il vincitore. Comincio io...». Qual è la probabilità che Anna si aggiudichi lo spizzico rimasto? $\left[\dfrac{2}{3}\right]$

215 Alex e Bob, appassionati di basket, gareggiano ai tiri liberi: alternandosi alla «lunetta», vince il primo che a parità di tentativi segna un canestro in più dell'altro. Sapendo che Alex ha il 60% di probabilità di centrare il canestro, mentre Bob il 40%, qual è la probabilità che Alex vinca la sfida? $\left[\dfrac{9}{13}\right]$

216 **E se?** Riferisciti all'esercizio precedente. La gara tra Alex e Bob, ora, prevede che a vincere sia il primo che realizza il proprio tiro libero, non importa se a parità di tentativi effettuati (ovviamente ciò avvantaggia chi comincia per primo). Qual è la probabilità che Alex vinca la sfida, se è proprio Alex il primo ad andare in «lunetta»?

- Se Alex cede a Bob il vantaggio di iniziare la serie, come cambia la probabilità di vittoria per Alex? E se invece i due concordano che sia il lancio di una moneta a stabilire chi avvierà la serie di tiri liberi? $\left[\dfrac{15}{19}; \dfrac{9}{19}; \dfrac{12}{19}\right]$

Problemi sulla formula di Bayes

217 ESERCIZIO GUIDATO

Nella dispensa sono disposte due ceste di mele. Nella prima cesta il 70% sono mature, mentre le altre sono acerbe; nella seconda cesta il 90% sono mature, mentre le altre sono acerbe. Si sceglie a caso una delle due ceste e si estrae da essa una mela a caso.

a. Qual è la probabilità che la mela estratta sia matura?

b. Sapendo che la mela è risultata matura, qual è la probabilità che provenga dalla prima cesta?

Considera gli eventi:

C_1: «ho scelto la prima cesta» C_2: «ho scelto la seconda cesta»

M: «ho scelto una mela matura»

In base ai dati è noto che:

$p(M|C_1) = 0{,}7$ $p(M|C_2) = \ldots\ldots$ $\underbrace{p(C_1) = p(C_2) = 0{,}5}_{\text{la cesta viene scelta a caso, quindi ciascuna delle due ha la stessa probabilità di essere scelta}}$

a. Per il teorema di disintegrazione: $p(M) = p(M|C_1) \cdot p(C_1) + p(M|C_2) \cdot p(C_2) = \ldots\ldots$

b. Per la formula di Bayes: $p(C_1|M) = \dfrac{p(M|C_1) \cdot p(C_1)}{p(M)} = \ldots\ldots$

$\left[\text{a. } \frac{4}{5}; \text{ b. } \frac{7}{16}\right]$

Matematica e controllo qualità

218 Un'azienda produce delle penne. La probabilità che una penna sia difettosa è uguale al 5%. Il controllo di qualità accetta tutte le penne senza difetti e scarta il 90% delle penne difettose. Scelta a caso una penna, calcola la probabilità:

a. che superi il controllo di qualità;

b. che sia difettosa, sapendo che ha superato il controllo di qualità.

$\left[\text{a. } \frac{191}{200}; \text{ b. } \frac{1}{191}\right]$

219 Un'azienda produce un bene che, prima di essere immesso sul mercato, viene sottoposto a un test di qualità. Se il bene è in perfette condizioni, supera sempre il test, mentre se presenta qualche difetto supera il test solo nel 5% dei casi. Supponiamo che il 2% della produzione presenti qualche difetto. Se una unità del bene supera il test, qual è la probabilità che sia ugualmente difettosa?

$\left[\frac{1}{981} \simeq 0{,}1\%\right]$

220 Un'azienda produce delle valvole di sicurezza per pentole a pressione, che vengono sottoposte a un test di qualità prima di essere messe in commercio. Se una valvola è perfettamente funzionante, supera certamente il test, mentre se presenta qualche difetto supera il test soltanto nel 4% dei casi. È noto inoltre, da statistiche precedentemente effettuate, che il 2% delle valvole prodotte risultano difettose. Ogni giorno l'azienda produce 5000 valvole, che vengono sottoposte nel giorno stesso al test di qualità; mediamente, quante valvole supereranno il test in una giornata? E quante, tra queste ultime, risulteranno in realtà difettose, nonostante abbiano superato il test? [4904; 4]

221 Un pezzo viene costruito assemblando due componenti, diciamo A e B. La probabilità che il componente A sia difettoso è il 4%; la probabilità che il componente B sia difettoso è il 5% ed è indipendente dal componente A. Il pezzo assemblato risulterà difettoso se e solo se almeno uno dei due componenti lo è.

a. Qual è la probabilità che il pezzo assemblato risulti difettoso?

b. Se un pezzo assemblato è difettoso, qual è la probabilità che il componente B sia difettoso?

c. Se un pezzo assemblato è difettoso, qual è la probabilità che soltanto il componente B sia difettoso?

[**a.** 8,8%; **b.** 56,8% (circa); **c.** 54,5% (circa)]

222 Una ditta ha due fornitori di componenti per personal computer. Il 30% dei componenti viene acquistato dal fornitore A e il restante 70% dal fornitore B. In base alle passate esperienze, si stima che il 4% dei componenti acquistati dal fornitore A e il 5% dei componenti acquistati dal fornitore B sono difettosi.

a. Scelto a caso un componente, qual è la probabilità che sia difettoso?

b. Avendo constatato che il componente scelto è difettoso, qual è la probabilità che provenga dal fornitore A?

$\left[\text{a. } \frac{47}{1000}; \text{ b. } \frac{12}{47}\right]$

223 Due scatole A e B, all'apparenza indistinguibili, contengono diversi tipi di cioccolatini. In particolare, la A contiene 10 gianduiotti, 8 boeri e 6 praline alla nocciola; nella B si trovano 5 gianduiotti, 6 boeri e 12 praline. Si sceglie a caso una delle due scatole e si pesca da essa un cioccolatino.

a. Qual è la probabilità che il cioccolatino pescato sia un boero?
b. Sapendo che il cioccolatino pescato è un boero, qual è la probabilità di avere scelto la scatola B?

$\left[\text{a. } \frac{41}{138}; \text{ b. } \frac{18}{41}\right]$

224 In un'università il 30% degli studenti ha frequentato il liceo classico. Il 70% degli studenti che hanno frequentato il liceo classico è di sesso femminile, mentre solo il 40% degli studenti che non hanno frequentato il liceo classico è di sesso femminile. Scelto a caso uno studente di quella università, qual è la probabilità:

a. che sia una ragazza che ha frequentato il liceo classico?
b. che sia di sesso femminile?
c. scelta a caso una ragazza, qual è la probabilità che provenga dal liceo classico?

$\left[\text{a. } \frac{21}{100}; \text{ b. } \frac{49}{100}; \text{ c. } \frac{21}{49}\right]$

225 Abbiamo tre urne A, B, C. L'urna A contiene 1 pallina nera e 2 bianche; l'urna B contiene 2 palline nere e 1 bianca; l'urna C contiene 3 palline nere e nessuna bianca. Scegliamo a caso un'urna ed estraiamo una pallina.

a. Qual è la probabilità che la pallina estratta sia nera?
b. Sapendo che abbiamo estratto una pallina nera, qual è la probabilità che provenga dall'urna A?

$\left[\text{a. } \frac{2}{3}; \text{ b. } \frac{1}{6}\right]$

226 L'insegnante sottopone agli studenti di una classe un quesito che ha cinque risposte, una sola delle quali è esatta. Se uno studente è preparato è certo che risponderà al quesito in modo corretto, mentre se è impreparato sceglierà la risposta in modo del tutto casuale. Nella classe solo il 40% degli studenti sono preparati. Sapendo che un dato studente ha risposto correttamente al quesito, qual è la probabilità che sia preparato?

$\left[\frac{10}{13}\right]$

Matematica e sport

227 Nella finale degli Internazionali di tennis di Roma, giocata al meglio dei tre set (vince chi conquista due set), il tennista A ha prevalso sull'avversario B. Calcola la probabilità che il giocatore A abbia vinto il primo set dell'incontro, assumendo per entrambi i giocatori una probabilità del 50% di vincere un singolo set.
(*Suggerimento*: indica con A l'evento «la partita è stata vinta da A» e con A_1 l'evento «il primo set è stato vinto da A») [75%]

228 La finale del torneo di Wimbledon viene giocata al meglio dei cinque set (vince chi conquista tre set). Supponi che i due finalisti abbiano entrambi il 50% di probabilità di aggiudicarsi un set. Calcola la probabilità che il tennista che vince il primo set vinca la finale. Inversamente, calcola poi la probabilità che il vincitore della finale si sia aggiudicato il primo set.

$\left[\frac{11}{16}; \frac{11}{16}\right]$

229 La probabilità che stasera piova è uguale a $\frac{1}{4}$. Se non piove, la probabilità che stasera vada al cinema è uguale a $\frac{4}{5}$; se piove, la probabilità che vada al cinema è $\frac{1}{10}$. Sapendo che stasera vado al cinema, qual è, in percentuale, la probabilità che non piova? [96%]

230 Uno studente svolge un test in cui ciascuna domanda prevede quattro risposte, di cui una sola è quella esatta. Per ogni domanda, se lo studente ha studiato l'argomento cui si riferisce la domanda (il che accade con probabilità p) allora risponde certamente in modo corretto, altrimenti risponde scegliendo a caso fra le quattro risposte proposte.

a. Sapendo che lo studente ha risposto correttamente a una data domanda, qual è la probabilità che non abbia risposto a caso?
b. Per quali valori di p la probabilità di cui al punto precedente è superiore al 50%?

$\left[\text{a. } \frac{4p}{3p+1}; \text{ b. } p > \frac{1}{5}\right]$

231 Due eventi A e B sono tali che:

$p(A) = 0{,}4 \qquad p(B|A) = 0{,}6 \qquad p(\overline{B}|\overline{A}) = 0{,}2$

Sapendo che si è verificato l'evento B, qual è la probabilità che si è verificato l'evento A? $\left[\frac{1}{3}\right]$

232 Un'urna A contiene 4 palline rosse e 6 palline nere. Un'urna B contiene 2 palline rosse e 8 palline nere. Un giocatore lancia un dado cubico regolare, le cui facce sono numerate da 1 a 6: se ottiene il numero 6, estrae una pallina a caso dall'urna A, altrimenti estrae una pallina a caso dall'urna B.

a. Quale è la probabilità che il giocatore estragga una pallina rossa?
b. Sapendo che il giocatore ha estratto una pallina rossa, qual è la probabilità che essa provenga dall'urna A?
c. Sapendo che il giocatore ha estratto una pallina rossa, è più probabile che tale pallina provenga dall'urna A o dall'urna B?

$\left[\textbf{a. } \frac{7}{30}; \textbf{ b. } \frac{2}{7}; \textbf{ c. } \text{è più probabile che provenga dall'urna } B\right]$

233 Il 20% degli abitanti di un paese soffre di ipertensione. Tra le persone ipertese il 60% sono fumatori. Tra le persone che non sono ipertese, il 50% non sono fumatori. Scelto a caso un fumatore del paese, qual è la probabilità che sia iperteso? $\left[\frac{3}{13}\right]$

234 Il 10% di un gruppo di persone ha contratto una data malattia. Ciascun individuo del gruppo viene sottoposto a un test diagnostico per rilevare la malattia. Se un individuo è malato, la probabilità che il test risulti positivo è uguale a p; se un individuo non è malato, la probabilità che il test risulti negativo è ancora uguale a p.

a. Il test relativo a una persona del gruppo risulta positivo. Qual è la probabilità che abbia davvero contratto la malattia?
b. Qual è il valore della probabilità calcolata al punto precedente, se $p = 95\%$?
c. Affinché la probabilità di cui al punto **a.** sia superiore al 90%, quali valori deve assumere p?

$\left[\textbf{a. } \frac{p}{9-8p}; \textbf{ b. } \frac{19}{28} \simeq 67{,}86\%; \textbf{ c. } p > \frac{81}{82}, \text{ ossia circa } p > 98{,}78\%\right]$

6. Le varie definizioni di probabilità e l'approccio assiomatico

Teoria p. 160

Negli Esercizi 235-244, indica quale definizione di probabilità ti sembra più adatta per assegnare la probabilità agli eventi relativi agli esperimenti aleatori descritti.

