Dati e previsioni

Gli argomenti del nucleo “Dati e previsioni” già presenti nel curricolo della scuola primaria e secondaria di primo grado per essere affrontati in modo prevalentemente intuitivo, ricevono in questo curricolo una organizzazione più strutturata, adatta all’età degli studenti. A livello di ciclo secondario, infatti, le conoscenze da acquisire, pur basandosi sempre sulla classificazione dei caratteri (qualitativi, sconnessi ed ordinati; quantitativi, discreti e continui) consistono nel mettere in evidenza la possibilità sia di sintetizzare la distribuzione statistica semplice con una pluralità di valori medi tra i quali scegliere in modo opportuno, tenendo conto della definizione di ciascuno di essi, sia di misurare la variabilità del carattere nel collettivo studiato. Lo studio della variabilità non è però fine a se stesso, ma ha uno scopo interpretativo. La sua giustificazione, perciò, richiede l’entrata in campo almeno di un secondo carattere con un ruolo esplicativo rispetto al primo. Ciò pone due problemi concettuali differenti: lo studio della connessione o interdipendenza statistica fra caratteri e lo studio (ma solo se entrambi i caratteri sono quantitativi), del loro variare simultaneo. Si introducono così il concetto di correlazione (concordanza e discordanza) e la ricerca di una semplice espressione funzionale che descriva la legge di dipendenza fra le variabili osservate (regressione lineare).

Non a caso le conoscenze che riguardano la probabilità seguono in questo curricolo quelle di statistica. Ciò suggerisce come sia opportuno iniziare la trattazione di tale tema, avendo già a disposizione motivazioni ed esempi accattivanti che permettono di introdurre la probabilità, le sue proprietà di base e le prime regole di calcolo. Anche il passaggio dagli eventi alle variabili aleatorie è favorito da questo approccio che vede “semplici distribuzioni di probabilità” introdotte su una base ormai solida, offerta dallo studio delle distribuzioni semplici. Il concetto di probabilità condizionata e il teorema di Bayes hanno importanti applicazioni concrete in vari settori.

La vita quotidiana e le proposte dei mezzi di comunicazione offrono sempre più l’opportunità di motivare gli studenti ad affrontare temi di statistica e di probabilità. L’insegnante potrà utilmente sfruttare la curiosità innata degli studenti per far loro raccogliere informazioni quantitative su argomenti che li coinvolgono direttamente, ma anche su argomenti che riguardano la fisica, l’economia, la storia, la geografia e che richiedono o la ripetizione della stessa esperienza o la gestione di un collettivo di osservazioni empiriche. Ciò che va evitato è di introdurre la statistica come un insieme di calcoli su numeri inventati e senza significato in un contesto reale. La statistica e la probabilità sono un valido aiuto per il cittadino e promuovono l’acquisizione di abilità utili nella vita quotidiana solo se aiutano a comprendere la realtà ed in particolare quel suo aspetto “disorientante” che è la variabilità dei fenomeni.

Tra l’altro operare in contesti quantitativi coinvolgenti ed interessanti, perché derivanti da fenomeni in parte conosciuti, può essere un utile supporto per passare dalla realtà alla sua astrazione simbolica, introducendo gradualmente il linguaggio formale della matematica, in modo che gli studenti arrivino a percepire che le formule non sono altro che un linguaggio che ha il vantaggio della concisione e della non ambiguità.

Elenco delle attività

Arrivare a scuola

Livello scolare: 1° biennio

Abilità interessate	Conoscenze	Nuclei coinvolti	Collegamenti esterni
Comprendere la differenza fra caratteri qualitativi, quantitativi discreti e continui. Passare dai dati grezzi alle distribuzioni statistiche di frequenze ed alle corrispondenti rappresentazioni grafiche.	Distribuzioni delle frequenze a seconda del tipo di carattere. Frequenze assolute, relative, percentuali. Principali rappresentazioni grafiche per le distribuzioni di frequenze.	<u>Dati e previsioni</u> Spazio e figure Misurare Risolvere e porsi problemi Laboratorio di matematica	Lingua italiana

Contesto
Extramatematico, sociale.

Il contesto di questa attività è di tipo extramatematico, in particolare riguarda l'ambito sociale.
Questa unità riguarda l'"arrivare a scuola". L'attività prevede la problematizzazione delle situazioni, la predisposizione del questionario, la raccolta dei dati, la loro elaborazione e rappresentazione grafica, la loro interpretazione.
Questa attività è consigliata nel 1° biennio, come uno dei possibili approcci al nucleo, in raccordo con quanto sviluppato nella scuola secondaria di primo grado. Il contesto è extramatematico e si basa sull'esperienza della vita quotidiana, ponendo i ragazzi di fronte all'analisi di problemi concreti.

Descrizione dell'attività
Rendere protagonisti gli studenti, con tematiche che li coinvolgono in prima persona, si è mostrato una metodologia vincente sotto l'aspetto della motivazione. Qui si propone un tema già largamente sperimentato nella scuola superiore: "l'arrivare a scuola".
L'argomento si presta per affrontare il problema del trasporto in cui tutti gli studenti sono coinvolti. Nel caso in cui l'indagine riguardi tutto l'istituto, il momento della restituzione dei dati potrà, ad esempio, avviare un dibattito sul tema dei ritardi nell'ambito dell'assemblea degli studenti.

<u>Prima fase</u>
Per accertare la conoscenza da parte degli allievi dei concetti di unità statistica, di collettivo statistico, come insieme delle unità indagate, di carattere e sue modalità, l'insegnante propone una attività di gruppo, chiedendo agli studenti di preparare una o più schede di lavoro sulla distinzione tra unità statistica, carattere e modalità. Per saggiare la conoscenza degli studenti con riferimento alla individuazione delle diverse tipologie di carattere il docente può sottoporli a prove di verifica (si vedano le proposte in fondo all'unità).

<u>Seconda fase</u>
Nel presupposto che il questionario sia lo strumento più adatto per raccogliere i dati, l'insegnante propone una attività di gruppo finalizzata alla sua preparazione.
I questionari costruiti dai ragazzi sono poi confrontati e discussi per fare emergere una versione definitiva, completa e corretta.

L'attenzione viene focalizzata, in particolare, sulla necessità di porre domande semplici, chiare e precise in modo da evitare che una domanda possa essere interpretata in modi diversi o risulti non completa nella sua formulazione.
E' preferibile che le domande siano a risposta chiusa, in modo che le alternative di risposta siano previste in anticipo; quando non è possibile individuarle tutte si introduce la modalità "altro".
Si consiglia di limitare le domande a risposta aperta a quando si richiede un parere personale. Perché i dati abbiano la stessa natura e siano confrontabili, occorre riferirsi ad opportune unità di misura. In questo caso particolare, ad esempio, l'unità di misura è necessaria con riferimento a: il tempo necessario per effettuare il percorso casa-scuola e il momento in cui ci si alza al mattino. Dato che il tempo è un carattere quantitativo continuo le domande che lo riguardano vengono espresse in intervalli. In questo modo è possibile tenere conto anche dell'errore di approssimazione da cui le misure sono affette.
Con riferimento al tempo abitualmente impiegato per arrivare a scuola, si è deciso di utilizzare la classificazione proposta dall'ISTAT nella scheda del censimento della popolazione. Ad eccezione della prima e dell'ultima classe, ciò significa utilizzare classi chiuse sia a destra sia a sinistra; con la conseguenza che ogni estremo viene inteso come il punto centrale di un intervallo di ampiezza unitaria. Rispetto al "momento" in cui ci si alza al mattino, si considera che si alzi alle 6.29 chi lo fa tra le 6.29.00 e le 6.29.59.
L'attenzione alla formulazione del questionario può costituire un utile momento di interdisciplinarità, in particolare con il collega di italiano.
L'insegnante orienta la discussione in modo che vengano prescelte le domande che soddisfano le finalità conoscitive del fenomeno e contemporaneamente in modo che siano presenti caratteri di tipo qualitativo, sconnesso e ordinato, e quantitativo, discreto e continuo, con l'obiettivo di mettere in evidenza il diverso trattamento dei dati a seconda del tipo di carattere.
Per collaudare il questionario predisposto, in analogia con le procedure usualmente adottate nelle indagini statistiche, risulta opportuno provare il questionario in un'altra classe.
Effettuata la somministrazione pilota ed apportate le eventuali modifiche si ottiene il questionario definitivo col quale eseguire l'indagine. Di seguito si fornisce un esempio di questionario frutto di un'attività realmente effettuata in un istituto scolastico superiore.

<u>Sezione 1: Dati relativi al rispondente</u>

Domanda 1: Qual è la tua età (in anni compiuti)?

□ Meno di 14
□ 14
□ 15
□ 16
□ 17
□ 18
□ 19
□ più di 19

Domanda 2: Sesso

□ Femmina
□ Maschio

Domanda 3: Qual è il tuo comune di residenza?

____

Domanda 4: Quale classe stai frequentando?

□ Prima

□ Seconda

□ Terza

□ Quarta

□ Quinta

Sezione 2: Andare a scuola

Domanda 5: A che ora ti alzi abitualmente il mattino per venire a scuola?

□ Prima delle 6.00

□ Dalle 6.00 alle 6.29

□ Dalle 6.30 alle 6.59

□ Dalle 7.00 alle 7.29

□ Dalle 7.30 in poi

Domanda 6: Con quale tipo di mezzo abitualmente raggiungi la scuola?

□ A piedi

□ Solo con mezzi privati (bicicletta, motorino, automobile,…)

□ Solo con i mezzi pubblici (bus urbano e/o extraurbano, treno,…)

□ Con mezzi sia privati che pubblici

Domanda 7: Quanto tempo impieghi abitualmente per arrivare a scuola, partendo da casa?

□ Fino a 15 min.

□ Da 16 a 30 min.

□ Da 31 a 45 min.

□ Da 46 a 60 min.

□ Oltre 60 min.

Domanda 8: Abitualmente arrivi a scuola in orario?

Si No

Terza fase

Una volta somministrato il questionario, è necessario elaborare i dati. Si ritiene utile, a questo livello scolare, procedere allo spoglio manuale di un adeguato numero di questionari per mettere in evidenza la difficoltà dell'operazione e il rischio, in questa fase molto delicata, di introdurre degli errori sia di lettura che di registrazione.
Si propone di utilizzare schemi di spoglio del tipo di Tabella 1 e di Tabella 2.
Qui, come nei prossimi esempi i dati riguardano un'indagine condotta presso alcune classi prime dell' I.T.C.S. "G. Salvemini" di Casalecchio di Reno (BO), nel febbraio 2003.

Prospetto di spoglio: carattere "Età"

Età in anni compiuti	Spoglio	N. studenti
14 anni	⧄ ⧄ ⧄ ⧄ ⧄ ⧄ ⧄ ⧄ ⧄ ⧄ ⧄ ⧄ |	61
15 anni	⧄ ⧄ ⧄ ⧄ ⧄ |	26
16 anni	⊔	3
Totale		90

Tabella 1

Prospetto di spoglio dei caratteri: "Sesso" e "Puntualità"

	Puntualità		
Sesso	Si	No	Tot
M	⧄ ⧄ ⧄ ⧄ ⧄ ⧄ ⧄ ⧄ ⧄ ⧄ ⧄	⧄ |	61
F	⧄ ⧄ ⧄ ⧄ ⧄ □		29
Totale	84	6	90

Tabella 2

E' fondamentale nel presentare i dati in forma tabellare inserire il titolo, indicando nell'ordine due elementi "chiave" che danno senso all'analisi: la definizione del collettivo statistico in esame e l'indicazione del carattere secondo il quale si sta effettuando lo studio del collettivo stesso. L'esigenza che entrambe le indicazioni siano esatte e sintetiche può rendere difficile sia la predisposizione del titolo della tabella che si sta costruendo, sia la comprensione di dati acquisiti già in forma tabellare.
L'operazione di spoglio effettuata produce rispettivamente la Tabella 3 e la Tabella 5a. La prima contiene una distribuzione statistica secondo un solo carattere (distribuzione semplice), la seconda una distribuzione statistica contemporaneamente secondo due caratteri (distribuzione doppia).

Quarta fase
L'insegnante propone di esaminare in dettaglio le risposte ad alcune domande.
Rispetto all'età:

Studenti della classe prima dell'I.T.C.S. "G. Salvemini"
per età (in anni compiuti)
(7 febbraio 2003)

Età (in anni compiuti)	N° studenti
14	61
15	26
16	3
Totale	90

Tabella 3

Qual età è la più frequente? *(moda = 14 anni)* Qual è l'età minima? Qual è l'età massima?
Se ci si vuole confrontare con la prima classe di un liceo scientifico vicino, è possibile utilizzare la tabella ottenuta? La risposta dipende dal fatto se le due scuole hanno lo stesso numero di studenti iscritti alla classe prima, se così non fosse occorrerà ricorrere al calcolo delle frequenze relative, o a quelle percentuali, di entrambe le tabelle, che diventeranno in tal modo commensurabili.