235 La probabilità che il primo estratto sulla ruota di Napoli sia il numero 9.

236 La probabilità che, scelto un cittadino italiano, esso abbia più di 40 anni.

237 La probabilità che, lanciando un dado, esca un numero divisibile per 3.

238 La probabilità che il Milan vinca la prossima partita.

239 La probabilità che domani piova.

240 La probabilità che un fumatore fumi più di 20 sigarette al giorno.

241 La probabilità che Barbara passi l'esame.

242 La probabilità che una famiglia italiana scelta a caso abbia una connessione a Internet.

243 La probabilità che un farmaco sia efficace.

244 La probabilità che lanciando due dadi esca un doppio 6.

---

245 Si estrae una carta da un mazzo di 52 carte. Quale probabilità si può attribuire all'evento «è uscito un Asso»? Quale definizione di probabilità si è utilizzata per assegnare la probabilità all'evento?

246 Un individuo A è disposto a scommettere 50 euro su un evento aleatorio, al verificarsi del quale si vincono 1000 euro; un altro individuo B è disposto a puntare 60 euro sullo stesso evento. Quale probabilità attribuisce ciascuno dei due individui a quell'evento aleatorio? Secondo quale definizione di probabilità?

247 Da una scatola che contiene 30 palline colorate si effettuano 250 estrazioni casuali, sempre con reinserimento. Nelle 250 estrazioni si osservano 23 volte palline verdi. Quale stima si può dare della probabilità di estrarre una pallina verde dall'urna? Secondo quale definizione di probabilità?

248 **Realtà e modelli** Nella seguente tabella sono riportati i dati d'ascolto relativi ai telegiornali trasmessi dalle principali emittenti televisive italiane il giorno 7/09/2016.

Telegiornale	Orario	Ascoltatori
TG1	13:30 20:00	3 344 000 5 029 000
TG2	13:00 20:30	2 259 000 2 193 000
TG3	14:25 19:00	1 573 000 1 332 000
TG5	13:00 20:00	2 810 000 3 569 000
TG LA7	13:30 20:00	548 000 1 284 000

Rispondi alle seguenti domande (supponi che il telespettatore scelto a caso sia sicuramente sintonizzato su uno dei cinque canali TV in cui vengono trasmessi i telegiornali indicati in tabella).

a. Qual è la probabilità che, scelto a caso un telespettatore che sta guardando un telegiornale delle ore 13:00, egli stia seguendo il TG1?
b. Qual è la probabilità che, scelto a caso un telespettatore che sta guardando un telegiornale delle ore 20:00, egli stia seguendo il TG5?
c. Qual è la probabilità che, scelto a caso un telespettatore che sta guardando un telegiornale delle ore 19:00, egli stia seguendo il TG3?
d. Qual è la probabilità che, scelto a caso un telespettatore che sta guardando un telegiornale delle ore 20:30, egli non stia seguendo il TG2?

[**a.** 0; **b.** circa 36,12%; **c.** 1; **d.** 0]

---

249 Un consiglio di amministrazione, composto da dieci persone, deve eleggere al suo interno un presidente. I membri con possibilità di essere eletti sono tre: si ritiene che il primo abbia la stessa probabilità di essere eletto del secondo, mentre il terzo è accreditato della metà delle chance di ciascuno degli altri due avversari. Per ciascuno dei tre membri candidati a diventare presidente, determina la probabilità che sia eletto. [40%, 40%, 20%]

250 Un dado è truccato in modo tale che la probabilità che esca 2 è doppia di quella che esca 1, la probabilità che esca 3 è doppia di quella che esca 2, la probabilità che esca 4 è doppia di quella che esca 3, la probabilità che esca 5 è doppia di quella che esca 4 e la probabilità che esca 6 è doppia di quella che esca 5. Si lancia il dado una volta; qual è la probabilità che esca un numero divisibile per 3? $\left[\frac{12}{21}\right]$

251 Un dado, con le facce numerate da 1 a 6, è truccato in modo tale che:

- tutti i numeri pari hanno la stessa probabilità di uscita;
- tutti i numeri dispari hanno la stessa probabilità di uscita;
- la probabilità che esca un numero pari è il triplo di quella che esca un numero dispari.

Si lancia un dado; qual è la probabilità di ottenere un numero minore o uguale a 3? $\left[\frac{5}{12}\right]$

252 Una roulette francese (in cui i numeri vanno da 0 a 36) è truccata in modo tale che:

- tutti i numeri pari hanno la stessa probabilità di uscita;
- tutti i numeri dispari hanno la stessa probabilità di uscita;
- la probabilità che esca un numero pari è doppia di quella che esca un numero dispari (considerando lo 0 un numero pari).

Qual è la probabilità che esca un 4 o un 5? $\left[\frac{3}{56}\right]$

ESERCIZI

Esercizi di riepilogo

Esercizi interattivi

253 **Vero o falso?**

Un'urna contiene dieci palline, numerate da 1 a 10. Si estrae a caso una pallina dall'urna e si indica con A l'evento «È stata estratta una pallina corrispondente a un numero pari» e con B l'evento «È stata estratta una pallina corrispondente a un multiplo di 5».

a. I due eventi A e B sono indipendenti V F

b. i due eventi A e B sono incompatibili V F

Nelle stesse ipotesi precedenti, supponi questa volta che l'urna contenga unidici palline, numerate da 1 a 11. In queste condizioni:

c. I due eventi A e B sono indipendenti V F

d. I due eventi A e B sono incompatibili V F

254 **Vero o falso?**

Due eventi A e B sono tali che $p(A) = \frac{2}{5}$, $p(B) = \frac{1}{2}$, $p(A \cap B) = \frac{3}{10}$, allora

a. $p(A \cup B) = 1$ V F

b. $p(\overline{A}) = p(A \cup B)$ V F

c. $p(B) = p(\overline{B})$ V F

d. $p(A \cap \overline{B}) = p(A) - p(A \cap B)$ V F

e. $p(A \cap \overline{B}) = \frac{1}{10}$ V F

f. $p(A \cup \overline{B}) = \frac{3}{5}$ V F

Test

255 Di due eventi indipendenti si sa che la probabilità dell'unione è uguale a $\frac{1}{2}$ e la probabilità dell'intersezione è uguale a $\frac{1}{12}$. Che cosa possiamo dire in proposito?

A La situazione descritta non si può realizzare
B Uno dei due eventi si verifica con certezza
C Uno dei due eventi ha probabilità $\frac{1}{4}$ e l'altro $\frac{1}{3}$
D I due eventi sono equiprobabili

256 Una scatola contiene 8 palline numerate da 1 a 8. Se ne estraggono 3 e si considera la probabilità che tutte rechino un numero pari a seconda che: l'estrazione sia simultanea (probabilità p_1), l'estrazione avvenga in sequenza con reimmissione (probabilità p_2), l'estrazione avvenga in sequenza senza reimmissione (probabilità p_3). Che cosa possiamo dedurre?

A $p_1 < p_2 < p_3$ B $p_2 < p_1 = p_3$ C $p_2 > p_1 = p_3$ D $p_3 < p_2 < p_1$

257 Il mese di maggio 2014 è cominciato di giovedì e terminato di sabato. Considerando quattro bambini nati in tale mese, Paolo afferma che la probabilità che nessuno di essi sia nato di domenica è circa 200 volte maggiore della probabilità che tutti siano nati di domenica, Barbara afferma che Paolo si sbaglia di un ordine di grandezza, mentre Luigi afferma che Paolo si sbaglia di due ordini di grandezza. Chi ha ragione?

A Paolo B Barbara C Luigi D Sbagliano tutti e tre

258 Paolo, in un tiro al bersaglio, ha probabilità pari a 0,4 di fare centro in ciascun lancio e 12 lanci a disposizione. Qual è la probabilità che faccia centro esattamente 7 volte, supponendo i lanci indipendenti tra loro?

A Circa l'1% B Circa lo 0,1% C Circa il 10% D Circa 1/100

259 Si estrae una pallina da un'urna che ne contiene 20 bianche, 15 verdi, 10 rosse e 30 di altri colori. Calcola la probabilità che essa:

a. sia bianca o rossa;
b. non sia né bianca né rossa;
c. non sia verde.

$\left[\mathbf{a.}\ \frac{2}{5};\ \mathbf{b.}\ \frac{3}{5};\ \mathbf{c.}\ \frac{4}{5}\right]$

260 Da un mazzo di 52 carte se ne estrae una a caso; calcola la probabilità che:

a. sia una carta di fiori;
b. sia un fante;
c. sia una carta di fiori o un fante;
d. non sia una figura;
e. sia una carta di cuori o di picche.

$\left[\mathbf{a.}\ \frac{1}{4};\ \mathbf{b.}\ \frac{1}{13};\ \mathbf{c.}\ \frac{4}{13};\ \mathbf{d.}\ \frac{10}{13};\ \mathbf{e.}\ \frac{1}{2}\right]$

261 In un'urna ci sono 7 biglie rosse, 4 biglie nere e n biglie verdi. Estraendo a caso una biglia dall'urna, la probabilità di estrarre una biglia che non sia nera è 0,75. Quante biglie verdi ci sono nell'urna? [5]

262 Si lancia una moneta equilibrata tre volte, consecutivamente. Determina la probabilità:

a. che si ottenga «croce» esattamente una volta;
b. che si ottenga «croce» esattamente due volte;
c. che si ottenga «croce» tre volte;
d. che si ottenga «croce» almeno una volta.

$\left[\mathbf{a.}\ \frac{3}{8};\ \mathbf{b.}\ \frac{3}{8};\ \mathbf{c.}\ \frac{1}{8};\ \mathbf{d.}\ \frac{7}{8}\right]$

Matematica e informatica

263 Un programma informatico è sottoposto a tre test fra loro indipendenti. Se il programma contiene un errore, i tre test sono in grado di individuare l'errore rispettivamente con una probabilità uguale al 25%, al 40% e al 60%. Supponi che il programma contenga un errore; qual è la probabilità che l'errore venga individuato da almeno uno dei tre test? $\left[\frac{41}{50}\right]$

264 Un nuovo virus informatico può essere trasmesso in due modi: mediante allegati e-mail o mediante file scaricati da un sito web. La probabilità che il virus contagi un dato sistema tramite e-mail è del 20%, mentre la probabilità che il virus contagi lo stesso sistema via sito web è del 50%. Infine, la probabilità che il virus contagi il sistema sia tramite e-mail sia tramite web è del 30%. Qual è la probabilità che il virus **non** contagi il sistema? $\left[\frac{3}{5}\right]$

265 Un'urna contiene 5 biglie rosse e 10 bianche. Si estraggono dall'urna, successivamente, due biglie, *senza rimettere* nell'urna la prima biglia estratta. Determina la probabilità:

a. di estrarre due biglie rosse;
b. di estrarre due biglie dello stesso colore;
c. di estrarre due biglie di colori diversi.

$\left[\mathbf{a.}\ \frac{2}{21};\ \mathbf{b.}\ \frac{11}{21};\ \mathbf{c.}\ \frac{10}{21}\right]$

266 Un'urna contiene 5 biglie rosse e 10 bianche. Si estraggono dall'urna, successivamente, due biglie, *rimettendo* nell'urna la prima biglia estratta. Determina la probabilità:

a. di estrarre due biglie rosse;
b. di estrarre due biglie dello stesso colore;
c. di estrarre due biglie di colori diversi.

$\left[\mathbf{a.}\ \frac{1}{9};\ \mathbf{b.}\ \frac{5}{9};\ \mathbf{c.}\ \frac{4}{9}\right]$

267 Un'urna contiene 25 palline bianche e 15 rosse. Si estraggono, successivamente, tre palline *senza reinserire*, dopo ciascuna estrazione, la pallina estratta nell'urna. Qual è la probabilità che:

a. siano estratte tutte palline dello stesso colore;
b. solo una delle palline estratte sia bianca.

$\left[\mathbf{a.}\ \frac{29}{104} \simeq 27{,}88\%;\ \mathbf{b.}\ \frac{525}{1976} \simeq 26{,}57\%\right]$

268 Si lancia quattro volte una moneta regolare.

a. Qual è la probabilità di ottenere quattro «testa»?
b. Qual è la probabilità di ottene 2 «testa» e 2 «croce»?

$\left[\mathbf{a.}\ \frac{1}{16};\ \mathbf{b.}\ \frac{3}{8}\right]$

269 Una password è costituita da cinque cifre, con la possibilità che le cifre siano ripetute. Scegliendo a caso una password, determina la probabilità:

a. che non contenga alcuno 0; **b.** che contenga almeno uno 0; **c.** che contenga esattamente uno 0.