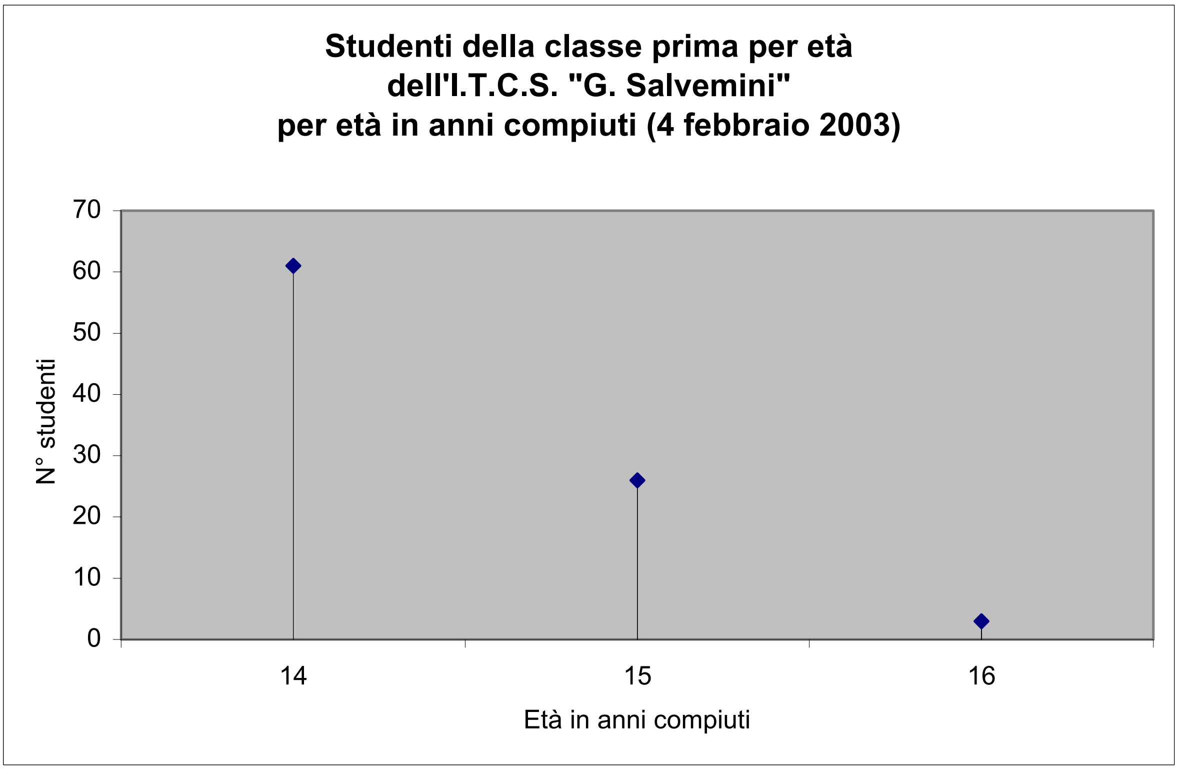


Figura 1

Dovendo costruire la rappresentazione grafica quale tipo di grafico è più opportuno? L'insegnante apre la discussione cercando di condurre gli studenti ad osservare che, essendo il carattere quantitativo discreto, la coppia ordinata (modalità, frequenza) individua un punto del piano cartesiano, con ciò consentendo la costruzione di un diagramma ad aste.
Ogni asta è proporzionale alla frequenza della corrispondente modalità del carattere.
Rispetto al tempo abituale di percorrenza casa scuola:

Studenti della classe prima dell'I.T.C.S. "G. Salvemini"
per tempo abituale di percorrenza casa-scuola
(7 febbraio 2003)

Tempo di percorrenza (min.)	N° studenti
fino a 15 min.	30
da 16 a 30 min.	24
da 31 a 45 min.	30
da 46 a 60 min.	5
oltre i 60 min.	1
Totale	90

Tabella 4

Nella tabella che classifica gli studenti rispetto ai tempi di percorrenza si può effettuare una rappresentazione grafica ad aste? Si? No? Perché?
Si ritrovano le stesse condizioni della tabella precedente? In effetti ora si ha a disposizione la coppia ordinata (intervallo, frequenza), non è quindi più possibile costruire un grafico ad aste.
Occorre fare una nuova ipotesi di lavoro: la frequenza viene rappresentata mediante un rettangolo in cui la base rappresenta l'intervallo e l'altezza la densità di frequenza (tale concetto potrà essere opportunamente approfondito ed utilizzato in seguito).

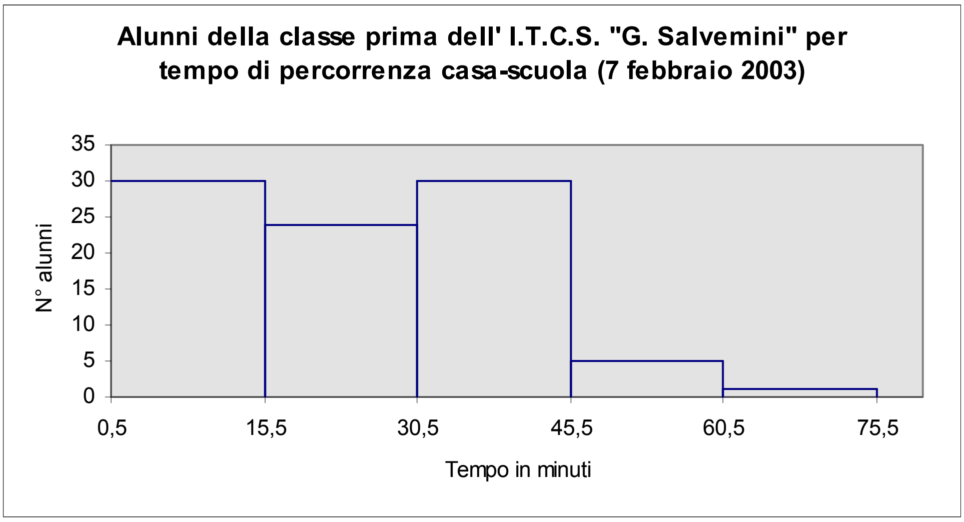


Figura 2

Inoltre occorre procedere ad una modifica degli estremi di ogni intervallo; ciò è possibile considerando l'approssimazione delle misure sicché gli intervalli modificati sono: 0,5 - 15,5; 15,5 – 30,5; 30,5 – 45,5; 45,5 – 60,5; 60,5 – 75,5. Si è in questo modo risolto anche il problema del primo e dell'ultimo intervallo che in origine erano aperti e che sono stati chiusi rispettivamente a sinistra con 0,5 e a destra con 75,5. Gli estremi di sinistra e di destra sono stati scelti in modo da mantenere invariata l'ampiezza degli intervalliA questo punto l'insegnante, dopo aver osservato che le aree di due rettangoli di uguale base sono proporzionali alle rispettive altezze, porta gli studenti a costruire **l'istogramma** di frequenze di Figura 2.

Contemporaneamente rispetto al sesso e alla puntualità:

Studenti della classe prima
dell'I.T.C.S. "G. Salvemini"
per sesso e puntualità
(7 febbraio 2003)

	Puntualità		
Sesso	Sì	No	
Femmine	55	6	61
Maschi	29	0	29
Totale	84	6	90

Tabella 5a

Studenti della classe prima
dell'I.T.C.S. "G. Salvemini"
per sesso e puntualità (dati in percentuale)
(7 febbraio 2003)

	Puntualità		
Sesso	Sì	No	
Femmine	90,2	9,8	100
Maschi	100	0	100
Totale	93,3	6,7	100

Tabella 5b

Lo scopo dell'indagine è volta anche a capire come gli studenti rispettano l'orario d'ingresso a scuola. Può essere che il sesso abbia influenza sul comportamento di questi studenti?
Guardando la tabella c'è differenza fra il comportamento dei maschi e quello delle femmine? Può essere utile calcolare le frequenze percentuali di riga?
I dati nel loro complesso possono meravigliare per l'alta percentuale di studenti che dichiara di arrivare abitualmente in orario. Gli studenti hanno "barato"? Ci sono stati errori di rilevazione o di spoglio? C'è qualche motivo per questo comportamento"virtuoso"? Un docente della scuola richiesto di giustificare questi dati informa che la scuola effettua un rigido controllo sui ritardi.

Possibili sviluppi
In modo analogo si possono costruire le distribuzioni degli altri caratteri facendo attenzione a che le rappresentazioni grafiche siano adeguate per ogni tipo di carattere. Grafici a settori circolari grafici, a colonne e grafici a nastro sono adatti per caratteri qualitativi; istogrammi a basi uguali sono adatti per caratteri qualitativi ordinati. Sarà poi cura dell'insegnante condurre gli studenti ad interpretare i dati elaborati.

Elementi di prove di verifica

Individua la risposta corretta fra quelle proposte, barrando la relativa lettera.

1. Qual è la definizione corretta di **carattere statistico**?
a) Un aspetto di una determinata entità, che può essere misurato o classificato
b) Un gruppo di unità omogenee per qualità
c) La misura dell'altezza di un individuo espressa in centimetri
d) La manifestazione di un giudizio su una unità statistica

2. Quale tra le seguenti è la definizione corretta di **unità statistica**?
a) Sono le entità che formano un collettivo
b) Il sesso degli individui che formano un collettivo
c) Un collettivo di unità statistiche a cui rivolgere l'indagine
d) Una popolazione che può essere classificata o misurata sotto vari aspetti

3. Fra i seguenti distingui quali sono **caratteri statistici**:

a) Sesso di un individuo	Sì	No
b) Reddito di una persona	Sì	No
c) Campione	Sì	No
d) Famiglia	Sì	No
e) Azienda	Sì	No
f) Numero di dipendenti di una azienda	Sì	No
g) Valutazione di un esame	Sì	No
h) Perimetro toracico	Sì	No
i) Colore dei capelli	Sì	No
l) Regioni di uno Stato	Sì	No
m) L'insieme delle città italiane	Sì	No

4. Tra quelli che hai individuato come caratteri, indica quali sono **qualitativi** e quali **quantitativi**

Caratteri qualitativi	Caratteri quantitativi

5. Per ciascuno fra i seguenti caratteri statistici, indica di quale tipo si tratta, scegliendo fra: **qualitativo sconnesso**, **qualitativo ordinato**, **quantitativo discreto**, **quantitativo continuo.** Spiega la tua risposta.

a) Età

____

b) Numero degli elettori alle elezioni politiche

____

c) Sesso

____

d) Statura

____

e) Reddito annuo di una persona

____

f) Temperatura rilevata in una certa ora della giornata

____

g) Professione di un individuo

____

Griglia di correzione

1. a
2. d
3. a, b, f, g, h, i

Grafico … è bello

Livello scolare: 1° biennio

Abilità interessate	**Conoscenze**	**Nuclei coinvolti**	**Collegamenti esterni**
Passare dai dati grezzi alle distribuzioni statistiche di frequenze ed alle corrispondenti rappresentazioni grafiche.	Frequenze assolute, relative, percentuali. Principali rappresentazioni grafiche per le distribuzioni di frequenze. Serie storiche e loro rappresentazioni.	<u>Dati e previsioni</u> Spazio e figure Risolvere e porsi problemi Laboratorio di matematica	Lingua italiana, storia ed educazione civica

Contesto
Distribuzioni semplici, grafici.

Il contesto è di tipo matematico, in particolare riguarda l'ambito statistico (distribuzione semplice e rappresentazioni grafiche) con aspetti extramatematici in quanto si basa su un'attività interdisciplinare con la storia e l'educazione civica, nella quale gli studenti si trovano di fronte a problemi di rappresentazione di dati che riguardano il "Processo di industrializzazione nella seconda metà dell'Ottocento".

Descrizione dell'attività
In accordo con il gruppo interdisciplinare che tratta questo modulo, relativamente alla parte di matematica, l'insegnante fa precedere all'esame delle tabelle che riguardano l'argomento di storia, alcune riflessioni su una tabella riguardante l'andamento degli iscritti alla classe prima in un istituto tecnico. In particolare propone la Tabella 1.

Andamento degli iscritti alla classe prima in un istituto superiore

anno	Iscritti
1993	216
1994	171
1995	204
1996	260
1997	247
1998	273
1999	281
2000	307
2001	324

Tabella 1

Successivamente si presentano la Tabella 2 e la Tabella 3 che riguardano l'attività interdisciplinare.

Percentuale della popolazione agricola sulla popolazione attiva

Anni	Gran Bretagna	Francia	Stati tedeschi	U.S.A.
1850	22	64	65	65
1870	15	49	49	50
1910	6	42	18	33

Tabella 2

Scambi internazionali nel 1914

Tra paesi europei	40
Da paesi non europei a paesi europei	21,5
Da paesi europei a paesi non europei	15,5
Tra paesi non europei	23
Totale	100

Tabella 3

La lettura dei dati delle tre tabelle comporta difficoltà crescenti, particolarmente per quelli di Tabella 3. Perché i dati non sono facilmente leggibili? Cosa manca nella Tabella 3? L'insegnante chiede agli studenti l'individuazione di un modo efficace per la loro lettura ed interpretazione e li conduce a rendersi conto che un grafico, non meno della parola, di uno scritto, consente una comunicazione efficace dei risultati di un'indagine, di uno studio su un problema sociale, ecc. I grafici consentono infatti di comunicare, comprendere, analizzare, riassumere, confrontare e visualizzare subito tutto! Tuttavia occorre imparare ad usare e non ad abusare dei grafici. I prerequisiti minimi per la costruzione dei grafici sono: piano cartesiano, proporzioni, percentuali.