[**a.** 0,59049; **b.** 0,40951; **c.** 0,32805]

270 In un compito in classe Alessandro deve rispondere a tre quesiti del tipo «vero o falso». Supponi che Alessandro risponda a caso a tutti e tre i quesiti. Calcola la probabilità che Alessandro:

a. abbia riposto correttamente a tutti e tre i quesiti;
b. abbia risposto correttamente a solo due quesiti;
c. abbia dato almeno una risposta corretta;
d. abbia dato almeno una risposta sbagliata.

$\left[\textbf{a. } \frac{1}{8}; \textbf{ b. } \frac{3}{8}; \textbf{ c. } \frac{7}{8}; \textbf{ d. } \frac{7}{8}\right]$

Matematica in azienda

271 Viene indetta una gara di appalto per la costruzione di un ponte. Il management di un'impresa specializzata nella costruzione di ponti pensa di potersi aggiudicare l'appalto con una probabilità del 50%. Dopo avere presentato il proprio progetto, vengono chieste all'impresa informazioni aggiuntive. In base a statistiche effettuate sulla base di precedenti gare di appalto, è noto che sono state chieste informazioni aggiuntive nell'80% dei casi in cui un progetto presentato ha vinto l'appalto e nel 35% dei casi in cui non l'ha vinto. Dopo avere saputo che sono state richieste informazioni aggiuntive, con quale probabilità l'impresa può ritenere che il progetto vinca l'appalto?

$\left[\frac{16}{23} \simeq 69{,}6\%\right]$

272 Una grande società fa mandare in onda in televisione uno spot pubblicitario per reclamizzare uno dei suoi prodotti. In base a un'indagine statistica effettuata sui potenziali clienti si è stabilito che:

- la probabilità che un potenziale cliente acquisti il prodotto pubblicizzato è uguale a 0,25;
- la probabilità che un potenziale cliente abbia visto lo spot è uguale a 0,5;
- la probabilità che un potenziale cliente acquisiti il prodotto pubblicizzato e abbia visto lo spot è uguale a 0,15.

a. Qual è la probabilità che un potenziale cliente acquisti il prodotto pubblicizzato, se l'individuo ha visto lo spot?

b. Qual è la probabilità che un potenziale cliente acquisti il prodotto pubblicizzato, se l'individuo non ha visto lo spot?

c. Vedere lo spot accresce la probabilità che il potenziale cliente acquisti il prodotto? Supponendo che i potenziali clienti siano 1 000 000, che per ogni unità di prodotto venduto l'azienda abbia un guadagno di 3 euro e che per la trasmissione dello spot l'azienda abbia un costo di 200 000 euro, ritieni che per l'azienda sia utile continuare a fare trasmettere lo spot?

[**a.** 0,3; **b.** 0,2]

---

273 Un'urna contiene n palline. Se si estrae una pallina, la probabilità che essa sia gialla è $\frac{4}{7}$. Se invece si estraggono successivamente due palline, *senza* rimettere la prima pallina estratta nell'urna prima di estrarre la seconda, la probabilità che esse siano entrambe gialle è $\frac{38}{119}$. Quanto vale n? Quante sono le palline gialle contenute nell'urna?

[$n = 35$, 20 palline gialle]

274 Un'urna A contiene 100 biglie, di cui 20 rosse e 80 bianche. Un'urna B contiene n biglie rosse e 85 bianche. Determina il minimo valore di n per cui la probabilità di estrarre una biglia rossa dall'urna B sia maggiore della probabilità di estrarre una biglia rossa dall'urna A.

[$n = 22$]

275 Un'urna contiene tre biglie rosse e quattro biglie nere. Si estraggono a caso dall'urna due biglie, senza rimettere la prima biglia estratta nell'urna. Determina la probabilità che:

a. la prima biglia estratta sia rossa e la seconda nera;
b. le due biglie estratte siano entrambe nere;
c. almeno una delle due biglie estratte sia rossa;
d. le due biglie estratte abbiano lo stesso colore;
e. le due biglie estratte abbiano colori differenti.

$\left[\textbf{a. } \frac{2}{7}; \textbf{ b. } \frac{2}{7}; \textbf{ c. } \frac{5}{7}; \textbf{ d. } \frac{3}{7}; \textbf{ e. } \frac{4}{7}\right]$

276 Un giocatore lancia un dado regolare, successivamente, per cinque volte. In ciascuno dei cinque lanci, il giocatore perde se esce «1» e vince altrimenti.

a. Qual è la probabilità che perda esattamente tre volte?

b. Qual è la probabilità che vinca esattamente quattro volte?

$\left[\textbf{a. } \frac{125}{3888}; \textbf{ b. } \frac{3125}{7776}\right]$

277 Il 10% di una certa popolazione è contagiato da un virus. Si scelgono a caso 4 persone dalla popolazione. Supponendo le scelte tra loro indipendenti, qual è la probabilità che almeno due delle quattro persone siano contagiate?

$\left[\frac{523}{10\,000}\right]$

Matematica e controllo qualità

278 Una azienda produce bulloni che vengono venduti in stock. La probabilità che un bullone prodotto non sia conforme alle caratteristiche dichiarate è uguale a 0,02.

a. Ai fini di un controllo della qualità, viene estratto un campione di 6 bulloni da un lotto sufficientemente numeroso perché il campionamento si possa assimilare a una estrazione con reimmissione. Qual è la probabilità che il campione estratto non contenga alcun pezzo difettoso? E quale la probabilità che contenga esattamente due pezzi difettosi?

b. Da quanti pezzi al massimo può essere costituito uno stock di vendita, affinché la probabilità che esso non contenga alcun pezzo difettoso sia superiore al 90%? Supponi che l'eventuale difettosità di un bullone sia indipendente dagli altri. [**a.** circa 88,6%, circa 0,55%; **b.** 5]

279 Una azienda produce delle protesi, che vengono sottoposte a un test di controllo della qualità. In base a statistiche precedenti è noto che la percentuale di protesi prodotte difettose è del 5%.

a. Viene estratto un campione di 10 protesi da un lotto sufficientemente numeroso perché il campionamento si possa assimilare a una estrazione con reimmissione. Qual è la probabilità che al massimo una delle protesi del campione sia difettosa?

b. Viene estratto un campione di 20 protesi da un lotto sufficientemente numeroso perché il campionamento si possa assimilare a una estrazione con reimmissione. Qual è la probabilità che al massimo due delle protesi del campione siano difettose? [**a.** circa 91,4%; **b.** circa 92,5%]

280 Un'azienda si prepara a lanciare sul mercato un nuovo elettrodomestico. L'azienda ha identificato un evento che può potenzialmente causarne un guasto. La probabilità che tale evento causi effettivamente il guasto è stimata uguale a 0,1. In base ai parametri di qualità fissati, si vuole che la probabilità di guasto effettivo, nell'ipotesi che l'evento potenzialmente pericoloso si realizzi per cinque volte indipendenti l'una dall'altra, sia inferiore al 10%. L'azienda può immettere il prodotto sul mercato, oppure deve rivederne la progettazione? [Deve rivederne la progettazione]

Matematica e telecomunicazioni

281 In un canale di trasmissione binario la probabilità che un singolo bit sia trasmesso correttamente è uguale a p. Si vuole trasmettere una sequenza costituita da 4 bit. Per aumentare la probabilità che la sequenza sia trasmessa correttamente, si utilizza un codice che prevede la trasmissione di 3 bit extra: in tal caso, per trasmettere una sequenza di 4 bit, si invia una sequenza di 7 bit (i 4 bit originari + i 3 bit extra). Tale codice è in grado di correggere 1 errore (al massimo) tra quelli che potrebbero verificarsi durante la trasmissione. Supponi che eventuali errori di trasmissione siano indipendenti e determina:

a. la probabilità che la sequenza originaria di 4 bit venga trasmessa correttamente, se **non** si utilizza il codice;

b. la probabilità che la sequenza originaria di 4 bit venga trasmessa correttamente, se si utilizza il codice correttore che prevede la trasmissione dei 3 bit extra;

c. per quali valori di p l'utilizzo del codice correttore aumenta la probabilità che la sequenza originaria sia trasmessa correttamente.

$\left[\textbf{a. } p^4;\ \textbf{b. } 7p^6 - 6p^7;\ \textbf{c. } p > \frac{1}{2}\right]$

282 Un segnale binario può assumere solo due stati (0 o 1). Un segnale emesso come 0 attraversa due successivi canali di trasmissione, prima di essere ricevuto. In ciascun canale, il segnale viene trasmesso correttamente con una probabilità del 95%, altrimenti si verifica un errore che comporta un'inversione del segnale (se era 0 diventa 1 e viceversa). Supponi che gli eventuali errori di trasmissione che si verificano nei due canali siano indipendenti.

a. Determina la probabilità che il segnale emesso inizialmente come 0 venga ricevuto correttamente (cioè ancora come 0).

b. Se il segnale emesso inizialmente come 0 viene ricevuto correttamente, qual è la probabilità che il primo canale lo abbia trasmesso correttamente?

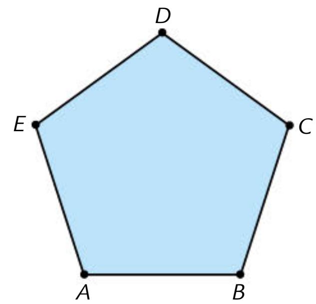
c. Vengono inviati 5 segnali emessi come 0, indipendentemente uno dall'altro. Qual è la probabilità che esattamente due di essi **non** vengano ricevuti correttamente? [**a.** 90,5%; **b.** circa 99,7%; **c.** circa 6,7%]

Esercizi *più*

283 Considera i cinque vertici di un pentagono *regolare*. Qual è la probabilità che tre di essi, scelti a caso, siano i vertici di un triangolo:

a. acutangolo;
b. ottusangolo;
c. isoscele.

[**a.** 50%; **b.** 50%; **c.** 100%]



Dalle gare

284 Tra i partecipanti a un campo estivo il numero delle ragazze supera del 40% quello dei ragazzi. La probabilità che una delegazione di due partecipanti scelta casualmente sia formata da una ragazza e da un ragazzo è esattamente $\frac{1}{2}$. Quanti sono i partecipanti al campo?

A 20 B 24 C 36 D 38
E La situazione descritta è impossibile

(Kangourou 2018)

285 Un'urna contiene più di 10 palline, ma meno di 50, tutte uguali eccetto per il colore. La probabilità di pescarne una rossa è del 20%, ma se aggiungo all'urna alcune palline rosse e una nera tale probabilità aumenta a $\frac{4}{9}$. Quante palline rosse ho aggiunto?

A 14 B 17 C 20 D 22 E 25

(Gran Premio di Matematica Applicata 2017)

286 Tira un dado tradizionale: qual è la probabilità che il prodotto dei numeri che compaiono sulle cinque facce che rimangono visibili sia divisibile per 6?

A $\frac{1}{3}$ B $\frac{1}{2}$ C $\frac{2}{3}$ D $\frac{5}{6}$ E 1

(Kangourou 2007)

287 Tiriamo due dadi regolari, con le facce numerate da 1 a 6, osserviamo i punteggi sulle facce superiori e calcoliamone la differenza. Qual è il valore più probabile per il valore assoluto di tale differenza?

A Tutti i numeri fra 0 e 5 sono equiprobabili
B 0 C 1 D 2 E 3

(Kangourou 2003)

288 Due gabbiani bianchi e otto gabbiani grigi volano su un fiume. All'improvviso atterrano su una delle sponde del fiume, disponendosi in linea retta in ordine casuale. Qual è la probabilità che i due gabbiani bianchi si trovino uno accanto all'altro?

A $\frac{1}{5}$ B $\frac{1}{6}$ C $\frac{1}{7}$ D $\frac{1}{8}$ E $\frac{1}{9}$

(Kangourou 2003)

289 Carla si è dimenticata la password di accesso al suo nuovissimo computer! Si ricorda però che è una sequenza di 4 vocali, non necessariamente distinte, di cui due sono maiuscole e due sono minuscole. Quante password diverse deve provare Carla, al massimo, per accedere al suo computer?