<u>Prima fase</u>
L'insegnante induce gli studenti ad osservare che i dati della Tabella 1 riguardano la serie storica delle iscrizioni nella classe iniziale di una scuola mentre quelli della Tabella 2 riguardano serie storiche riferite a più paesi. Si tratta di distribuzioni? Sì, no, perché?
La tabella 3 è o non è una distribuzione? Sì, no e perché? Ha un titolo completo?
I dati contenuti nelle tabelle sono dati direttamente osservati o derivano da un'elaborazione?
L'insegnante guida gli studenti a scrivere il rapporto che dà origine a ciascuno dei dati della Tabella 2 e confronta tale procedimento con quello che dà origine ai dati di Tabella 3.
L'insegnante illustra agli studenti le caratteristiche dei principali strumenti di rappresentazione grafica utilizzabili in questo contesto ed il ruolo che in essi ha l'asse orizzontale. Dipende dal tipo di modalità osservate? E' solo una base di appoggio o deve avere una scala? In particolare, analizza le caratteristiche del grafico a settori circolari e dei grafici a colonne.

<u>Seconda fase</u>
Ogni studente partendo dalle tabelle iniziali produce il grafico che ritiene più opportuno, utilizzando i mezzi che ha a propria disposizione (carta, compasso e matita, calcolatore).

<u>Terza fase</u>
L'insegnante discute con gli studenti i grafici che hanno elaborato. Nel prosieguo dell'unità vengono riportati i lavori prodotti da studenti di una prima classe di un istituto superiore. In particolare, disponendo di un Foglio elettronico Excel, vi sono stati studenti che hanno scelto grafici, alcuni dei quali sono particolarmente suggestivi, anche se non privi di errori.

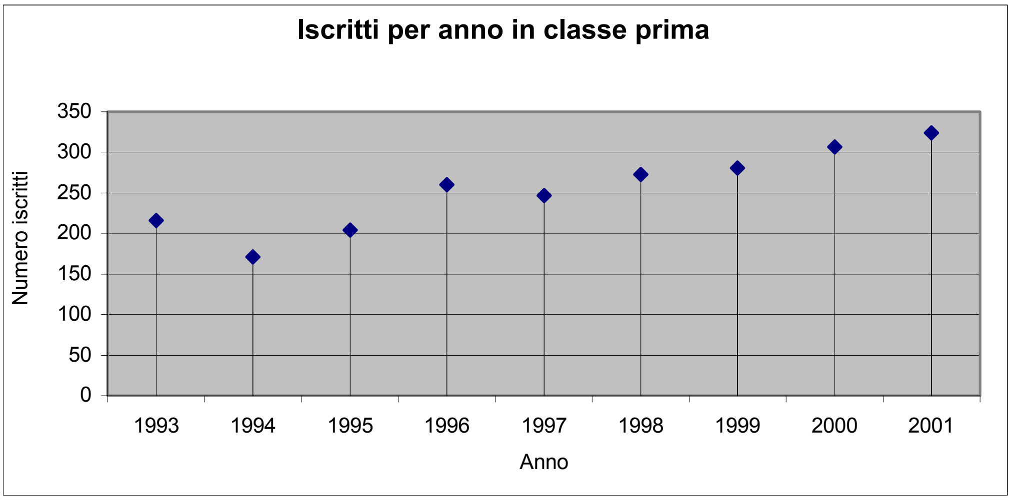


Figura 1

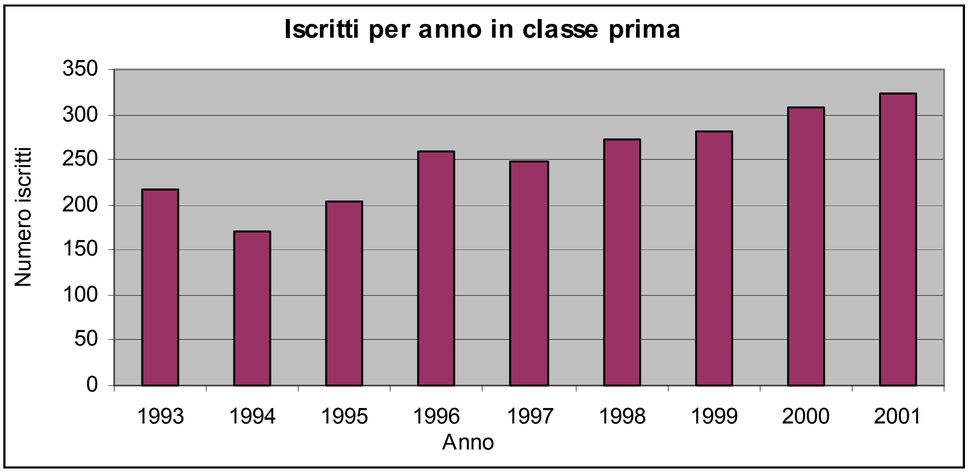


Figura 2

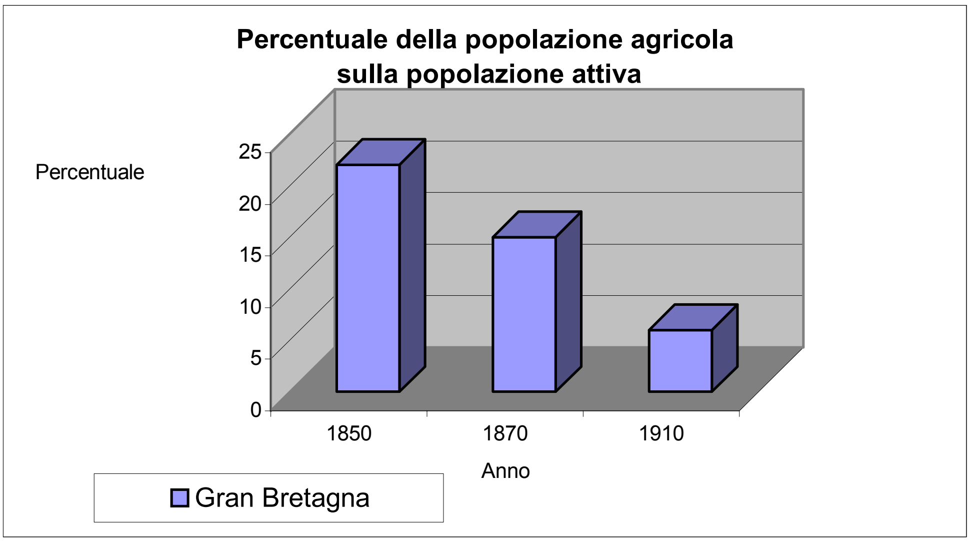


Figura 3

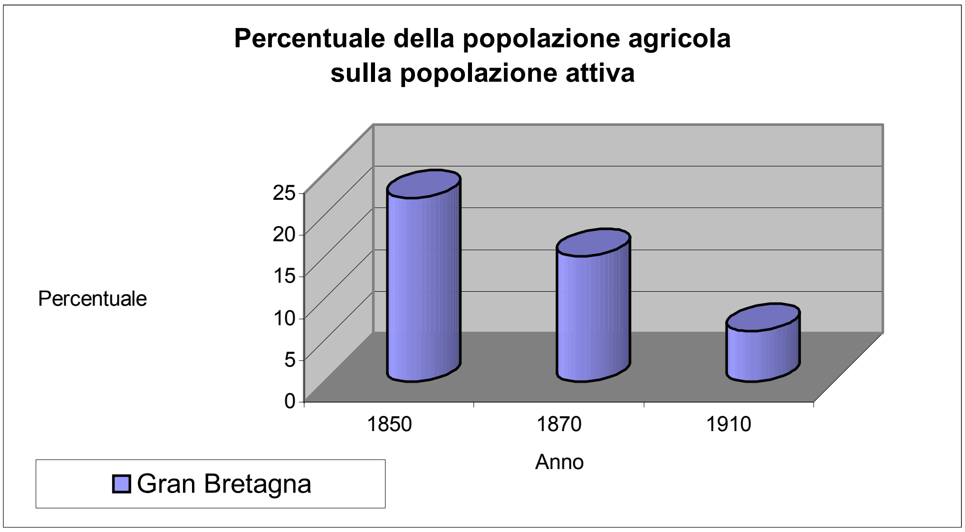


Figura 4

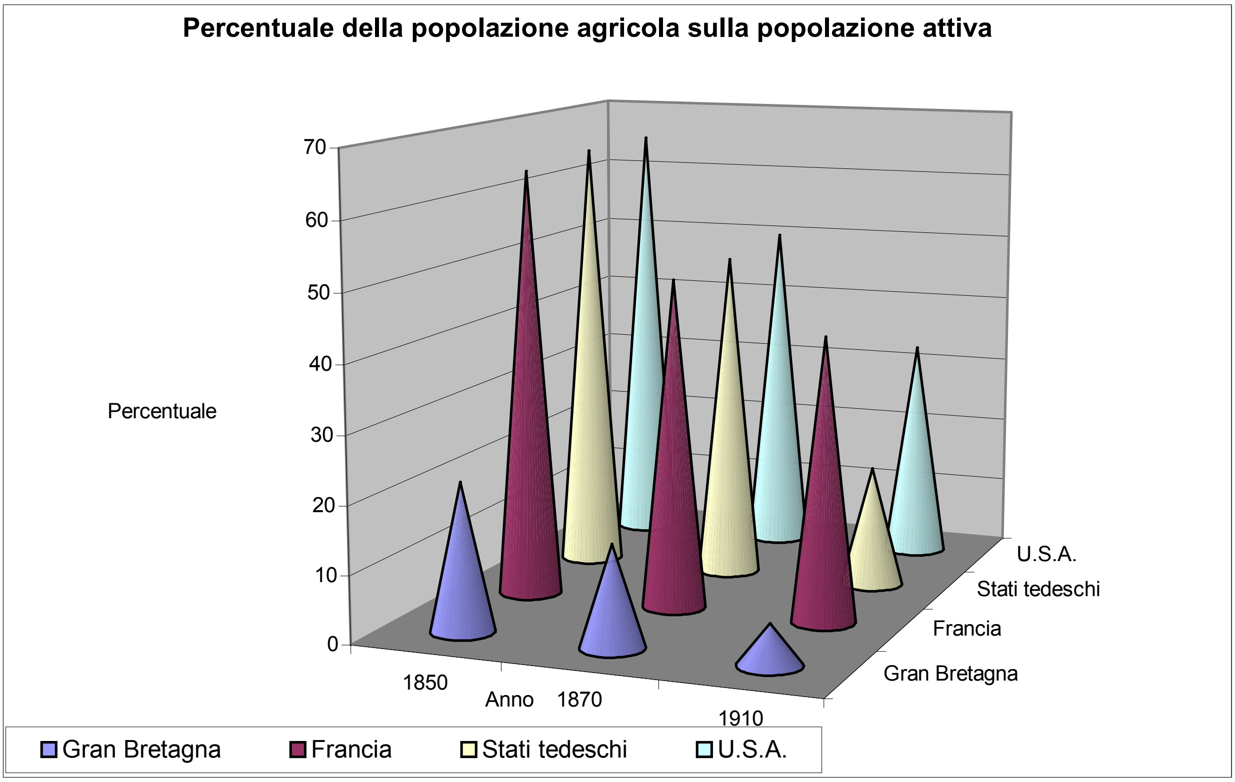


Figura 5

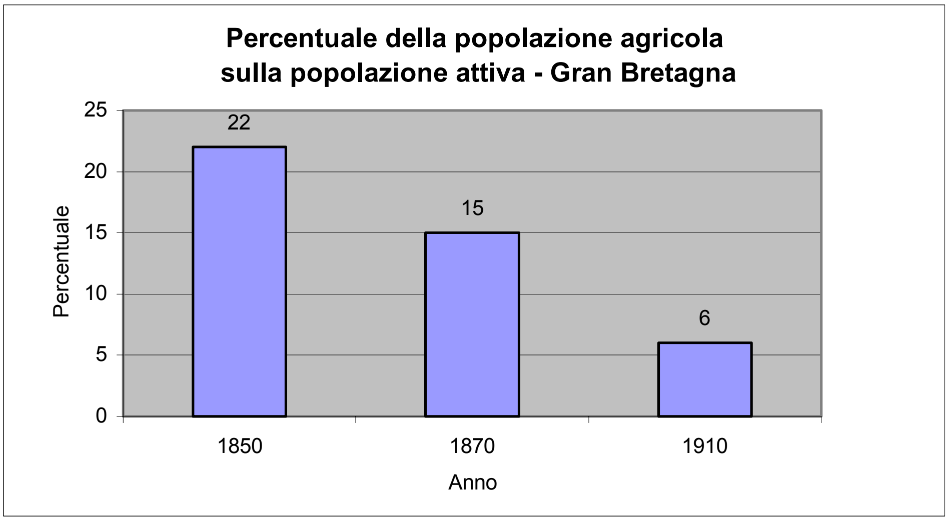


Figura 6a

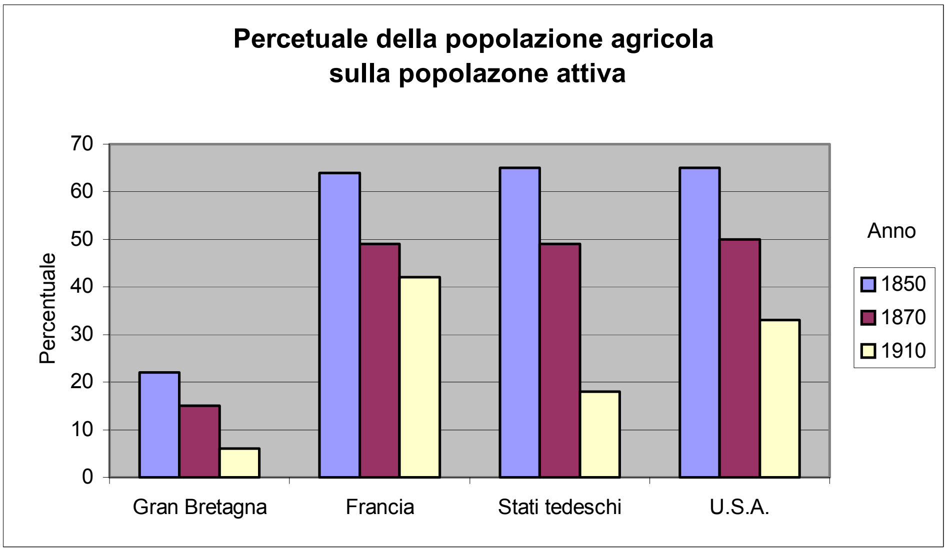


Figura 6b

Per la rappresentazione grafica dei dati della Tabella 1 sono stati usati due grafici che esprimono in maniera corretta, (perché?) l'informazione che la tabella offre. L'insegnante fa emergere con la discussione che un buon grafico è quello che riesce a combinare l'esigenza della completezza e della correttezza con la facilità e l'immediatezza della lettura.
Per la Tabella 2 sono stati scelti, in generale, grafici che introducono una terza dimensione non esistente nelle tabelle, producendo prismi (Figura 3) o cilindri (Figura 4) nei quali viene affidato all'altezza il ruolo di rappresentare la percentuale. Si è anche prodotto un grafico molto particolare che ha utilizzato dei coni in uno spazio tridimensionale dove vengono rappresentati contemporaneamente gli anni, le percentuali, i paesi (Figura 5). La lettura di questo grafico risulta indubbiamente meno facile rispetto ai precedenti, anche se l'immagine che si ricava è sufficientemente fedele rispetto ai dati da rappresentare.
L'insegnante chiede agli studenti se, rispetto ai tre periodi considerati, gli intervalli temporali sono costanti. Ciò induce gli studenti a riflettere che tutti i grafici presentano un errore al riguardo perché non si è tenuto conto che gli intervalli fra un anno e il successivo sono diversi.

Dovendo convergere verso una sola rappresentazione si scelgono le Figure 6a e 6b, apportando, però, la correzione necessaria per tener conto dell'osservazione precedente. Dopo l'opportuna correzione, il grafico della Figura 6a dà luogo al grafico di Figura 7, che consente di illustrare correttamente i dati di un singolo paese. In maniera analoga l'insegnante conduce gli studenti a modificare il grafico di Figura 6b per effettuare in modo corretto il confronto simultaneo fra paesi diversi.

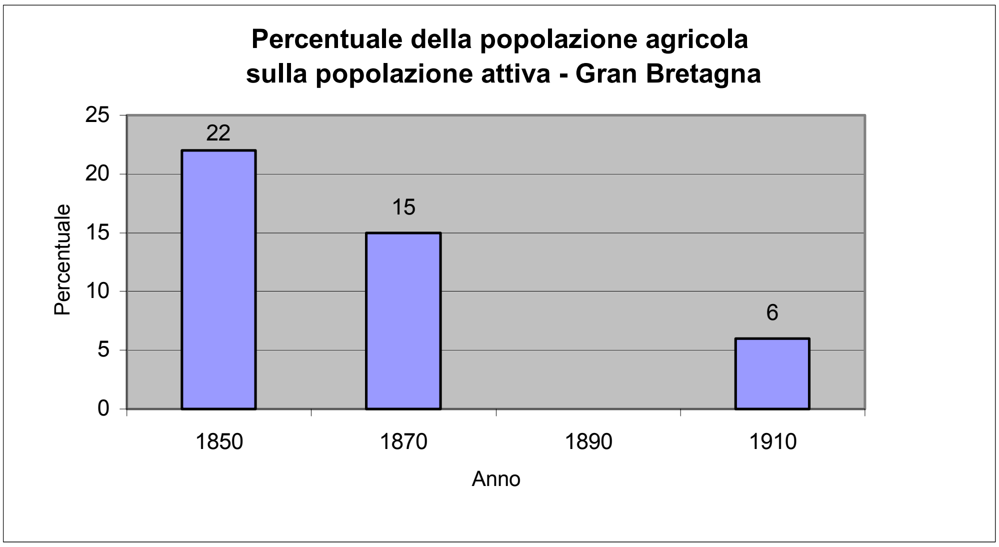


Figura 7

Rispetto alla distribuzione degli scambi internazionali nel 1914 (Tabella 3), i grafici delle Figure 8a e 8b risultano corretti. In particolare nel grafico a settori circolari si è evitato di introdurre una dimensione inesistente. L'insegnante guida gli studenti al confronto tra due grafici. Quale dei due consente la rappresentazione del totale? E' corretto usare il grafico a settori per i dati della prima colonna della Tabella 2? Perché?

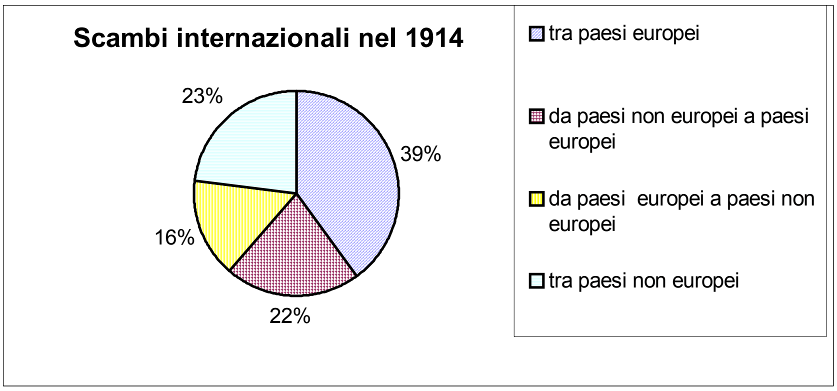


Figura 8a

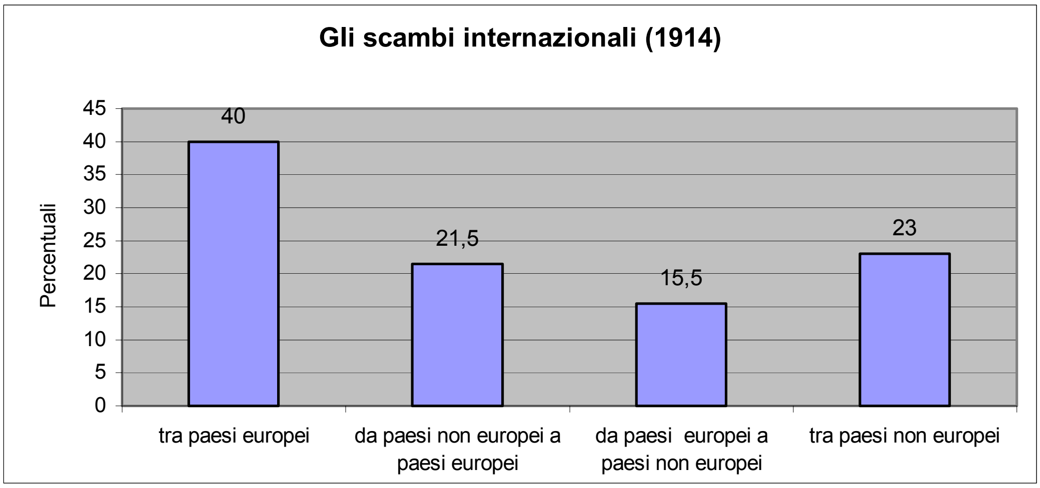


Figura 8b

In questa fase di studio l'insegnante indica i nomi tecnici dei grafici, fa osservare come sono stati costruiti e, soprattutto, che essi richiedono, per essere leggibili, la presenza di alcuni elementi fondamentali: titolo, legenda, indicazione delle unità di misura.
In seguito mostra uno schema simile al seguente, dove, per ciascuna tipologia di dati, viene suggerito il tipo di grafico più appropriato. L'insegnante termina l'attività, invitando gli studenti a completare l'ultima colonna con un esempio di distribuzione.

Tipologia di dati	Tipo di grafico	Esempio di distribuzione
Una distribuzione secondo un carattere *qualitativo* (con poche modalità)	Il diagramma a settori circolari (detto diagramma a torta)	
Una distribuzione secondo un carattere *qualitativo* (con molte modalità)	Il diagramma a colonne o nastri	
Una distribuzione secondo un carattere *quantitativo discreto*	Il diagramma cartesiano per punti o con diagramma ad aste	
Una serie storica	Il diagramma cartesiano o a nastri	
Una distribuzione secondo un carattere *quantitativo in classi* (prestando attenzione alla loro ampiezza)	L'istogramma di frequenza	
Una serie territoriale	Il cartogramma	

Di media *non* ce n'è una sola!

Livello scolare: 1° biennio

Abilità interessate	Conoscenze	Nuclei coinvolti	Collegamenti esterni
Calcolare i principali valori medi per caratteri quantitativi.	Proprietà dei principali valori medi.	Dati e previsioni Numeri e algoritmi Argomentare, congetturare, dimostrare Misurare Risolvere e porsi problemi	Viaggi Commercio Vita sociale

Contesto.
Vita quotidiana.

Il contesto è di tipo matematico, in particolare riguarda l'ambito statistico e l'ambito probabilistico. Partendo da situazioni problematiche legate all'esperienza degli studenti (vedi anche l'attività "Arrivare a scuola") si possono introdurre alcune attività che motivino l'introduzione di diversi valori medi. L'unità è rivolta a studenti del primo biennio con l'obiettivo di "fugare" la credenza, ampiamente diffusa nell'opinione pubblica, che esista soltanto la media aritmetica (la cosiddetta "media matematica"!) utile per giustificare o risolvere qualunque problema in cui occorra individuare un indice sintetico.

Descrizione dell'attività
Prima fase
L'insegnante presenta le seguenti situazioni problematiche:

a) Un aereo viaggia da Roma a New York. All'andata le correnti favorevoli permettono all'aereo di viaggiare alla velocità di crociera di 932 Km/h; al ritorno la velocità è, invece, di 856 Km/h. Qual è la velocità media dell'aereo nell'intero percorso andata-ritorno?

b) Una agenzia che effettua indagini di mercato ha rilevato per una rete televisiva i seguenti dati medi giornalieri di ascolto, nel periodo invernale e nella fascia oraria dalle 20 alle 21:

Giorno della settimana	Lunedì	Martedì	Mercoledì	Giovedì	Venerdì	Sabato	Domenica
Numero medio spettatori (in migliaia)	1.200	1.800	2.000	1.600	1.200	800	900

Una agenzia di pubblicità in quale giorno potrebbe consigliare a un proprio cliente di inserire uno spot pubblicitario di un prodotto per la neve, volendo usare la fascia oraria 20 – 21?

c) In un ciclo di lavorazione tre apparecchiature lavorano in serie: la prima macchina ha un rendimento del 90 % , la seconda dell'80% , la terza del 30%. Qual è il rendimento medio complessivo?

d) Uno studente nella pagella del primo quadrimestre ha riportato i seguenti voti:

Italiano	7
Storia	8
Geografia	7
Lingua inglese	6
Scienze	5
Matematica	4
Educazione Fisica	9

Il padre gli ha promesso un regalo se la media dei suoi voti fosse stata superiore al 7.
Otterrà lo studente il regalo?

e) Uno studente universitario iscritto al corso di laurea in Matematica ha superato durante il primo anno i seguenti esami[1] riportando le seguenti votazioni:

Esame	Punteggio in trentesimi	Crediti
Laboratorio di Matematica	25	9
Analisi Matematica	24	12
Geometria	21	6
Algebra	27	6
Calcolo delle probabilità	23	9
Fisica generale	24	9
Lingua inglese	30	3
Fondamenti di Informatica	28	3
Abilità relazionali	30	3

Lo studente accede ad una borsa di studio se ha conseguito una media superiore a 27/30.
Otterrà il nostro studente la borsa di studio?

f) In una prova multidisciplinare di Storia, Inglese, Matematica, Diritto, gli studenti vengono valutati con un punteggio da 0 a 15 per ogni materia. Il voto finale è dato dalla media dei quattro punteggi parziali. La prova non si considera superata se uno studente prende 0 punti in una delle materie. Quale valore medio consente di rappresentare adeguatamente questo modo di valutare?