A $3 \cdot 5^4$ B 5^5 C $6 \cdot 5^4$ D 5^6 E $3 \cdot 5^6$

(Giochi di Archimede 2009)

290 Quattro ragazzi vogliono telefonare tutti contemporaneamente alle rispettive ragazze. Ogni loro cellulare può funzionare su quattro frequenze distinte. Se due cellulari si attivano sulla stessa frequenza, la comunicazione cade. Se ogni ragazzo non sa che frequenza scelgono gli atri tre, qual è la probabilità che tutti e quattro riescano a parlare con le loro ragazze?

A $\frac{3}{32}$ B $\frac{3}{64}$ C $\frac{1}{256}$ D $\frac{1}{16}$ E $\frac{9}{129}$

(Giochi di Archimede 2003)

291 Lanciando due dadi regolari con 12 facce, numerate da 1 a 12, la probabilità che la somma dei valori delle facce sia 13 è uguale a:

A $\frac{1}{24}$ B $\frac{1}{12}$ C $\frac{13}{144}$ D $\frac{1}{6}$ E $\frac{13}{72}$

(Giochi di Archimede 2001)

292 Un comune dado con facce numerate da 1 a 6 viene lanciato tre volte e ogni volta si prende un bastoncino di lunghezza pari al risultato del lancio. Qual è la probabilità che i tre bastoncini costituiscano i lati di un triangolo rettangolo?

A $\frac{1}{6}$ B $\frac{1}{36}$ C $\frac{1}{216}$ D $\frac{5}{18}$ E $\frac{1}{72}$

(Giochi di Archimede 2000)

293 La percentuale di femmine che nascono nei parti gemellari è del 48,5%. Supponendo che nei parti gemellari la probabilità che i due nati siano di sesso differente sia del 33%, qual è la probabilità che in un parto gemellare nascano due femmine?

A 32% B 33% C 33,33% D 35% E 50%

(Giochi di Archimede 2002)

Prova di autoverifica

ESERCIZI

Calcolo delle probabilità

1 **Vero o falso?**

a. due eventi incompatibili e aventi probabilità non nulla sono indipendenti V F

b. la probabilità di pescare, al primo tentativo, uno dei quattro assi da un mazzo di quaranta carte è il 10% V F

c. la probabilità che in 10 lanci di una moneta esca «testa» esattamente 8 volte è data dall'espressione
$\binom{10}{8}\left(\frac{1}{2}\right)^{10}$ V F

d. se A e B sono due eventi indipendenti di uno stesso spazio campionario Ω, allora $p(A \cup B) = p(A) + p(B)$ V F

2 Supponiamo che l'ordine con cui 5 congressisti devono prendere la parola sia stabilito per sorteggio e che ciascuno prenda la parola una e una sola volta.

a. Qual è la probabilità, espressa in percentuale, che l'intervento del Prof. Verdi **non** sia l'ultimo?

b. Qual è la probabilità che l'ordine degli interventi coincida con quello alfabetico crescente (dalla A alla Z) o decrescente (dalla Z alla A) dei cinque congressisti?

3 Due eventi A e B di uno stesso spazio campionario Ω sono indipendenti e hanno uguale probabilità k. Determina il valore di k, sapendo che $p(A \cup B) = \frac{3}{4}$.

4 Da un sacchetto contenente 2 palline blu, 4 rosa e 3 nere sono estratte due palline. Stabilisci qual è la probabilità che siano entrambe nere, nel caso in cui:

a. la prima pallina estratta sia reimmessa nel sacchetto;

b. la prima pallina estratta non sia reimmessa nel sacchetto.

5 Una ditta ha due fornitori di componenti per personal computer. Il 40% dei componenti viene acquistato dal fornitore A e il restante 60% dal fornitore B. In base alle passate esperienze, si stima che l'8% dei componenti acquistati dal fornitore A e il 6% dei componenti acquistati dal fornitore B sono difettosi.

a. Scelto a caso un componente, qual è la probabilità che sia difettoso?

b. Avendo constatato che il componente scelto è difettoso, qual è la probabilità che provenga dal fornitore A?

6 Alessandro stima che ciascuno dei cinque rigoristi della squadra per cui tifa abbia una probabilità dell'80% di segnare il proprio calcio di rigore. Qual è la probabilità che almeno quattro dei cinque rigoristi realizzino il proprio calcio di rigore?

7 Calcola la probabilità di realizzare tre volte «testa» e tre volte «croce» (in qualunque ordine) lanciando per sei volte una moneta regolare.

8 Alex e Bob si sfidano ai tiri liberi: vince chi, al termine del secondo turno di tiri, ne ha realizzati di più. Calcola la probabilità che la sfida finisca in pareggio, sapendo che entrambi hanno una percentuale di realizzazione sul singolo lancio del 50%. Assumi che i tentativi a canestro siano indipendenti.

Valutazione

Esercizio	1	2	3	4	5	6	7	8	**Totale**
Punteggio	$0{,}25 \cdot 4 = 1$	$0{,}5 \cdot 2 = 1$	1	$0{,}75 \cdot 2 = 1{,}5$	$0{,}75 \cdot 2 = 1{,}5$	1,5	1	1,5	10
Punteggio ottenuto									

Tempo indicativo: 1 h

Risposte p. 219

Risolvere problemi e costruire modelli

1 Per recarsi in ufficio il signor Francesconi incontra tre semafori agli incroci. Calcola la probabilità che al suo transito i tre semafori siano tutti verdi, sapendo che la probabilità che un singolo semaforo sia verde è del 40%. [6,4%]

2 Le proteine, come è noto, sono costituite da sequenze di aminoacidi. Classicamente esistono in natura 20 differenti tipi di aminoacidi (anche se studi più recenti suggeriscono un numero maggiore). Fra tutte le possibili proteine costituite da una sequenza di 10 aminoacidi, qual è la percentuale di quelle per cui gli aminoacidi che compaiono nella sequenza sono tutti differenti? $\left[\dfrac{20 \cdot 19 \cdot 18 \cdot \ldots \cdot 11}{20^{10}} \simeq 6,5\%\right]$

3 Si vuole che il codice di accesso ai servizi web di una società di assicurazioni sia costituito da n numeri (compresi tra 0 e 9, inclusi 0 e 9) seguiti da due lettere dell'alfabeto italiano. Qual è il minimo valore di n per poter generare almeno 10 milioni di codici diversi? $[n = 5]$

4 Una coppia di coniugi ha tre figli. Supponendo che la probabilità di avere un figlio maschio sia uguale a quella di avere una figlia femmina, calcola la probabilità:

a. che i tre figli siano tutti dello stesso sesso;
b. che i tre figli siano tre femmine;
c. che dei tre figli almeno uno sia maschio;
d. che la primogenita sia femmina. $\left[\text{a. } \frac{1}{8}; \text{ b. } \frac{1}{4}; \text{ c. } \frac{7}{8}; \text{ d. } \frac{1}{2}\right]$

5 In un liceo di 500 allievi si è svolta un'indagine sul numero delle assenze degli studenti in un dato anno scolastico. Si è rilevato che:

- il 45% degli allievi sono stati assenti almeno un giorno;
- il 25% degli allievi sono stati assenti almeno due giorni;
- il 10% degli allievi sono stati assenti almeno tre giorni;
- il 5% degli allievi sono stati assenti almeno quattro giorni.

Scegliendo a caso uno studente di quel liceo, calcola la probabilità che:

a. lo studente sia stato assente almeno un giorno;
b. lo studente non sia mai stato assente;
c. lo studente sia stato assente esattamente un giorno;
d. lo studente sia stato assente esattamente due o tre giorni;
e. lo studente sia stato assente esattamente due giorni.

Esprimi tutti i risultati tramite frazioni ridotte ai minimi termini. $\left[\text{a. } \frac{9}{20}; \text{ b. } \frac{11}{20}; \text{ c. } \frac{1}{5}; \text{ d. } \frac{1}{5} \text{ e. } \frac{3}{20}\right]$

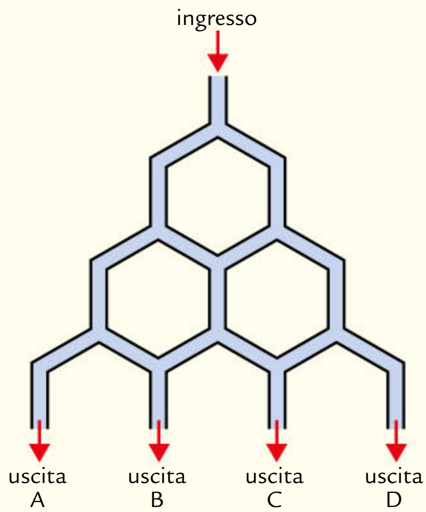
6 La famiglia Peterson ha tre figli. Calcola la probabilità che siano tutti e tre maschi, nei due casi seguenti:

a. sapendo che almeno due di essi sono maschi;
b. sapendo che il primogenito è maschio, così come almeno uno degli altri due.
c. Come spieghi il fatto che le probabilità appena calcolate sono diverse? Perché le due informazioni ai punti **a.** e **b.** non sono equivalenti? $\left[\text{a. } \frac{1}{4}; \text{ b. } \frac{1}{3}\right]$

7 In un dato giorno si è rilevato che $\frac{1}{4}$ delle persone presenti sulle piste da sci di una località di montagna praticano snowboard, mentre gli altri sono sciatori. Inoltre, $\frac{2}{3}$ di coloro che praticano snowboard hanno meno di 25 anni, mentre la metà delle persone presenti sulle piste hanno 25 anni o più. Si sceglie a caso una persona sulla pista; calcola la probabilità che tale persona:

a. pratichi snowboard e abbia 25 anni o più;
b. sia uno sciatore con meno di 25 anni;
c. sia uno sciatore con 25 anni o più. $\left[\text{a. } \frac{1}{12}; \text{ b. } \frac{1}{3}; \text{ c. } \frac{5}{12}\right]$

8 **Dispositivo di Galton.** Una pallina scende lungo la guida rappresentata in figura. A ogni diramazione, la probabilità di imboccare la via di sinistra è uguale a quella di imboccare la via di destra.



a. Rappresenta con un diagramma ad albero tutti i possibili percorsi attraverso i quali la pallina può uscire dalla guida.

b. Da quali delle uscite A, B, C, D è più probabile che la pallina esca dopo aver completato il suo percorso?

c. Da quali delle uscite A, B, C, D è meno probabile che la pallina esca dopo aver completato il suo percorso?

[**b.** Dalle uscite B e C, da ciascuna delle quali la pallina esce con probabilità $\frac{3}{8}$; **c.** dalle uscite A e D, da ciascuna delle quali la pallina esce con probabilità $\frac{1}{8}$]

Realtà e modelli

9 **Win for life.** Il gioco del Win for life consiste nello scegliere 10 numeri tra 20. Ogni numero dei 10 scelti che viene estratto (in una estrazione di 10 dei 20 numeri senza reimmissione) fornisce 1 punto.

a. Qual è la probabilità di fare 7 punti?

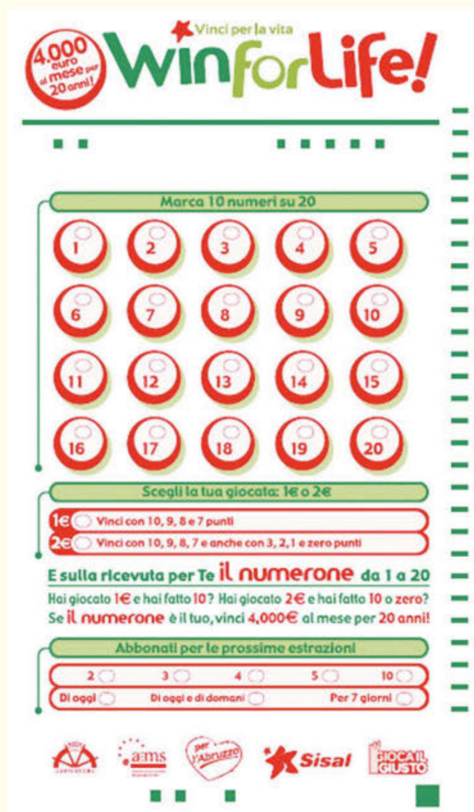
b. Qual è la probabilità di fare 8 punti?

c. Qual è la probabilità di fare 9 punti?

d. Qual è la probabilità di fare 10 punti?

e. Se hai giocato 1 euro e hai fatto 10 punti, devi controllare anche il «numerone» (che è un numero, sempre tra 1 e 20, assegnato casualmente dal sistema): se è uguale a quello estratto (in una seconda estrazione indipendente dalla prima, per cui il «numerone» può eventualmente coincidere con uno dei 10 numeri estratti nella prima estrazione), vinci la rendita di 6000 euro al mese per 20 anni (la rendita è aumentata da 4000 a 6000 euro al mese dall'8 giugno 2010). Qual è la probabilità di vincere tale rendita?

f. Se hai giocato 2 euro, vinci la rendita, oltre che nel caso descritto al punto precedente, anche se hai fatto 0 punti ed esce il «numerone» In tal caso qual è la probabilità di vincere la rendita?



[**a.** $\frac{3600}{46\,189} \simeq 7{,}80\%$; **b.** $\frac{2025}{184\,756} \simeq 1{,}09\%$; **c.** $\frac{25}{46\,189} \simeq 0{,}05\%$; **d.** $\frac{1}{184\,756}$; **e.** $\frac{1}{3\,695\,120}$; **f.** $\frac{1}{1\,847\,560}$]

10 I pantaloni prodotti da un'azienda possono presentare due difetti, diciamo A e B. La presenza del difetto A è indipendente dalla presenza del difetto B; inoltre la probabilità che un pantalone presenti il difetto A è del 5% e la probabilità che presenti il difetto B è del 2%.

Scelto a caso un pantalone tra quelli prodotti dall'azienda, determina qual è la probabilità che:

a. non presenti il difetto A;
b. presenti entrambi i difetti;
c. presenti almeno uno dei due difetti;
d. presenti il difetto A, ma non il difetto B;
e. non presenti nessuno dei due difetti;
f. presenti soltanto uno dei due difetti.

Esprimi tutti i risultati sotto forma di percentuale.

11 Una azienda produce dei componenti elettronici; la probabilità che un componente prodotto risulti difettoso è uguale a 0,002. Ai fini di un controllo della qualità, viene estratto un campione di 20 componenti da un lotto. Nell'ipotesi che il campionamento avvenga tramite 20 estrazioni successive con reimmissione, calcola:

a. la probabilità che il campione estratto non contenga alcun pezzo difettoso;
b. la probabilità che il campione estratto contenga esattamente due pezzi difettosi;
c. la probabilità che il campione estratto contenga esattamente uno o esattamente due pezzi difettosi;
d. la probabilità che il campione estratto contenga al massimo due pezzi difettosi;
e. da quanti componenti al massimo può essere costituito uno stock di vendita, affinché la probabilità che esso non contenga alcun pezzo difettoso sia superiore al 95% (supponendo che ciascun componente possa essere difettoso indipendentemente dagli altri)? [**a.** Circa 96%; **b.** circa 0,073%; **c.** circa 3,92%; **d.** circa 99,9%; **e.** 25]

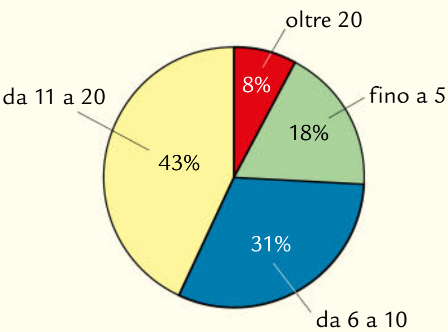
Interpretare grafici e dati

12 Osserva il diagramma. Scelto a caso un fumatore, qual era, nel 2008, la probabilità:

a. che fumasse al massimo 5 sigarette al giorno?
b. che fumasse più di 10 sigarette al giorno?
c. che fumasse meno di 6 sigarette al giorno o più di 20 sigarette al giorno?

Esprimi tutti i risultati tramite frazioni ridotte ai minimi termini.

Fumatori per numero di sigarette fumate al giorno nel 2008



$\left[\textbf{a. } \frac{9}{50}; \textbf{ b. } \frac{51}{100}; \textbf{ c. } \frac{13}{50}\right]$

13 Le lampadine di una nota marca vengono prodotte in tre diversi stabilimenti: uno a Milano, uno a Bologna e uno a Roma. Si effettua un controllo sulla qualità delle lampadine prodotte nel mese di giugno 2018 in ciascuno dei tre stabilimenti. I dati raccolti sono riportati, in parte, nella seguente tabella.

Sede dello stabilimento	Numero di lampadine difettose	Numero di lampadine funzionanti regolarmente	Totale
Milano	140	....	3240
Bologna	....	....	1264
Roma	152	....	....
Totale	360	8200	....

a. Completa la tabella.
b. Scelta a caso una lampadina tra quelle prodotte nel giugno 2018, calcola la probabilità dei seguenti eventi:
- la lampadina è stata prodotta a Milano;
- la lampadina è stata prodotta a Bologna o a Roma;
- la lampadina è difettosa;
- la lampadina è stata prodotta a Milano ed è difettosa;
- la lampadina è stata prodotta a Milano o è difettosa.

c. Quale dei tre stabilimenti sembra garantire una maggiore qualità nella produzione?

$\left[\mathbf{b.}\ \frac{81}{214}, \frac{133}{214}, \frac{9}{214}, \frac{7}{428}, \frac{173}{428}\right]$

14 Un campione rappresentativo di 200 ragazzi dai 18 ai 25 anni viene interrogato sul numero medio di SMS che invia in una giornata. I risultati ottenuti sono riportati nella seguente tabella.

Numero medio di SMS inviati in un giorno	0	1	2	3	4	Più di 4
Frequenza	10	30	64	36	40	20

Scegliendo a caso un ragazzo appartenente al campione considerato, qual è la probabilità che egli:

a. non invii SMS;
b. invii in media 1 SMS al giorno;
c. invii almeno due SMS al giorno;
d. invii al massimo due SMS al giorno. [**a.** 0,05; **b.** 0,15; **c.** 0,8; **d.** 0,52]

15 In un villaggio turistico gli animatori hanno richiesto alla reception alcuni dati riguardanti gli ospiti maggiorenni per organizzare una serata di intrattenimento: infatti si vuole organizzare una gara in cui i partecipanti saranno divisi in squadre in base alla provenienza.
I dati sono riportati in tabella.

	Femmine	Maschi
Sud Italia	15	20
Centro Italia	12	17
Nord Italia	8	5
Stranieri	9	13

Determina:

a. la probabilità che scelte a caso due femmine tra i partecipanti, esse siano entrambe della squadra del Sud Italia;
b. la probabilità che scelti a caso due concorrenti della squadra degli stranieri, essi siano un maschio e una femmina;
c. la probabilità che scelti a caso due partecipanti, estraendo a sorte i nominativi sia della squadra sia dei concorrenti, essi siano un maschio e una femmina;
d. la probabilità che scelto a caso un concorrente estraendo il nominativo a sorte e verificato che sia un nome maschile, il concorrente sia della squadra del Nord Italia. [**a.** Circa 11%; **b.** circa 51%; **c.** circa 51%; **d.** circa 9%]

16 Osserva la seguente tabella (*fonte*: Italia in cifre, Istat, 2010).

SCUOLE, CLASSI E ALUNNI PER TIPO DI SCUOLA

Anno scolastico 2008/2009

	Dell'infanzia	Primarie	Secondarie di primo grado	Secondarie di secondo grado
Scuole	24 518	18 009	7 921	6 809
Classi*	72 889	150 345	82 751	130 784
Alunni	1 651 713	2 819 193	1 758 384	2 723 562
% femmine sul totale	48,1	48,3	47,9	49,0
% iscritti a scuole pubbliche	69,6	93,1	95,9	94,5
Stranieri per 1000 iscritti	75,7	83,1	79,6	48,0
Ripetenti per 100 iscritti	–	0,3	3,4	7,7

* *Per le scuole dell'infanzia si fa riferimento alle sezioni.*

Facendo riferimento alle scuole, classi e alunni dell'anno scolastico 2008/2009, rispondi alle seguenti domande. Esprimi i risultati sotto forma di percentuale.

a. Scelta a caso una scuola, qual è la probabilità che essa sia una scuola primaria?
b. Scelta a caso una classe, qual è la probabilità che appartenga alla scuola secondaria di primo o secondo grado?
c. Scelto a caso un alunno della scuola secondaria di primo grado, qual è la probabilità che sia una femmina?
d. Scelto a caso un alunno della scuola secondaria di secondo grado, qual è la probabilità che sia un maschio?
e. Scelto a caso un alunno della scuola primaria, qual è la probabilità che sia iscritto a una scuola privata?
f. Scelto a caso un alunno della scuola primaria, qual è la probabilità che sia straniero?
g. Scelto a caso un alunno della scuola secondaria di primo o secondo grado, qual è la probabilità che sia ripetente?

[**a.** Circa 31,45%; **b.** circa 48,89% ; **c.** 47,9%; **d.** 51%; **e.** 6,9% ; **f.** 8,31%; **g.** circa il 6%]

Esporre, argomentare e dimostrare

17 Siano A e B due eventi. Considera la proposizione «se $A \subset B$, allora $p(A) \leq p(B)$».

a. Stabilisci se è vera.
b. Scrivi la proposizione *inversa* di quella data e stabilisci se è vera.

In ciascuno dei due casi, se la risposta è che la proposizione è vera, dimostrala, altrimenti trova un controesempio.

18 Spiega, tramite facili esempi, la differenza tra eventi incompatibili ed eventi indipendenti. Vi è una qualche relazione tra i due concetti?

19 Dimostra che se A e B sono due eventi indipendenti, allora anche $\overline{A}$ e $\overline{B}$ sono indipendenti. Scrivi l'enunciato inverso di quello dimostrato e stabilisci se è vero: in caso affermativo, dimostralo, altrimenti esibisci un controesempio.

20 Da un'urna contenente n palline, numerate da 1 a n, se ne estraggono k, con $k \leq n$. Stabilisci in quanti modi si possono estrarre le k palline, nell'ipotesi che:

a. le palline siano estratte successivamente, con reimmissione;
b. le palline siano estratte successivamente, senza reimmissione;
c. le palline siano estratte simultaneamente.

21 Dimostra che $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$ nei seguenti due modi:

a. algebricamente, utilizzando la definizione di coefficiente binomiale;
b. dal punto di vista insiemistico, osservando che, fissato un elemento, diciamo a, di un insieme A di n elementi, il numero di sottoinsiemi di A di k elementi è uguale alla somma tra il numero di sottoinsiemi di k elementi di A che contengono l'elemento a e il numero di sottoinsiemi di k elementi di A che non contengono a.

22 Dimostra che, per ogni numero naturale n maggiore di 0, risulta $\sum_{k=0}^{n} \binom{n}{k} = 2^n$.

Verso la prova Invalsi

1 Una fabbrica utilizza due diverse macchine M_1 ed M_2 che lavorano indipendentemente l'una dall'altra. Ciascuna delle due macchine produce chiavette USB da 16 GB e da 32 GB nelle percentuali descritte dalla seguente tabella.

	Chiavette USB da 16 GB	Chiavette USB da 32 GB	Totale
M_1	18%	42%	60%
M_2	22%	18%	40%
Totale	40%	60%	100%

a. Qual è la probabilità di estrarre dalla produzione della fabbrica una chiavetta da 16 GB prodotta da M_1?

Risposta: .... %

b. Qual è la probabilità che una chiavetta USB estratta dalla produzione della fabbrica sia da 16 GB?

Risposta: .... %

(Prova Invalsi 2017)

2 Il PIN delle carte Bancomat è usualmente formato da cinque cifre scelte tra le dieci a disposizione (0, 1, 2, 3, 4, 5, 6, 7, 8, 9). Un codice è accettabile se non è formato dalla stessa cifra ripetuta 5 volte. Quanti sono i codici accettabili? Risposta: ....

(Pretest 2015)

3 Un test è composto da 5 domande del tipo «vero» o «falso».

a. Quanti sono i modi diversi di compilare il test rispondendo a caso a ogni domanda?

A 5 B 10 C 32 D 25

b. Uno studente risponde a caso a tutte le domande. Qual è la probabilità che tutte le risposte siano corrette?

Risposta: ....

(Pretest 2016)

---

4 La probabilità di estrarre una pallina bianca da un'urna è $\frac{3}{5}$. Quale delle seguenti affermazioni è compatibile con la precedente?

A L'urna contiene 6 palline bianche, 2 rosse e 3 nere.
B L'urna contiene 6 palline bianche, 1 rossa e 3 nere.
C L'urna contiene 3 palline bianche e 5 rosse.
D L'urna contiene 6 palline bianche e 10 rosse.