<u>Seconda fase</u>
Si invitano gli studenti, preferibilmente in attività di gruppo, a riflettere sul modo di risolvere i problemi proposti.
Gli errori ricorrenti potrebbero essere, relativamente ad ogni esercizio, i seguenti:

a) la prima idea risolutiva per tale problema è frequentemente quella di utilizzare la media aritmetica, ovvero di calcolare

$$v_m = \frac{932 + 856}{2} = 894 Km/h$$

In tale situazione non è sempre semplice riuscire a spiegare allo studente dove è l'errore commesso. Può essere opportuno "esasperare" il problema assegnando valore 0 alla velocità

[1] Secondo il nuovo ordinamento universitario ad ogni esame è associato un numero di crediti: ciascun credito corrisponde a circa 25 ore di lezione-tutoraggio-impegno individuale dello studente. Ogni anno lo studente è tenuto ad accumulare 60 crediti.

media del ritorno, il che significa ammettere implicitamente di non ritornare; in tal caso appare evidente che la media aritmetica, conservando i totali dei valori da sintetizzare, non risolve il problema. L'insegnante condurrà allora gli studenti ponendo loro la domanda: nell'andata quanti Km si sono percorsi? E quanti nel ritorno? Come si può allora esprimere la velocità del "sistema" andata ritorno? E' ragionevole ammettere che il numero dei Km percorsi nell'andata e nel ritorno sia identico? Se sì, qual è la formulazione teorica che risolve il problema? Se no, il problema ha soluzione? L'insegnante può utilmente avvalersi del calcolo letterale, guidando gli studenti alla soluzione corretta che consiste nell'utilizzo della media **armonica.** Tale media può essere formalizzata in collegamento con il nucleo Numeri e algoritmi. Si avrà pertanto, ammettendo uguali i Km percorsi nell'andata e nel ritorno:

$$v_m = \frac{2}{\frac{1}{v_A} + \frac{1}{v_B}}$$

b) Una lettura non sufficientemente attenta della distribuzione statistica potrebbe indurre gli studenti a calcolare la media aritmetica delle frequenze. L'insegnante ricorda il significato rispettivamente delle modalità e delle frequenze e conduce gli studenti ad individuare, come risposta al problema proposto, la moda. Essa in questo caso è il mercoledì, il giorno della settimana con la massima frequenza di ascolto.

c) A questo punto del percorso, gli studenti dovrebbero avere chiaro che la media aritmetica risponde al problema della equidistribuzione di un totale che viene conservato.
L'insegnante invita gli studenti a cogliere il senso del problema. La prima macchina in origine quanto deve rendere? Se la risposta è 100 pezzi, effettuata la lavorazione, quanto ha reso? Se si danno 90 pezzi alla seconda macchina quanti pezzi si ottengono? Se si immettono 72 pezzi nella terza ed ultima macchina quanti pezzi si hanno? Dunque da 100 pezzi teorici si ottengono in realtà, con l'intervento in serie delle 3 macchine, 21,6 pezzi. Il problema è allora di trovare tre macchine tutte con la stessa resa che diano alla fine dei tre passaggi 21,6 pezzi. Il percorso logico può essere coadiuvato da una semplice schematizzazione formale che porta a concludere che quando è necessario mantenere il risultato di un prodotto la media idonea è quella **geometrica.**

d) Poiché in questo caso va rispettata la somma dei punteggi ottenuti, la sintesi corretta è la media aritmetica. E' da far notare un errore che gli studenti potrebbero commettere: calcolare le medie parziali dell'area umanistica (Italiano, Storia, Geografia, Lingua Inglese), scientifica (Scienze, Matematica) e di Educazione Fisica e poi fare la media semplice delle medie.
Riflettere su ciò può essere didatticamente efficace perché dà l'opportunità di offrire un esempio di operazione non associativa. Infatti nel fare una media di medie parziali occorre tener conto di quante sono le unità statistiche che ciascuna media parziale sintetizza (collegamento con Numeri e algoritmi).

e) Questo problema non dovrebbe creare eccessive difficoltà per convincere gli studenti ad utilizzare la media aritmetica ponderata. Nel caso in cui qualcuno commettesse l'errore di trascurare il credito si potrebbe proporre al solito un esempio "estremo": una valutazione di 30/30 in un esame con 3 crediti e una valutazione di 30/30 in un esame con 10 crediti può contare allo stesso modo? Evidentemente no!
Una situazione simile potrebbe essere il calcolo della media **aritmetica** con valori ripetuti (per esempio i voti conseguiti da una classe di studenti in un compito di matematica).

f) Quale dei valori medi utilizzati nei casi precedenti potrebbe andar bene in questo esempio?

Dalla discussione dovrebbe emergere che ciò che si vuole è che, quando si presenta uno zero, la sintesi si annulli. Dunque occorre utilizzare un valore medio fondato sul prodotto. E' forse la media geometrica la soluzione del problema posto? Essa in effetti è l'unica media che, assumendo in tal caso il valore 0, consente di escludere uno studente che abbia riportato un punteggio nullo in almeno una delle quattro materie.

Terza fase
L'insegnante a questo punto sistematizza le conoscenze sui diversi tipi di medie, fornendo in ogni caso la formula e chiarendo le relative proprietà.

Quarta fase
Verifica.
Si allegano elementi di prove di verifica sui valori medi.

Elementi di prove di verifica

1. Il "colore dei capelli" viene osservato su tre individui, ottenendo:

A	B	C
Capelli BIONDI	Capelli ROSSI	Capelli NERI

La mediana di questa distribuzione è:
a) Capelli BIONDI
b) Capelli ROSSI
c) E' B
d) Non si può calcolare
e) Capelli NERI

La moda di questa distribuzione è:
a) E' A
b) Capelli NERI
c) Non esiste
d) Capelli ROSSI
e) Capelli BIONDI

2. In una prova nove studenti vengono valutati, assegnando loro uno dei tre livelli: O = ottimo, B = Buono, S = Sufficiente, ottenendo:

Alberto	Gino	Raffaella	Maria	Anna	Mario	Giuseppe	Carla	Roberto
O	O	B	S	S	O	B	B	S

La mediana della distribuzione è:
a) Anna
b) O
c) 4,5
d) 5
e) B

3. Il reddito medio mensile di cinque famiglie italiane nel mese di giugno nel 2002 è stato di € 1.705. Il reddito complessivo di queste famiglie è di:

a) circa € 10.000
b) minore di € 7.000
c) non si può calcolare
d) € 8.525
e) maggiore di € 10.000

4. Ad una certa data, l'età media in anni compiuti dei componenti di una famiglia di quattro persone è pari ad anni 32. Se tre dei componenti hanno rispettivamente 15, 50 e 47 anni, l'età del quarto componente è:

a) 16 anni
b) 11 anni
c) Minore di 10 anni
d) Non si può calcolare
e) 18 anni

5. I voti in matematica di 8 studenti di una scuola secondaria sono:

Studente	1	2	3	4	5	6	7	8
Voto	8	3	6	8	7	8	4	7

Per la metà degli studenti più bravi il voto minimo è stato almeno:

a) 7,1
b) 7
c) 5
d) 8
e) 6

Griglia di correzione

1. d, c
2. e
3. d
4. a
5. b

Pivot è bello

Livello scolare: 1° biennio

Abilità interessate	**Conoscenze**	**Nuclei coinvolti**	**Collegamenti esterni**
Predisporre la struttura della matrice dei dati grezzi con riguardo a una rilevazione pianificata e inserire i dati rilevati anche in un foglio elettronico. Passare dai dati grezzi alle distribuzioni statistiche di frequenze ed alle corrispondenti rappresentazioni grafiche. Calcolare i principali valori medi e le misure di variabilità.	Distribuzioni delle frequenze a seconda del tipo di carattere. Frequenze assolute, relative, percentuali. Principali rappresentazioni grafiche per le distribuzioni di frequenze. Proprietà dei principali valori medi e delle misure di variabilità.	<u>Dati e previsioni</u> <u>Laboratorio di matematica</u> Argomentare, congetturare, dimostrare Risolvere e porsi problemi	Vita sociale Organizzazione di attività

Contesto
Numeri, grafici.

Il contesto è di tipo matematico, in particolare riguarda l'uso di dati numerici e delle loro rappresentazioni grafiche.
Questa attività può essere introdotta, nella forma che qui viene proposta, in una classe del primo biennio quando gli studenti hanno acquisito abilità di base sull'utilizzo del foglio elettronico (selezionare, copiare, incollare, indirizzamenti assoluti e relativi, creazione guidata dei grafici, …).
Dopo che gli studenti hanno imparato ad effettuare lo spoglio di tipo manuale di un numero limitato di semplici questionari, l'attività proposta diventa molto utile quando lo spoglio riguarda molti questionari. Trasferire i dati in una tabella di un foglio elettronico rende comodo e agevole la manipolazione flessibile di distribuzioni univariate tratte dall'intera base di dati o da un suo sottoinsieme.

Descrizione dell'attività
Per lo sviluppo di questa attività si è utilizzato il foglio elettronico Excel.
Nell'esempio si presenta una parte di una tabella, tratta dall'unità "Arrivare a scuola" (Tabella 1) e la si descrive nelle sue parti principali al fine di chiarire i diversi concetti in essa contenuti.
Ogni riga rappresenta una unità statistica e ogni informazione da essa ricavata è disposta nelle rispettive colonne, una per la modalità di ciascun carattere.
Perché nelle colonne C e D, che contengono caratteri qualitativi sconnessi, si trovano dei numeri?
Perché proprio quei numeri?
Per poter utilizzare in modo semplice lo strumento delle tabelle pivot è importante che nelle singole celle ci siano dei valori numerici. Pertanto anche le modalità qualitative sono state codificate in forma numerica. In questo caso la scelta è stata la seguente: carattere "Sesso": modalità "Femmina" $\rightarrow$ 1, modalità "Maschio" $\rightarrow$ 2. Anche per il carattere "Comune di residenza" è stata effettuata una scelta analoga.
Tenuto conto che il questionario è stato somministrato in una scuola superiore, come è possibile trovare le età riportate in tabella? Evidentemente è stata fatta, a livello di codifica, una scelta analoga alla precedente, anche se il carattere quantitativo non lo richiedeva espressamente. E' stato infatti attribuito codice 1 all'età "Meno di 14" che corrisponde alla prima risposta data alla

domanda 1 del questionario reperibile all'unità "Arrivare a scuola", codice 2 ad una età "14", e così via. In questo modo è stata legata direttamente la modalità al numero di risposta del questionario stesso.

	A	B	C	D	E
1	codice	età	sesso	comune	classe
2	1	7	2	3	4
3	2	6	1	4	4
4	3	5	1	16	4
5	4	5	1	4	4
6	5	6	1	2	4
7	6	5	2	2	4
8	7	5	1	2	4

Tabella 1

Nella colonna A sono riportati i codici attribuiti ai singoli questionari in fase di spoglio. Ciò permette di individuare ogni record e di effettuare il controllo della correttezza dell'inserimento dei dati di ciascun questionario

Nella colonna B sono riportati i codici assegnati alle diverse età, secondo quanto indicato sopra: 7 corrisponde ad uno studente la cui età è di 19 anni, che è la settima risposta alla domanda 1 del questionario.

<u>Prima fase</u>

Una prima proposta riguarda il modo di ottenere informazioni circa la distribuzione univariata delle frequenze del carattere "Età" del collettivo considerato.

Si utilizza la procedura guidata per "interrogare" l'insieme dei dati oggetto di interesse (database).

La prima cosa da fare è selezionare una cella qualsiasi del database. Successivamente si sceglie dal menu dati la voce "Report della tabella pivot e del grafico pivot …".

Si ottiene così la seguente figura:

Figura 1

La scelta del pulsante “Avanti” apre una finestra che indica la zona nella quale sono residenti i dati oggetto di interesse; dopo aver confermato nuovamente con il pulsante “Avanti”, si ottiene la seguente finestra:

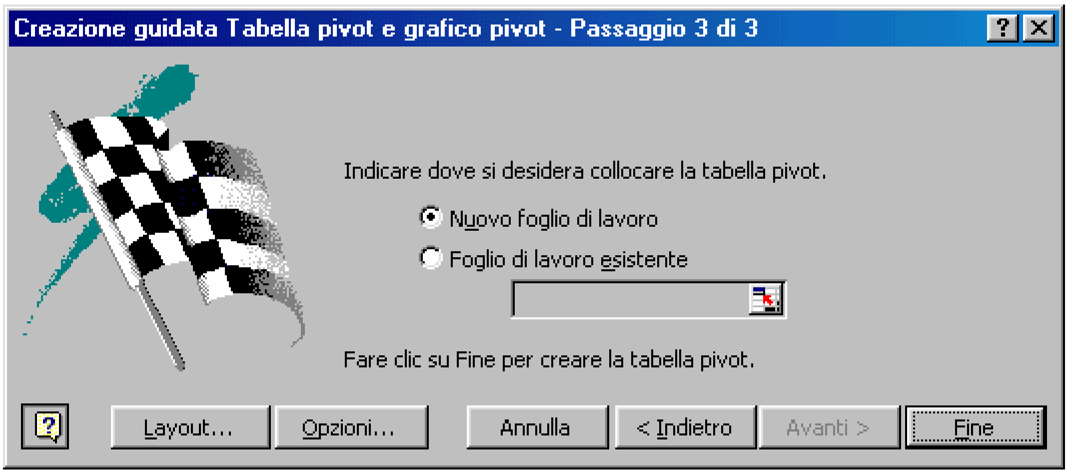


Figura 2

In questa finestra si deve scegliere dove posizionare i risultati dell’interrogazione (Tabella pivot) del database; è consigliabile, per maggior chiarezza, collocare la tabella pivot in un nuovo foglio di lavoro (come indicato in Figura 2).
Scegliendo il pulsante “Layout” si giunge alla finestra successiva che permette la costruzione della tabella pivot.