5 Si sceglie a caso un numero tra i multipli di 9 minori di 1 000 000. Qual è la probabilità che si verifichi l'evento: «esce il numero 111111»?

A 0 C $\frac{1}{111111}$

B $\frac{1}{3}$ D 1

6 È stato effettuato un sondaggio su un campione di 1500 donne di età compresa tra i 25 e i 55 anni per conoscere la loro opinione su una rivista mensile dedicata alla salute. Si sono ottenuti i seguenti risultati.

	Occupate	Disoccupate
Giudizio positivo	450	276
Giudizio negativo	367	407

a. Quante sono le donne che hanno espresso un giudizio positivo?
Risposta: ....

b. Quante sono le donne disoccupate intervistate?
Risposta: ....

c. Scegliendo a caso una delle donne intervistate, qual è la probabilità che abbia espresso un giudizio negativo?
Risposta: ....

d. Scegliendo a caso una delle donne intervistate tra quelle che hanno espresso un giudizio positivo, qual è la probabilità che sia una donna occupata?
Risposta: ....

(Prova Invalsi 2014)

7 Da un mazzo di 52 carte da gioco (composto da 13 carte per ognuno dei semi: cuori, quadri, fiori, picche) sono stati tolti i 4 assi.

a. Si estrae una carta a caso. Qual è la probabilità che sia di cuori?
Risposta: ....

b. Da un mazzo di 52 carte uguale al precedente sono state tolte alcune carte di fiori. Dopo questa operazione la probabilità di estrarre, a caso, una carta di fiori è $\frac{6}{45}$. Quante carte di fiori sono state tolte?
Risposta: ....

(Prova Invalsi 2015)

8 Nel foglietto illustrativo contenuto nella confezione di un farmaco, alla voce «Effetti collaterali» si legge che:

- il 2% dei pazienti trattati con il farmaco ha accusato vertigini;
- il 7% dei pazienti trattati con il farmaco ha avuto bruciori di stomaco.

I due tipi di effetti collaterali sono indipendenti l'uno dall'altro.

a. Qual è la probabilità che un paziente che ha assunto il farmaco **non** abbia bruciori di stomaco? Esprimi il risultato in forma percentuale.

Risposta: .... %

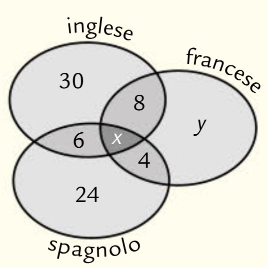
b. Qual è la probabilità che un paziente che ha assunto il farmaco manifesti **entrambi** gli effetti collaterali?

A 9% B 0,14% C 14% D 0,9%

(*Prova Invalsi 2015*)

9 Nelle classi prime di una scuola ci sono 100 studenti. Tutti studiano almeno una lingua straniera.

50 studiano inglese
40 studiano francese
40 studiano spagnolo
8 studiano solo l'inglese e il francese
6 studiano solo l'inglese e lo spagnolo
4 studiano solo il francese e lo spagnolo



a. Il numero x di studenti che studiano tutte e tre le lingue è ....

Il numero y di studenti che studiano solo il francese è ....

b. Qual è la probabilità che uno studente, preso a caso dall'elenco delle classi prime della scuola, studi solo l'inglese?

Risposta: ....

(*Prova Invalsi 2016*)

10 La corriera passa alle 6.30 alla fermata dove sale Giorgio. Nel 40% dei casi la corriera è in orario, nel 50% dei casi ha un ritardo di 5 minuti e nei rimanenti casi ha un ritardo di 10 minuti. Se Giorgio arriva alla fermata alle 6.34, che probabilità ha di prendere la corriera?

A 10% B 40% C 50% D 60%

(*Prova Invalsi 2011*)

11 La probabilità che su cinque lanci di un comune dado a sei facce non truccato si ottenga per cinque volte il numero 2 è:

A minore di $\frac{1}{6}$

B uguale a $\frac{1}{6}$

C maggiore di $\frac{1}{6}$

D i dati sono insufficienti a stabilirlo

12 In un dado truccato avente le facce numerate da 1 a 6, la probabilità di uscita di un numero è direttamente proporzionale al numero stesso. Quanto vale la probabilità che, lanciando un dado, esca il numero 3?

A $\frac{1}{3}$ B $\frac{1}{5}$ C $\frac{1}{6}$ D $\frac{1}{7}$

13 La probabilità di estrarre una pallina nera da un'urna che contiene 12 palline nere è $\frac{1}{5}$. Quante sono le palline contenute complessivamente nell'urna?

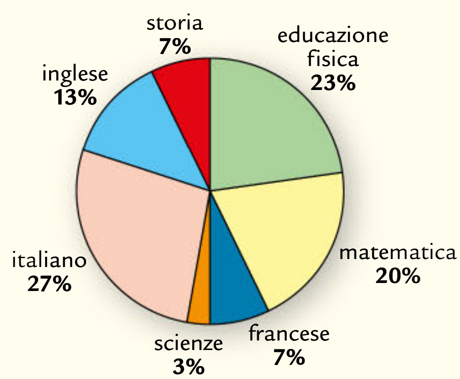
A 50

B 60

C 70

D Le informazioni date non sono sufficienti per stabilirlo.

14 In una scuola frequentata da 800 studenti si sceglie un campione di 300 studenti per un sondaggio sulla materia da essi preferita. I risultati del sondaggio sono rappresentati nel seguente diagramma.



a. Qual è il numero di studenti del campione che *non* hanno indicato come materia preferita matematica?

Risposta: ....

b. Qual è la probabilità che uno studente, scelto a caso dal campione, abbia indicato come materia preferita la matematica?

A $\frac{1}{15}$ B $\frac{1}{7}$ C $\frac{1}{5}$ D $\frac{1}{20}$

(*Prova Invalsi 2013*)

15 Viene lanciato un dado regolare a forma di ottaedro (solido regolare a otto facce), le cui facce sono numerate da 1 a 8. Qual è la probabilità che esca una faccia il cui numero è multiplo di 4?

A $\frac{1}{8}$ B $\frac{1}{4}$ C $\frac{3}{8}$ D $\frac{1}{2}$

16 In una lotteria i 4 premi sono assegnati per estrazioni successive, partendo dal 1° fino al 4°. Pietro ha acquistato uno solo dei 100 biglietti venduti. Egli è presente all'estrazione dei premi, e le estrazioni del 1° e del 2° premio lo vedono perdente. Qual è la probabilità che Pietro vinca il 3° o il 4° premio?

A $\frac{1}{97}$ B $\frac{1}{49}$ C $\frac{2}{97}$ D $\frac{1}{98}$

17 Una fabbrica utilizza due diversi macchinari, M_1 ed M_2, per produrre tondini. M_1 ha un indice di qualità uguale a 0,96 (cioè la probabilità che un tondino che esce da M_1 non sia difettoso è del 96%), mentre M_2 ha un indice di qualità uguale a 0,98.

a. La probabilità che un tondino esca da M_2 difettoso è:

A 0,98 B 0,02 C 0,04 D 0,96

b. Per la realizzazione di tondini metallici, M_1 e M_2 lavorano in serie, cioè ogni tondino viene lavorato prima da M_1 e poi da M_2. Supponiamo che gli eventi «M_1 produce un tondino non difettoso» ed «M_2 produce un tondino non difettoso» siano tra loro indipendenti; allora la probabilità che un tondino non sia difettoso alla fine del ciclo di produzione (cioè dopo essere stato lavorato sia da M_1 sia da M_2) è:

A 94,08% B 6% C 1,94% D 98%

(*Prova Invalsi 2013*)

18 Una password è costituita da 3 lettere minuscole dell'alfabeto italiano e da 3 cifre da 1 a 9, inclusi 1 e 9. Le lettere e le cifre non sono ripetute. Scegliendo a caso una password, qual è la probabilità che inizi con una lettera?

A $\frac{7}{10}$ B $\frac{8}{11}$ C $\frac{11}{13}$ D $\frac{17}{30}$

19 Un dado non truccato è stato lanciato 70 volte di seguito. La seguente tabella riporta la frequenza con cui ciascun numero è uscito.

Numero	Frequenza
1	11
2	10
3	11
4	16
5	9
6	13

Indica se ciascuna delle seguenti affermazioni è vera o falsa.

a. Poiché il 5 è uscito meno volte, la probabilità che esca 5 nel lancio successivo al settantesimo è maggiore rispetto agli altri numeri. V F

b. La probabilità che esca 5 nel lancio successivo al settantesimo è uguale a quella che esca 6. V F

c. La probabilità che nei due lanci successivi al settantesimo esca entrambe le volte 6 è minore del 3%. V F

d. Se al 71° lancio esce 3, la probabilità che al 72° lancio esca il numero consecutivo a 3 è superiore al 7% V F

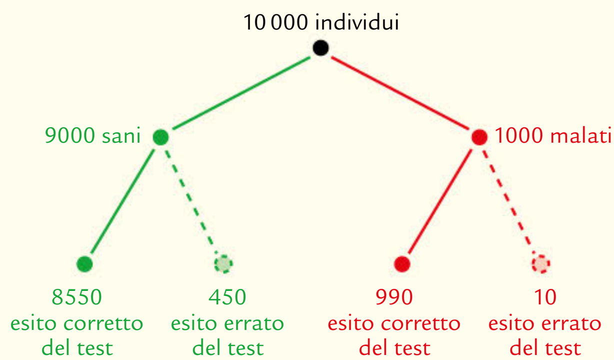
e. La probabilità che nel settantaduesimo lancio esca un numero consecutivo a quello uscito nel settantunesimo è uguale a $\frac{5}{36}$. V F

20 Quanti sono i possibili ambi che si possono giocare al gioco del lotto?

A 2000 B 2005 C 4000 D 4005

21 Si sa che in una popolazione di 10 000 individui il 10% è affetto da una malattia, mentre il 90% è sano. Il test che diagnostica la presenza della malattia è affidabile solo parzialmente: nel 5% dei casi rileva la malattia su un individuo sano e nell'1% dei casi non rileva la malattia su un individuo malato. Il seguente diagramma riassume la situazione.

a. Utilizzando i dati del diagramma ad albero, completa la seguente tabella.



	Esito corretto del test	Esito errato del test	Totale
Sani	....	450	....
Malati	....	....	....
Totale	9540	....	10 000

b. Qual è la probabilità che l'esito dei test sia corretto per una persona scelta a caso tra quella popolazione?

A 90,0% B 97,0% C 95,4% D 85,5%

c. Qual è la probabilità che un individuo, preso a caso tra tutti quelli che hanno avuto un esito corretto al test, sia sano? Scrivi il risultato in percentuale, con una cifra dopo la virgola.

Risposta: ....%

(Prova Invalsi 2012)

22 Due urne A e B contengono ciascuna tre bigliettini numerati con i numeri 1, 2 e 3. Si estrae un bigliettino dall'urna A e poi un bigliettino dall'urna B.

a. Completa l'elenco di tutti i possibili esiti che si possono ottenere:

(1, 1), (1, 2), (1, 3), (2, 1), ...

b. Si estrae un bigliettino dall'urna A e poi uno dall'urna B e si esegue la somma dei due numeri estratti. Fra tutte le possibili somme che si possono ottenere, qual è la più probabile?

Risposta: ....%

(Prova Invalsi 2017)

23 In una gara motociclistica la moto M ha probabilità di vincere la gara:

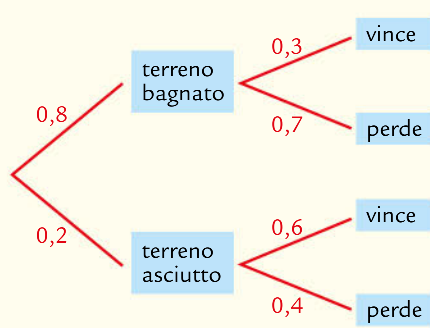
- 0,3 se il terreno è bagnato;
- 0,6 se il terreno è asciutto.

La probabilità che il giorno della gara il terreno sia asciutto è 0,2.

Il diagramma può aiutare a determinare, per esempio, la probabilità che il terreno sia asciutto e che la moto M perda la gara. Essa è $0{,}2 \cdot 0{,}4 = 0{,}08$.
Qual è la probabilità che la moto M vinca la gara?

Risposta: ....%

(Prova Invalsi 2017)



1 Un test è costituito da 20 quesiti, ognuno dei quali ha 5 opzioni di risposta. In quanti modi diversi si può rispondere all'intero test, assegnando esattamente una risposta a ciascun quesito?

A $20 \cdot 5$ B 20^5 C 5^{20}
D Nessuna delle altre risposte è corretta E $20 \cdot 19 \cdot 18 \cdot 17 \cdot 16$

(Test d'ingresso CISIA, Facoltà di Scienze, 2016)

2 La probabilità che, lanciando due volte un dado a 6 facce, il numero ottenuto al primo lancio sia minore del numero ottenuto al secondo lancio è:

A $\frac{5}{6}$ B $\frac{5}{12}$ C $\frac{3}{7}$ D $\frac{1}{2}$ E $\frac{1}{3}$

(Test d'ingresso CISIA, Facoltà di Scienze, 2016)

3 Avendo 6 palline numerate da 1 a 6, di cui 3 di colore rosso e 3 di colore blu, in quanti modi è possibile disporle in fila in modo che due palline adiacenti non abbiano lo stesso colore?

A 72 B 90 C 360 D 720 E 729

(Facoltà di Statistica e informatica per l'Azienda, Università di Trieste, 2016)

4 In una partita di calcio tra due squadre A e B si supponga di sapere che la probabilità che A vinca o pareggi è 0,8 mentre la probabilità che B vinca o pareggi è 0,4. Qual è la probabilità che si verifichi un pareggio?

A 0,2 B 0,1 C 0,4 D 0,8 E 0

(Facoltà di Statistica e informatica per l'Azienda, Università di Trieste, 2016)

5 In una scatola si trovano nove palline, su ciascuna delle quali è scritta una delle cifre tra 1 e 9. Si estrae a caso una pallina dalla scatola e si legge il numero, poi si rimette la pallina nella scatola. Si fa una nuova estrazione casuale e si legge il numero. Qual è la probabilità che il prodotto dei numeri letti sulle due palline estratte sia dispari?

A $\frac{56}{81}$ B $\frac{1}{3}$ C $\frac{3}{4}$ D $\frac{25}{81}$ E $\frac{1}{4}$

(Test d'ingresso CISIA, Facoltà di Scienze, 2015)

6 Due arcieri tirano ciascuno una sola freccia verso un bersaglio. Un arciere colpisce il bersaglio con probabilità del 30%, l'altro con probabilità del 40%. Qual è la probabilità che il bersaglio venga colpito almeno una volta?

A 35% B 46% C 58% D 42% E 70%

(Test d'ingresso CISIA, Facoltà di Scienze, 2014)

7 Ho 8 vaschette di gelato, con gusti tutti diversi tra loro: tra essi, fragola e liquirizia. In quanti modi diversi posso servire gelati con tre gusti differenti, se escludo di mettere insieme fragola e liquirizia?

A 50 B 56 C 21 D 44 E 42

(Test d'ingresso CISIA, Facoltà di Scienze, 2013)

8 Un mazzo di carte è formato da 4 re, 4 donne e 4 fanti. Qual è la probabilità che, dopo averle mescolate, le prime quattro carte del mazzo siano quattro donne?

A $\frac{1}{864}$ B $\frac{1}{495}$ C $\frac{1}{64}$ D $\frac{1}{3}$ E $\frac{1}{81}$

(Test d'ingresso CISIA, Facoltà di Scienze, 2013)

9 Considera tutti gli anagrammi della parola FUNGHI, ovvero tutte le parole (anche prive di senso) che si ottengono permutando le sei lettere. Tra esse, quante sono le parole che non cominciano per F?

A 360 B 600 C 720 D 120

(Test di ingresso, Facoltà di Scienze, 2009)

10 Utilizzando solo i caratteri «0» e «1», quante sequenze diverse di 5 caratteri si possono scrivere?

A 50 B 10 C 25 D 32

(Test di ingresso per i corsi di laurea scientifici, 2008)

11 Aldo, Bea, Carlo, Dario, Ebe, Franco vanno in treno e trovano uno scompartimento a sei posti libero. Considerando che Aldo e Bea devono stare vicino al finestrino, quanti modi diversi hanno i sei amici di disporsi nello scompartimento?

A 48 B 4 C 240 D 8 E 10

(Prova di ammissione, Ingegneria, 2007)

12 Il Circolo Canottieri Santerno è formato da sei rematori, tutti ugualmente bravi e affiatati fra loro. Il circolo deve mandare una rappresentanza di quattro atleti al campionato regionale. In quanti diversi modi può essere formata una tale rappresentanza?

A 720 B 5 C 15 D 4 E 6

(Prova di ammissione, Ingegneria, 2009)

13 Durante una vacanza, sette amici prendono in affitto due automobili. Una di esse ha due posti, l'altra ne ha cinque. In quanti modi differenti possono distribuirsi i sette amici sulle due automobili?

A 21 B 14 C 28 D 35

(Test di ingresso, Facoltà di Scienze, 2009)

14 Qual è la probabilità che lanciando 6 volte una moneta escano esattamente 4 «testa»?

A $\frac{15}{64}$ B $\frac{1}{64}$ C $\frac{15}{16}$ D $\frac{1}{16}$ E $\frac{5}{32}$

(Prova di ammissione, Corso di laurea in Medicina 2008)

15 Sulle sei facce di un dado compaiono le cifre da 1 a 6. Si lancia un dado due volte; qual è la probabilità che il 3 non esca al primo lancio ed esca al secondo?

A $\frac{18}{36}$ B $\frac{6}{36}$ C $\frac{25}{36}$ D $\frac{5}{36}$ E $\frac{1}{36}$

(Test di ingresso, Facoltà di Scienze, 2010)

16 Mario lancia quattro volte una moneta non truccata. Qual è la probabilità che esca «testa» in almeno tre lanci?

A $\frac{5}{16}$ B $\frac{1}{8}$ C $\frac{1}{4}$ D $\frac{9}{16}$

(Test di ingresso, Facoltà di Scienze, 2009)

17 Due sacchetti contengono ciascuno i numeri 1, 2, 3, 4, 5. Si estrae un numero da ciascun sacchetto. Qual è la probabilità che i due numeri siano entrambi dispari?

A $\frac{6}{5}$ B $\frac{3}{5}$ C $\frac{4}{5}$ D $\frac{9}{25}$

(Test di ingresso, Facoltà di Scienze, 2008)

18 Il codice per aprire un lucchetto è costituito da una sequenza di quattro cifre (da 0 a 9). Ho dimenticato il codice, ma mi ricordo che le cifre sono tutte distinte e che tra le prime tre cifre ci sono sicuramente i numeri 6 e 9. Quante sequenze di quattro numeri dovrei provare per essere certo di aprire il lucchetto?

A 100 B 118 C 336 D 600

(Test di ingesso, Facoltà di Scienze 2008)

19 La probabilità che, lanciando due dadi a 6 facce, si ottenga come somma 3 è:

A $\frac{1}{3}$ B $\frac{1}{12}$ C $\frac{1}{18}$ D $\frac{1}{36}$

(Test di ingresso per i corsi di laurea scientifici, 2008)

Compito di realtà 1

Totocalcio

L'obiettivo del gioco del Totocalcio è indovinare gli esiti delle 14 partite inserite in schedina. Per ogni singola partita inserita in schedina si deve marcare «1» se si pronostica la vittoria della squadra che gioca in casa, «X» se si prevede un pareggio, «2» se si prevede la vittoria della squadra ospite. A ogni concorso settimanale il Totocalcio premia la colonna con tutte le 14 partite indovinate, ma anche (in misura minore) le colonne con 13 oppure 12 partite indovinate (se si fa «14» viene pagato solo il «14», non anche il «13» e il «12»; analogamente, se si fa «13» non viene pagato anche il «12»).

1 Calcola il numero di tutti i modi possibili di compilare una colonna.

2 Quanto dovrebbe spendere uno scommettitore per giocare tutte le possibili colonne senza «2»? Il costo di una colonna è 0,50 euro.

3 Verifica che le colonne vincenti settimanalmente sono in tutto 393.

Ipotizziamo ora che per ciascuna partita la probabilità che vincano i padroni di casa sia del 50%, mentre la probabilità che vincano gli ospiti sia del 20%.

4 Qual è la probabilità che tra le 14 partite in schedina non si registri alcuna vittoria esterna, cioè nessun «2»? Barra la risposta giusta tra le quattro riportate qui di seguito.

A $\left(\frac{5}{10}\right)^{14}+\left(\frac{3}{10}\right)^{14} \simeq 0{,}01\%$

B $\left(\frac{5}{10}+\frac{3}{10}\right)^{14} \simeq 4{,}40\%$

C $1-\left(\frac{2}{10}\right)^{14} \simeq 99{,}99\%$

D $1-\left(\frac{5}{10}+\frac{3}{10}\right)^{14} \simeq 95{,}60\%$

5 Qual è la probabilità di vincere (cioè di fare almeno 12) giocando una colonna interamente costituita da «1»?

Supponi infine che uno scommettitore scelga a caso i 14 esiti di una colonna e che i tre esiti di ciascuna partita siano equiprobabili.

6 Qual è probabilità di vincere (cioè di fare almeno 12)? Tale probabilità è maggiore o minore di quella di fare un «terno al lotto», cioè che si giochino tre numeri e questi facciano parte dei cinque estratti?

Risposte p. 220

Compito di realtà 2

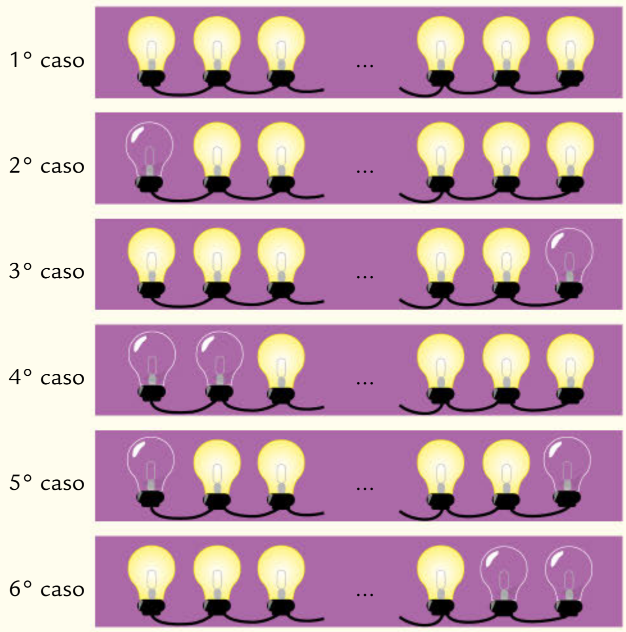
Luci per l'albero di Natale

Assunta e Raffaele stanno finendo di addobbare l'albero di Natale. Per completarlo, devono acquistare un set di luci da disporre intorno all'albero. Sono però indecisi se acquistare un set di 10 luci collegate *in serie* oppure un set di 10 luci collegate *in parallelo*. Il primo è poco costoso, 5 euro, ma ha un difetto: una sola lampadina fulminata rende inservibile tutto l'insieme. L'alternativa è più costosa, 10 euro, ma con un vantaggio: una lampadina fulminata non compromette il regolare funzionamento delle altre nove. Supponi che la probabilità che una lampadina si fulmini sia del 10% (equivalentemente, che la probabilità che essa funzioni regolarmente sia del 90%).

1 Qual è la probabilità che il set di lampadine in serie funzioni regolarmente? In altre parole: qual è la probabilità che *tutte* le dieci lampadine, nessuna esclusa, stiano accese insieme?

2 Consideriamo ora il caso delle luci collegate in parallelo. Assunta e Raffaele considerano l'effetto luminoso soddisfacente se almeno 8 delle 10 lampadine funzionano, purché le eventuali lampadine fulminate assicurino la regolare continuità della sequenza luminosa; ciò si verifica nei sei casi (non equiprobabili!) rappresentati in figura.

Qual è la probabilità che le luci producano l'effetto desiderato?



Assunta ha un'idea: acquistare *due* set di luci in serie e installarne uno soltanto, confidando nel fatto che *almeno uno* di essi si accenda regolarmente. La spesa infatti pareggia quella per le luci collegate in parallelo ma, stando ai suoi conti, la probabilità di vedere l'albero illuminato per le festività natalizie supera l'altra.

3 Verifica che Assunta ha ragione.

Anche Raffaele riconsidera attentamente la questione. Diversamente da quanto convenuto al punto 2, Raffaele ritiene che l'effetto luminoso complessivo sia accettabile se si fulmina al massimo una delle 10 lampadine collegate in parallelo, indipendentemente dalla posizione della lampadina.