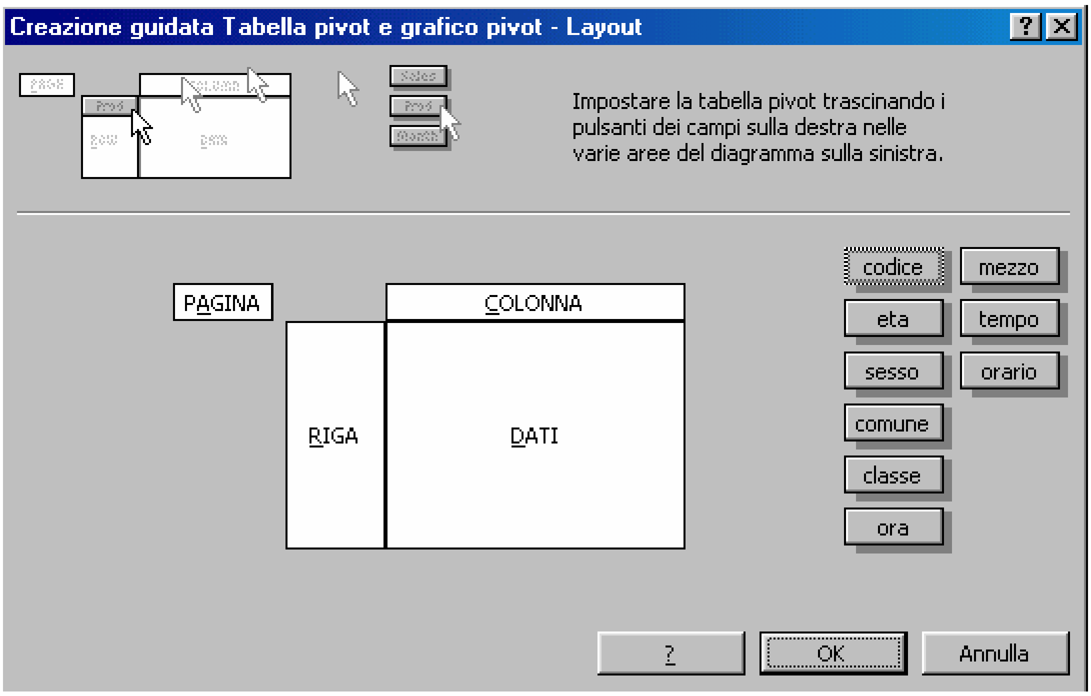


Figura 3

Nella finestra sono visibili i campi del database (i caratteri) e un riquadro dove, posizionando i singoli campi, è possibile ottenere la tabella desiderata. Per semplice trascinamento del campo sesso sulla zona contrassegnata dal nome Riga si inserisce il carattere “Sesso”. Trascinando il pulsante

sesso nella zona Dati si ottiene il conteggio delle unità che posseggono tali modalità. A questo riguardo è importante scegliere la modalità “Conta( )” effettuando un doppio clic sul pulsante attivo nella zona Dati dopo il trascinamento.
Per poter effettuare una interrogazione per sottoinsiemi del database, ad esempio per classe, si deve trascinare il pulsante del carattere discriminatore prescelto sul campo Pagina.
Il risultato delle scelte proposte è visibile nella figura seguente:

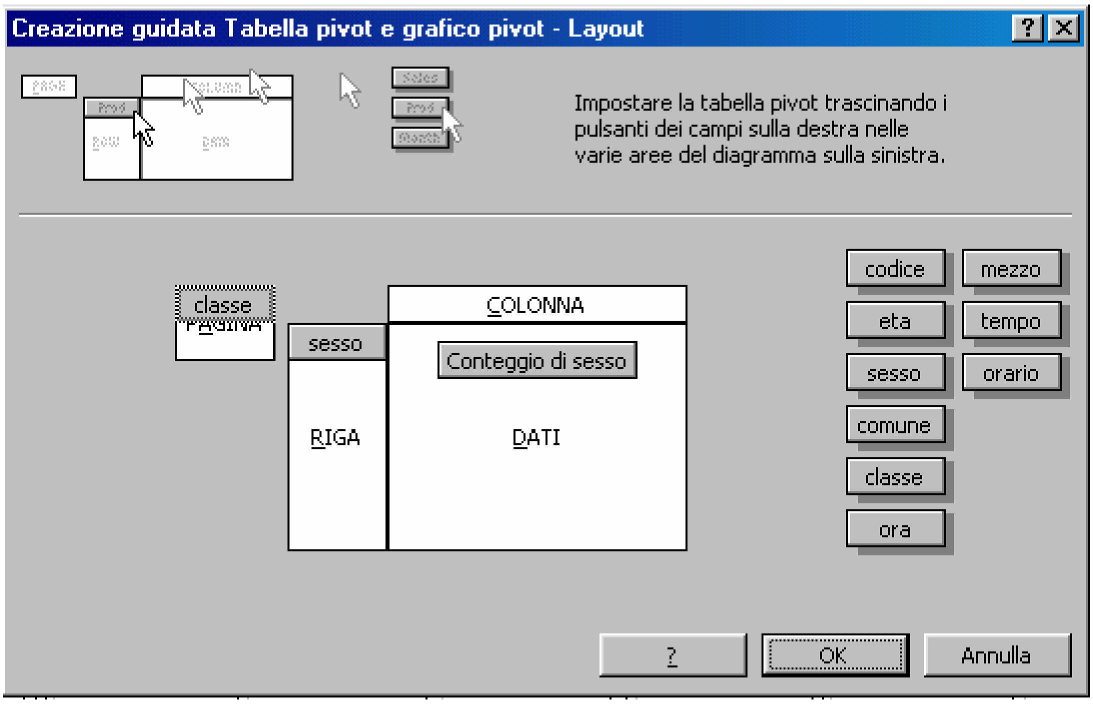


Figura 4

<u>Seconda fase</u>
Il risultato della procedura illustrata nella fase precedente permette di ottenere la tabella mostrata nella figura di seguito riportata:

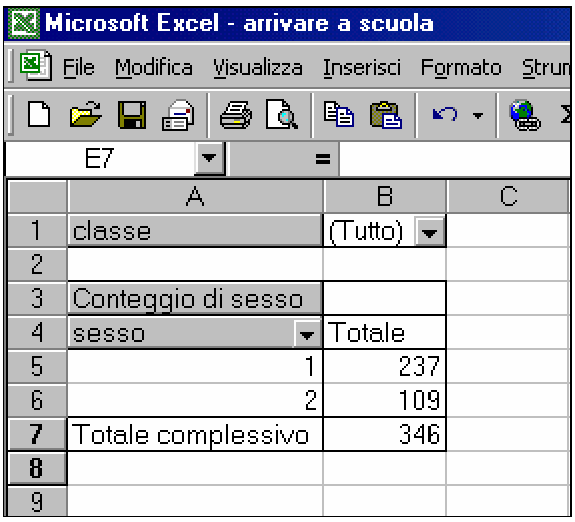


Figura 5

La tabella pivot così ottenuta, fornisce, rispetto a tutte le classi, le informazioni richieste.

Il pulsante contenuto nella cella B1 *(nome in Excel)* permette poi di ottenere lo stesso tipo di distribuzione riferita alla modalità prescelta tra quelle proposte dal carattere “Classe”.
Bisogna tuttavia osservare che per poter lavorare sui dati ottenuti è opportuno ricopiare il contenuto della tabella e incollare i valori in un altro foglio con la procedura: “Incolla speciale”, selezionando l’opzione “Valori”. A tal fine conviene effettuare la selezione partendo da una cella in basso a destra rispetto all’area di interesse.
In questo modo si ottiene una tabella indipendente da quella pivot e soprattutto dalle modifiche determinate su di essa da ulteriori interrogazioni.

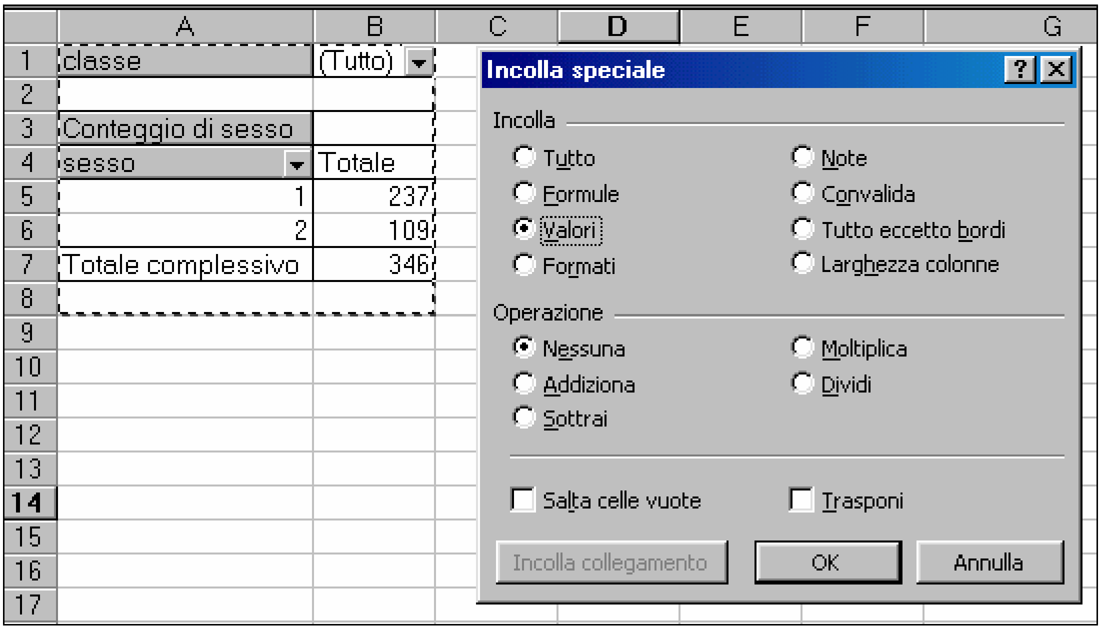


Figura 6

<u>Terza fase</u>
E’ possibile rappresentare graficamente le informazioni contenute nella Figura 5? Quale tipo di grafico rappresenta in maniera più appropriata la distribuzione di frequenze esaminata?
L’insegnante sollecita la discussione e guida l’attenzione degli studenti sul significato statistico delle coppie (“Modalità del carattere Sesso”, “Frequenza”) e sulla loro rappresentazione grafica.
Una soluzione possibile può essere ottenuta mediante i seguenti passaggi.
In primo luogo occorre decodificare le modalità del carattere sesso. E’ sufficiente sovrascrivere nelle celle corrispondenti la modalità “Femmina” e la modalità “Maschio”.
Con la selezione di una cella qualsiasi all’interno della tabella pivot si può utilizzare la procedura guidata per la composizione del grafico che è in grado di rappresentare in modo corretto la distribuzione in esame.
L’insegnante, per far emergere la scelta del grafico più opportuno per rappresentare la distribuzione di frequenze, guida gli studenti a riflettere sulla natura del supporto orizzontale sul quale si trovano le modalità del sesso e sulla scelta di un grafico adeguato tra quelli proposti dal foglio elettronico per rappresentare la distribuzione considerata.

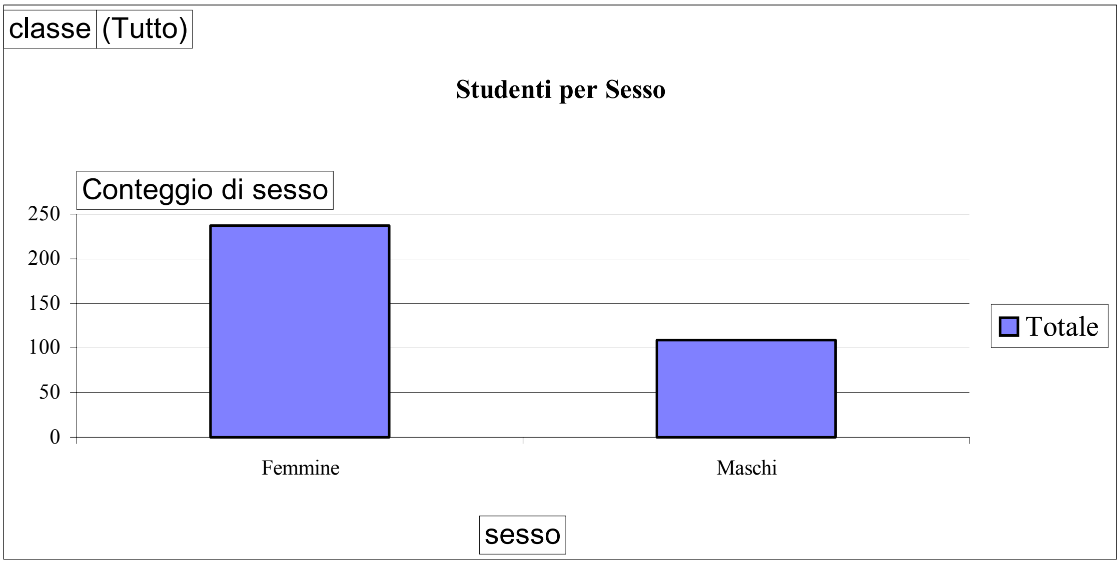


Figura 7

Che tipo di grafico propone in modo automatico il foglio elettronico? E' corretto? Perché?

<u>Quarta fase</u>

Sono di più i maschi in prima o in seconda? La tabella pivot può aiutare a rispondere a questa domanda? Si può rispondere alla prima domanda con un grafico ad aste?

L'insegnante guida gli studenti ad utilizzare la tabella pivot per trovare la risposta mediante l'interrogazione secondo la voce classe. I risultati sono riportati nelle tabelle seguenti:

classe	1
sesso	
Femmine	61
Maschi	29
Totale	90

Tabella 2

classe	2
sesso	
Femmine	36
Maschi	27
Totale	63

Tabella 3

Spontaneamente ci si attende che gli studenti rispondano che sono di più i maschi in prima. E' compito dell'insegnante guidarli al confronto fra due collettivi di diversa numerosità, pervenendo al concetto di frequenza relativa e al calcolo della Tabella 4 e della Tabella 5.

classe	1
sesso	
Femmine	67,78
Maschi	32,22
Totale	100,00

Tabella 4

classe	2
sesso	
Femmine	57,14
Maschi	42,86
Totale	100,00

Tabella 5

<u>Quinta fase</u>

Ci sono strumenti che consentono di sintetizzare una distribuzione semplice?

L'insegnante stimola l'attenzione degli studenti sull'individuazione della moda per una distribuzione di un carattere qualitativo e il calcolo della media aritmetica per un carattere quantitativo discreto.
Con l'utilizzo della tabella pivot si ottengono rapidamente le seguenti distribuzioni:

Classe	n. studenti
Prima	90
Seconda	63
Terza	117
Quarta	47
Quinta	29
Totale	346

Tabella 6

La Tabella 6 riporta i dati degli iscritti nelle varie classi della scuola in esame. La moda è "Terza" in quanto è la modalità che presenta la frequenza maggiore.
Se si prende in considerazione un carattere quantitativo discreto è possibile determinare la media aritmetica e ad essa associare due indici di variabilità: la varianza e lo scarto quadratico medio.
L'insegnante suggerisce agli studenti di estrarre la tabella che contiene le informazioni relative all'età degli studenti.

Età	Totale
14	61
15	69
16	107
17	61
18	40
19	5
> 19	3
Totale	346

Tabella 7

Qual è l'età media? La colonna dei totali contribuisce a rispondere alla domanda precedente? Come può essere trattata l'ultima modalità?
Il calcolo della media aritmetica può essere fatto anche senza l'ausilio dello strumento informatico che, tuttavia, permette di far convergere l'attenzione dello studente principalmente sugli aspetti concettuali piuttosto che su quelli calcolatori. E' fondamentale che l'insegnante induca lo studente a riflettere sul valore interpretativo degli indici trovati senza indulgere eccessivamente sulle tecniche utilizzate per la loro determinazione. A tal riguardo è importante che l'insegnante apra una discussione con gli studenti circa le soluzioni (arbitrarie) per "chiudere" l'ultima modalità.
Una soluzione possibile al problema suddetto può essere ottenuta predisponendo un'elaborazione come quella proposta nella Tabella 8.

xi	ni	xi*ni	xi - m	(xi - m)^2	(xi - m)^2*ni
14	61	854	-1,93	3,7385	228,0499
15	69	1035	-0,93	0,8715	60,1315
16	107	1712	0,07	0,0044	0,4728
17	61	1037	1,07	1,1374	69,3794
18	40	720	2,07	4,2703	170,8126
19	5	95	3,07	9,4033	47,0163
20	3	60	4,07	16,5362	49,6086
Totale	346	5513			625,4711
Media aritmetica (m)		15,93			
Varianza		1,81			
Scarto quad. med.		1,34			

Tabella 8

Il problema dell'ultima classe è stato risolto utilizzando informazioni aggiuntive disponibili in segreteria sull'età degli studenti di più di 19 anni.
Le prime due colonne sono state estratte dalla tabella pivot mentre le altre colonne sono state calcolate usando le formule e le funzioni di Excel.

	A	B	C	D	E	F
1	xi	ni	xi*ni	xi - m	(xi - m)^2	(xi - m)^2*ni
2	14	61	=A2*B2	=A2-C10	=D2*D2	=E2*B2
3	15	69	=A3*B3	=A3-C10	=D3*D3	=E3*B3
4	16	107	=A4*B4	=A4-C10	=D4*D4	=E4*B4
5	17	61	=A5*B5	=A5-C10	=D5*D5	=E5*B5
6	18	40	=A6*B6	=A6-C10	=D6*D6	=E6*B6
7	19	5	=A7*B7	=A7-C10	=D7*D7	=E7*B7
8	20	3	=A8*B8	=A8-C10	=D8*D8	=E8*B8
9	Totale	=SOMMA(B2:B8)	=SOMMA(C2:C8)			=SOMMA(F2:F8)
10	Media	aritmetica (m)	=C9/B9			
11	Varianza		=F9/B9			
12	Scarto quad. med.		=RADQ(C11)			
13						

Tabella 9

In Tabella 9 sono riportate le formule e le funzioni usate per il calcolo degli indici. L'insegnante fa notare l'uso dei riferimenti relativi e dei riferimenti assoluti nelle formule e l'impiego della funzione "Somma( )" per calcolare la somma dei valori in colonna e della funzione "Radq( )" per il calcolo della radice quadrata.
La terza colonna contiene il prodotto tra le singole modalità e le rispettive frequenze assolute. A tal riguardo l'insegnante può stimolare l'attenzione degli studenti sul fatto che tale operazione consente di "pesare" in modo diverso le singole modalità del carattere "Età".
Il rapporto tra la somma della terza colonna e la somma della seconda permette di calcolare la media aritmetica.
La quinta e la sesta colonna consentono di calcolare la varianza, ottenuta come rapporto tra la somma della sesta colonna con la somma delle seconda.

Quali informazioni si possono trarre dalla varianza? Cosa si può dire sulla sua unità di misura? E' la stessa del fenomeno del quale si vuol indagare la variabilità?
La radice quadrata della varianza fornisce lo scarto quadratico medio che è anch'esso una misura della variabilità.
Quali vantaggi ha lo scarto quadratico medio rispetto alla varianza? Quanto dista l'età 14 dalla media? E l'età 19? Possiamo esprimere queste due distanze prendendo come unità di misura lo scarto quadratico medio? Immaginiamo di avere un'altra distribuzione sempre con la stessa media 15,93 anni, e con scarto quadratico medio 3,2 anni, se uno studente ha 19 anni la sua distanza dalla media ha la stessa importanza di prima? In quale distribuzione lo studente è più vicino alla media? So lo scarto quadratico medio fosse pari a 0, cosa significherebbe?
La varianza e lo scarto quadratico medio sono misure idonee a confrontare la variabilità di due distribuzioni riferite a caratteri diversi? Perché?

Sesta fase
C'è la possibilità di calcolare indici sintetici di una distribuzione semplice rispetto ad un carattere continuo avente modalità suddivise in classi? C'è la possibilità di calcolare la media e lo scostamento quadratico medio, analogamente a quanto si è visto nella quinta fase?
Con l'utilizzo della tabella pivot applicata sul database si ottiene rapidamente la seguente distribuzione che fa riferimento ai tempi che gli studenti impiegano per arrivare a scuola.

Tempo di percorrenza (min.)	N° studenti
fino a 15 min.	145
da 16 a 30 min.	103
da 31 a 45 min.	65
da 46 a 60 min.	24
oltre i 60 min.	9
Totale	346

Tabella 10

L'insegnante stimola l'attenzione degli studenti sul calcolo della media aritmetica e degli indici di variabilità in questa nuova situazione.

ei	es	ni	xic	xic*ni	xic - m	(xic - m)^2	(xic - m)^2*ni
0,5	15,5	145	8	1160	-14,84	220,08	31912,31669
15,5	30,5	103	23	2369	0,16	0,03	2,79534064
30,5	45,5	65	38	2470	15,16	229,97	14948,00682
45,5	60,5	24	53	1272	30,16	909,91	21837,87678
60,5	70	9	70	630	47,16	2224,51	20020,6142
Totale		346		7901			88721,60983
media aritmetica (m)		22,84		legenda:	ei	estremo inferiore	
varianza		256,42			es	estremo superiore	
scarto quad. medio		16,01			xic	valore centrale	

Tabella 11

Il caso in esame pone allo studente problemi aggiuntivi rispetto al caso precedente? Dove deve porre l'attenzione lo studente? Nell'algoritmo di calcolo o altrove?
L'insegnante guida lo studente a riflettere sulla necessità di individuare con attenzione gli estremi della prima classe e la modalità di "chiusura" dell'ultima. Nell'esempio di Tabella 5 il valore rappresentativo dell'ultima classe è stato individuato in base alle informazioni aggiuntive a disposizione (i 9 studenti appartenenti all'intervallo in esame, risiedendo tutti nello stesso comune, impiegano 70 minuti per raggiungere la scuola).
E' possibile costruire l'istogramma di frequenze? L'insegnante guida gli studenti a costruire la Figura 8, usando il grafico dispersione (XY) di Excel e dando una tabella di punti xy che rappresentano i vertici di ogni rettangolo. Dovendo rappresentare sul grafico la media, dove la si posizionerà? Sull'asse delle ascisse o delle ordinate? Perché?

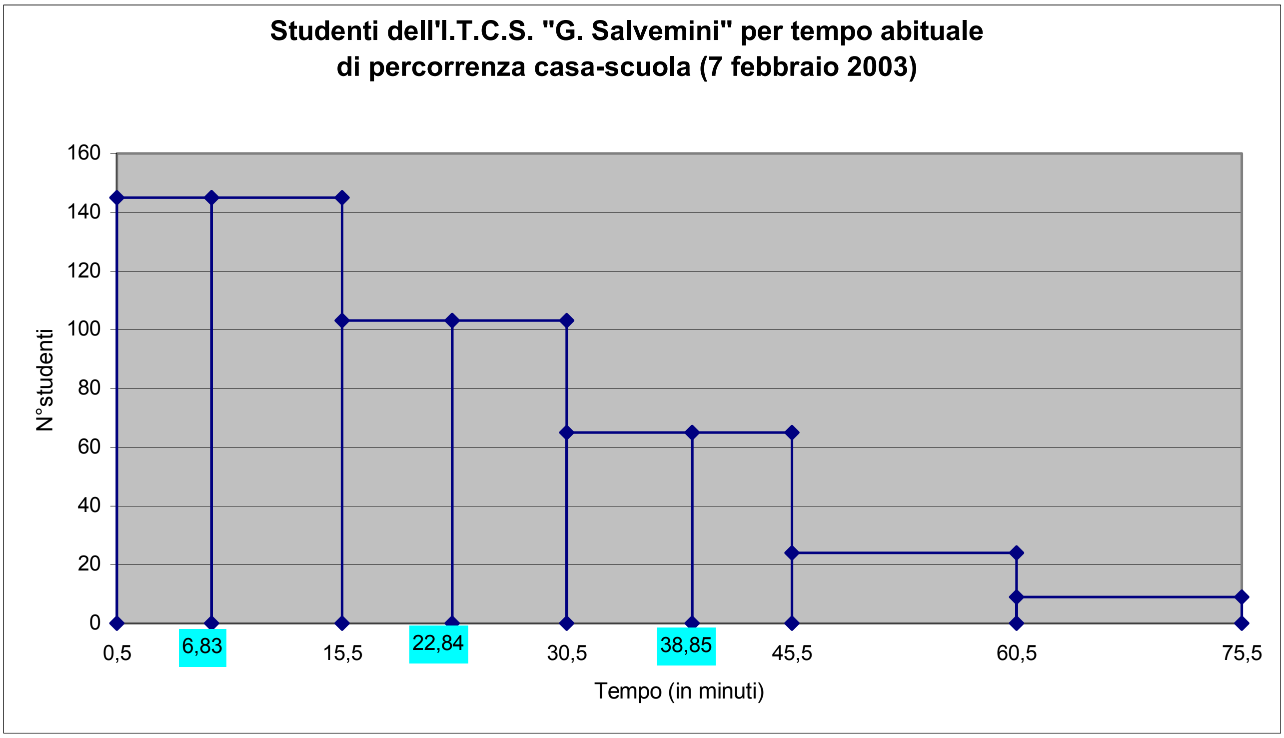


Figura 8

Dovendo rappresentare lo scarto quadratico medio come si farà? L'insegnante guida gli studenti a considerare che lo scarto dalla media non è altro che una distanza dalla media. Dunque lo scarto quadratico medio, come media quadratica di distanze, dal punto di vista geometrico è un segmento, che è possibile disegnare sull'asse delle ascisse, a sinistra e a destra della media aritmetica. Dunque come si sono ottenuti i punti di ascissa 6,83 e 38,85? Si può calcolare qual è il numero di studenti che impiegano fra 6,83 e 38,85 minuti? Quale percentuale rappresentano sul totale degli studenti?