4 Verifica che, in questo caso, è preferibile acquistare un set di 10 luci collegate in parallelo anziché due del tipo economico.

Risposte p. 220

Risposte alle prove proposte nel volume

Prove di autoverifica

Unità 1

1. **a.** derivazione; **b.** densità; **c.** coesistenza; **d.** composizione; **e.** derivazione o densità; **f.** coesistenza

2. Vedi tabella:

Anno	2012	2013	2014	2015	2016	2017
Numero studenti	800	840	798	840	924	1040
Indice a base fissa	100%	105%	99,75%	105%	115,5%	130%
Indice a base mobile		105%	95%	105,26%	110%	112,55%

3. **a.** 9,4; **b.** 7; **c.** $\simeq$ 4,03.

4. **a.** La media e la mediana aumentano di 1 punto, mentre la deviazione standard resta invariata.
b. La media, la mediana e la deviazione standard risultano moltiplicate di 1,1 rispetto alle precedenti.

5. **a.** Compresa tra 81,3 kg e 82,7 kg; **b.** compresa tra 81,5 kg e 82,5 kg.

6. **a.** Compresa tra il 68,9 e l'83,1; **b.** compresa tra il 64,9% e l'87,1%.

Unità 2

1. **a.** Vedi tabella:

Y / X	Negativo	Positivo	Ottimo	Totale
Giovani	4,55	7,95	12,5	25
Adulti	9,45	16,55	26	52
Anziani	6	10,5	16,5	33
Totale	20	35	55	110

b. Le frequenze della tabella data sono molto lontane dalle frequenze della tabella teorica di indipendenza, dunque X e Y sono dipendenti. In effetti si ricava $\chi^2 \simeq 86{,}63$ e $\chi^2_{\text{normalizzato}} \simeq 0{,}394$. La connessione tra X e Y è circa il 39,4% della massima connessione possibile, dunque c'è una discreta connessione.

2. **a.** 36; **b.** $\frac{328}{715} \simeq 46\%$

3. $\left(\frac{5}{2}, 3\right)$; **b.** $y = -\frac{7}{5}x + \frac{13}{2}$

4. **a.** $-0{,}98$; **b.** $y = -0{,}82x + 209{,}2$; **c.** 156

Unità 3

1. V, F, V, V

2. $n = 0$

3. **a.** 2024; **b.** 253

4. 15; 12

5. **a.** $4^8 = 65536$; **b.** $\frac{8!}{(2!)^4} = 2520$

6. Esclusi il portiere e 6 giocatori che devono comunque occupare i tre reparti (due giocatori per ognuno dei tre reparti) restano 4 giocatori da disporre in campo a piacere. Abbiamo dunque $C^*_{3,4} = \binom{6}{4} = \binom{6}{2} = 15$ schemi di gioco possibili.

7. 280

Unità 4

1. F, V, V, F

2. **a.** La probabilità che Verdi sia l'ultimo a parlare è $\frac{1}{5}$, quindi la probabilità cercata è $\frac{4}{5} = 80\%$.

b. Ci sono $5! = 120$ modi diversi in cui i cinque possono prendere la parola, quindi la probabilità richiesta è $\frac{2}{120} = \frac{1}{60}$.

3. Poiché $p(A \cup B) = p(A) + p(B) - p(A)p(B)$ (in virtù dell'indipendenza di A e B), abbiamo l'equazione $\frac{3}{4} = 2k - k^2$ da cui $k = \frac{1}{2}$ (valore accettabile) e $k = \frac{3}{2}$ (valore da scartare).

4. **a.** $\frac{3}{9} \cdot \frac{3}{9} = \frac{1}{9}$; **b.** $\frac{\binom{3}{2}}{\binom{9}{2}} = \frac{1}{12}$.

5. **a.** 6,8%; **b.** $\frac{8}{17} \simeq 47\%$.

6. Circa il 74%.

7. I casi possibili sono $2^6 = 64$: i casi favorevoli sono tanti quanti gli anagrammi della «parola» CCCTTT, ossia $\frac{6!}{3! \cdot 3!} = 20$. In conclusione: $\frac{20}{64} = \frac{5}{16}$. In alternativa si può giungere allo stesso risultato calcolando la probabilità di avere 3 successi in 6 prove ripetute e indipendenti, in ciascuna delle quali la probabilità di successo è $\frac{1}{2}$.

8. La probabilità che la sfida termini 0-0 è pari a $\frac{1}{4} \cdot \frac{1}{4} = \frac{1}{16}$, così come la probabilità che si concluda 2-2; mentre la probabilità che finisca 1-1 è uguale a $\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$. In definitiva, la probabilità richiesta è $\frac{1}{16} + \frac{1}{4} + \frac{1}{16} = \frac{3}{8}$.

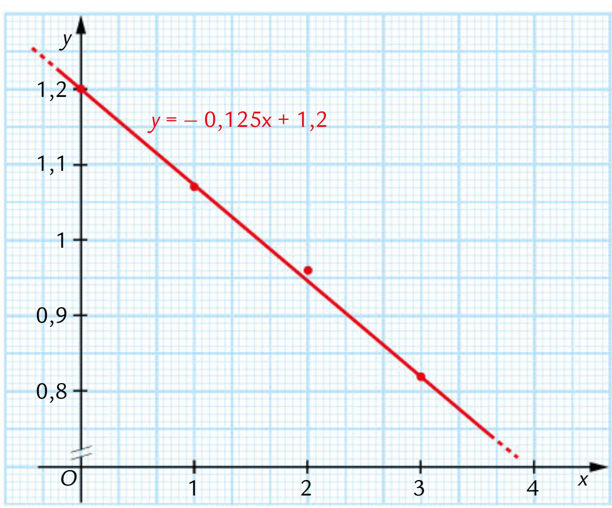
Compiti di realtà

Tema E

Compito di realtà 1

1. Risposta A.

2. Vedi la figura.



3. Circa 0,64 g/L (arrotondando alla seconda cifra decimale).

4. Dopo 5 ore e 36 minuti.

5. Dopo 9 ore e 36 minuti.

Compito di realtà 2

1. 26 bottiglie, 2 bottiglie.
2. 35%.
3. $\overline{x} = 73$, $\sigma = \sqrt{3{,}6} \simeq 1{,}9$.
4. A un livello di confidenza del 95%, l'intervallo di confidenza del contenuto medio delle bottiglie è [72,4; 73,6]; poiché il numero 75 non è compreso nell'intervallo di confidenza, dobbiamo rifiutare quanto dichiarato dall'azienda.
5. L'intervallo di confidenza in questo caso diviene [72,2; 73,8] e la risposta non cambia.
6. La media $\overline{x}$ aumenta di 1 cl, ma la deviazione standard e l'errore standard restano invariati; l'intervallo di confidenza al 95% diviene [73,4; 74,6] e bisogna ancora rifiutare quanto dichiarato dall'azienda.
7. 1,4 cl.

Tema F

Compito di realtà 1

1. 3^{14}.
2. $0{,}5 \cdot 2^{14} = 8192$ euro.
3. $\binom{14}{0} + 2\binom{14}{1} + 4\binom{14}{2} = 1 + 2 \cdot 14 + 4 \cdot 91 = 393$.
4. Risposta B.
5. La probabilità che almeno 12 partite si concludano con la vittoria della squadra di casa è data da:
$\binom{14}{14}\left(\frac{1}{2}\right)^{14}\left(\frac{1}{2}\right)^{0} + \binom{14}{13}\left(\frac{1}{2}\right)^{13}\left(\frac{1}{2}\right)^{1} + \binom{14}{12}\left(\frac{1}{2}\right)^{12}\left(\frac{1}{2}\right)^{2} \simeq 0{,}65\%$.
6. $\frac{393}{3^{14}}$ (*vedi* i punti **1** e **3**) oppure, utilizzando il modello delle prove ripetute:
$\binom{14}{14}\left(\frac{1}{3}\right)^{14}\left(\frac{2}{3}\right)^{0} + \binom{14}{13}\left(\frac{1}{3}\right)^{13}\left(\frac{2}{3}\right)^{1} + \binom{14}{12}\left(\frac{1}{3}\right)^{12}\left(\frac{2}{3}\right)^{2}$
Tale probabilità è all'incirca $\frac{1}{12170}$ (di poco inferiore), mentre la probabilità di fare un terno al lotto è $\frac{1}{11748}$: dunque la probabilità di vincere al Totocalcio è inferiore rispetto a quella di fare un terno al lotto.

Compito di realtà 2

1. $\left(\frac{9}{10}\right)^{10} \simeq 35\%$.
2. La probabilità che tutte le lampadine funzionino (cioè che si verifichi il primo caso rappresentato) è uguale a $\left(\frac{9}{10}\right)^{10}$; la probabilità che solo una delle due lampadine alle estremità si fulmini (cioè che si verifichino il secondo o il terzo caso rappresentato) è uguale a $2\left(\frac{1}{10}\right) \cdot \left(\frac{9}{10}\right)^{9}$; infine, la probabilità che si fulminino due lampadine senza interrompere la continuità della sequenza luminosa (cioè che si verifichi uno degli ultimi 3 casi rappresentati) è $3\left(\frac{1}{10}\right)^{2} \cdot \left(\frac{9}{10}\right)^{8}$. La probabilità richiesta è dunque:
$$\left(\frac{9}{10}\right)^{10} + 2\left(\frac{1}{10}\right)^{1} \cdot \left(\frac{9}{10}\right)^{9} + 3\left(\frac{1}{10}\right)^{2} \cdot \left(\frac{9}{10}\right)^{8} \simeq 44\%.$$
3. La probabilità che entrambi i set di luci collegate in serie siano inservibili è circa $0{,}6513^2 \simeq 42\%$, quindi la probabilità che almeno uno di essi resti illuminato regolarmente per le festività è intorno al 58%, maggiore della probabilità calcolata al punto **2**.
4. Alla probabilità (già calcolata) che tutte le 10 lampadine restino accese occorre sommare la probabilità che solo una (qualunque) delle 10 si fulmini:
$$\left(\frac{9}{10}\right)^{10} + 10\left(\frac{1}{10}\right)^{1} \cdot \left(\frac{9}{10}\right)^{9} \simeq 74\%.$$
La probabilità così ottenuta è maggiore di quella calcolata da Assunta (vedi il punto **3**); in questo caso risulta quindi preferibile l'acquisto del set di 10 luci collegate in parallelo.

Indice analitico

A, B

C

D

E

F, G, H

I

L

M

Petrini

internet: deascuola.it

e-mail: info@deascuola.it

Redattore responsabile: Monica Martinelli
Redazione: Giovanni Malafarina
Redazione multimediale: Rachele Ambrosetti
Progetto grafico: Carla Devoto, Marco Satta
Impaginazione: M.T.M.
Ricerca iconografica: Laura Fiorenzo
Ricerca iconografica per la copertina: Alice Graziotin
Copertina: Silvia Bassi, Simona Speranza
Disegni: Leprechaun

Art Director: Nadia Maestri

Le animazioni «Con GeoGebra» sono state realizzate dalla dott.ssa Simona Riva
Le Videolezioni sono a cura di Marco D'Isanto (*Lezionidimate*)

Attività interattive e laboratori realizzati con il software *GeoGebra* (www.geogebra.org)
Microsoft Excel è un marchio depositato di *Microsoft Corporation*
Per ogni informazione relativa agli esercizi tratti dalle gare *Kangourou della Matematica* visitare il sito www.kangourou.it

Numerosi insegnanti e colleghi hanno contribuito al perfezionamento dell'opera, fornendo agli Autori preziosi spunti didattici, stimolanti osservazioni e puntuali segnalazioni. A tutti è indirizzata la loro riconoscenza.
Leonardo Sasso ringrazia in modo particolare gli insegnanti dell'Istituto Tecnico «Guglielmo Marconi» di Verona, dell'Istituto Tecnico «Guglielmo Marconi» di Padova e dell'Istituto Tecnico «Galileo Ferraris» di Verona. Enrico Zoli ringrazia i colleghi dell'Istituto Tecnico «Luigi Bucci» di Faenza.

1ª edizione: gennaio 2019
Printed in Italy

Le fotografie di questo volume sono state fornite da: Archivio Dea Picture Library; Shutterstock; iStockphoto

Foto di copertina: Getty Images