Un gioco con tre dadi

Livello scolare: 1° biennio

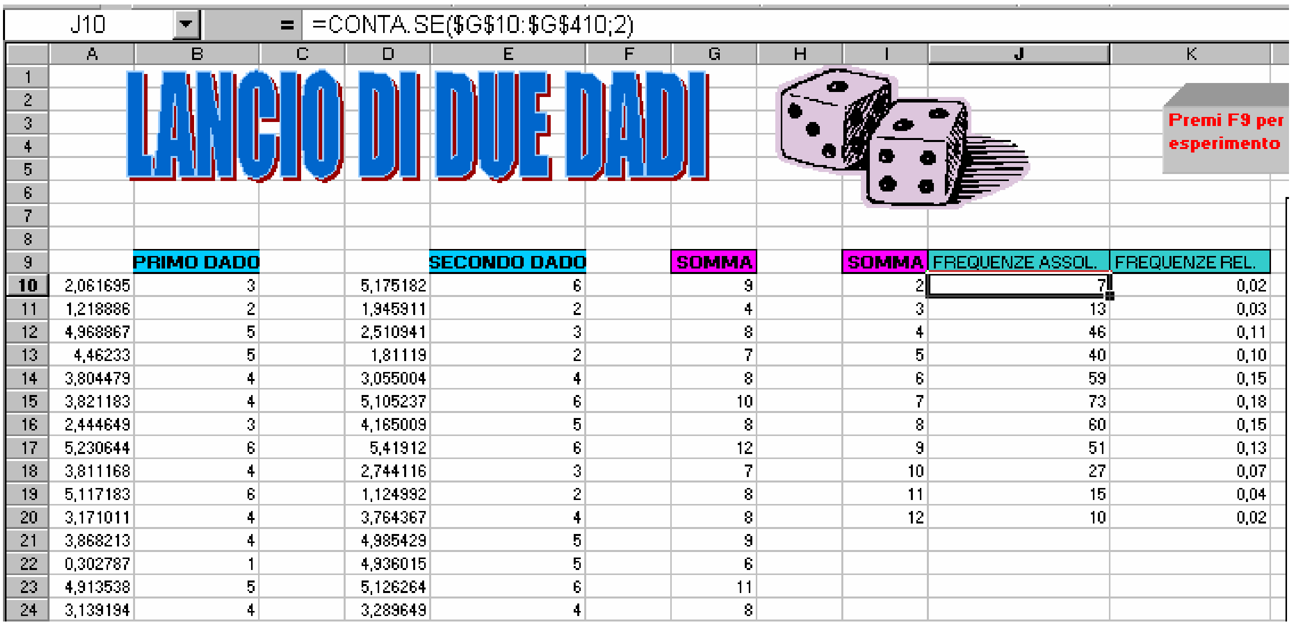
Abilità interessate	**Conoscenze**	**Nuclei coinvolti**	**Collegamenti esterni**
Costruire lo spazio degli eventi in casi semplici e determinarne la cardinalità. Valutare la probabilità in diversi contesti problematici. Distinguere tra eventi indipendenti e non.	Eventi e operazioni con gli eventi. Significato della probabilità e sue valutazioni.	Dati e previsioni Risolvere e porsi problemi Laboratorio di matematica	Giochi

Contesto
Giochi, probabilità.

Il contesto è di tipo matematico, in particolare riguarda l'ambito probabilistico; ha anche aspetti collegati al contesto dei giochi.
Questa attività può essere introdotta anche al primo anno di biennio dopo aver trattato le distribuzioni di frequenze e aver introdotto le definizioni di probabilità. Non occorrono né i teoremi sulle probabilità né altri concetti. L'attività ha lo scopo di indurre ad una individuazione corretta dello spazio degli eventi, in modo che gli studenti sappiano distinguere tra evento (uscita di un certo risultato nei dadi) ed evento elementare. L'obiettivo è condurre gli studenti alla scoperta che non tutti gli eventi hanno la stessa probabilità e che la probabilità dipende dal modo in cui l'esperimento è definito.
Per la simulazione al computer sono necessari alcuni prerequisiti di conoscenza del foglio elettronico: come si inseriscono i dati, come si inserisce una formula, come si copia una formula, riferimenti relativi e assoluti alle celle, come si crea un grafico. Le funzioni "Casuale( )" e "Conta.Se( )" possono essere introdotte anche in questo contesto.

Descrizione dell'attività
L'insegnante propone il seguente problema: lancia tre dadi e, ad ogni lancio, elimina il dado col punteggio maggiore annotando la somma dei due dadi rimasti (se quelli col punteggio maggiore sono due, eliminarne uno qualsiasi).
Per una migliore comprensione del problema l'insegnante suggerisce una lettura attenta del testo e chiede agli studenti di esemplificare qualche situazione che si può verificare. Pone alla fine la domanda:
I risultati sono gli stessi che nel lancio di due dadi?

<u>Prima fase</u>
Il problema si può affrontare dopo aver trattato preliminarmente il lancio di due dadi.
L'insegnante guida gli studenti a costruire il grafico di frequenze per il lancio di due dadi dopo aver effettuato l'esperimento manualmente o tramite la simulazione al computer ed aver calcolato la somma dei punteggi ottenuti.
Il foglio elettronico, riportato in Figura 1, è stato costruito riportando nella colonna B le uscite del primo dado e nella colonna E le uscite del secondo dado (ottenuti tramite la funzione casuale); nella colonna G si è effettuata la somma. Per poter calcolare le frequenze assolute e relative (colonne J e K) si è usata la funzione "Conta.Se( )".



	A	B	C	D	E	F	G	H	I	J	K
9		PRIMO DADO			SECONDO DADO		SOMMA		SOMMA	FREQUENZE ASSOL.	FREQUENZE REL.
10	2,061695	3		5,175182	6		9		2	7	0,02
11	1,218886	2		1,945911	2		4		3	13	0,03
12	4,968867	5		2,510941	3		8		4	46	0,11
13	4,46233	5		1,81119	2		7		5	40	0,10
14	3,804479	4		3,055004	4		8		6	59	0,15
15	3,821183	4		5,105237	6		10		7	73	0,18
16	2,444649	3		4,165009	5		8		8	60	0,15
17	5,230644	6		5,41912	6		12		9	51	0,13
18	3,811168	4		2,744116	3		7		10	27	0,07
19	5,117183	6		1,124992	2		8		11	15	0,04
20	3,171011	4		3,764367	4		8		12	10	0,02
21	3,868213	4		4,985429	5		9				
22	0,302787	1		4,936015	5		6				
23	4,913538	5		5,126264	6		11				
24	3,139194	4		3,289649	4		8				

Figura 1

Nel caso si sia effettuata la simulazione al computer l'insegnante invita ad osservare le colonne in cui sono riportati i risultati relativi al primo ed al secondo dado e chiede se i dadi sono truccati. Come sono tra loro i punteggi dei due dadi? Come avvengono i lanci?

Dall'esame del grafico in Figura 2 gli studenti possono dedurre che i casi centrali sono i più frequenti. Da cosa dipende?

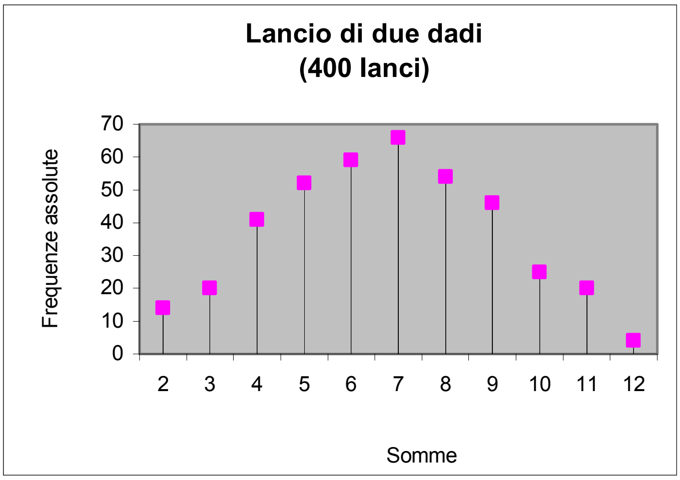


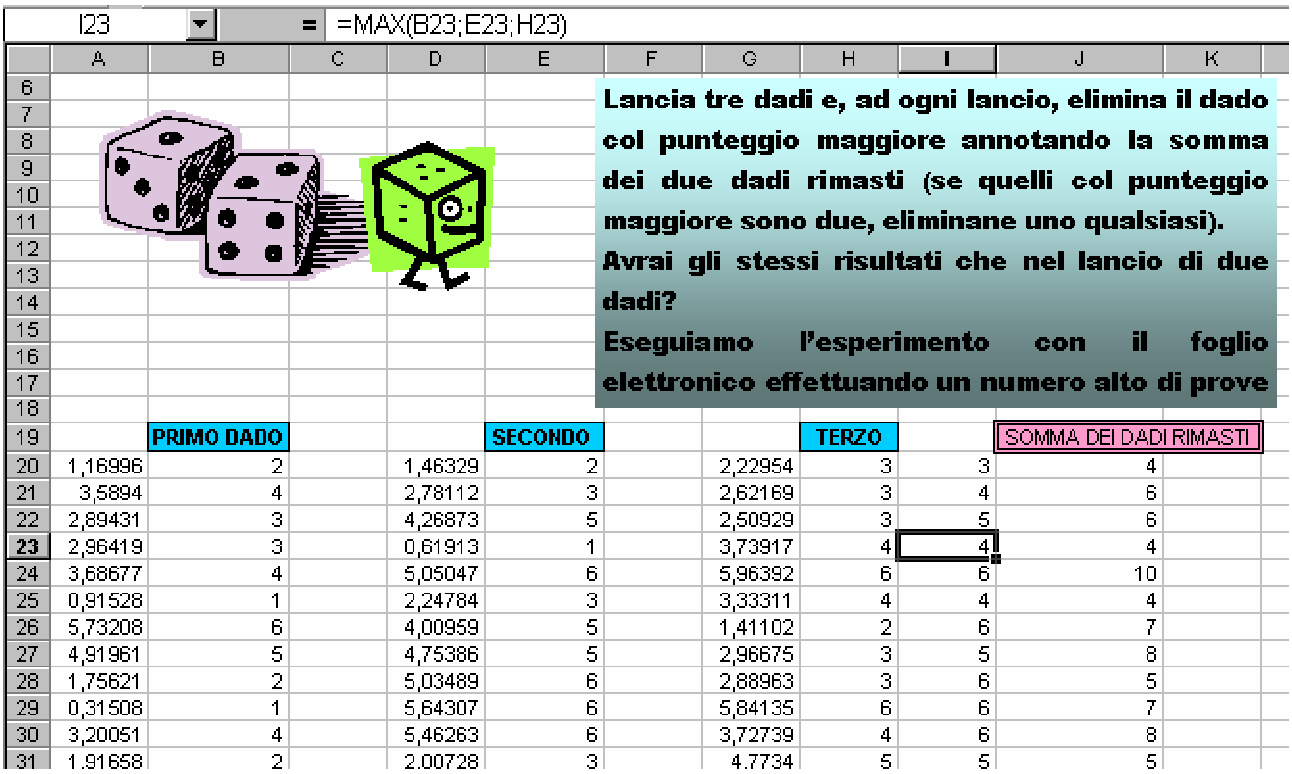
Figura 2

Si invitano gli studenti a valutare le probabilità dei diversi eventi possibili.
I casi possibili nel lancio di due dadi sono tutte le coppie che si possono ottenere associando una faccia del primo dado con una qualunque del secondo dado $6 \cdot 6 = 36$.
Non tutti i risultati, però, hanno la stessa probabilità: infatti, ad esempio, il numero 6 si può ottenere in più modi che non il numero 3.
Gli studenti saranno guidati a compilare la seguente tabella.

EVENTO	Modalità di presentazione	Numero casi possibili	Probabilità
uscita del 2	1+1	1	1/36
uscita del 3	1+2 2+1	2	2/36
uscita del 4	1+3 2+2 3+1	3	3/36
uscita del 5	1+4 2+3 3+2 4+1	4	4/36
uscita del 6	1+5 2+4 3+3 4+2 5+1	5	5/36
uscita del 7	1+6 2+5 3+4 4+3 5+2 6+1	6	6/36
uscita del 8	2+6 3+5 4+4 5+3 6+2	5	5/36
uscita del 9	3+6 4+5 5+4 6+3	4	4/36
uscita del 10	4+6 5+5 6+4	3	3/36
uscita del 11	5+6 6+5	2	2/36
uscita del 12	6+6	1	1/36
Totale		36	1

Si pongono agli studenti le seguenti domande: perché il grafico sperimentale non risulta perfettamente simmetrico come ci si aspetterebbe dalla valutazione di probabilità? Il grafico avrebbe avuto un andamento più vicino alle previsioni se il numero di lanci fosse stato maggiore?
Si può invitare gli studenti a ripetere la simulazione con un numero maggiore di lanci.
Il foglio elettronico permette agevolmente sia di ripetere più volte l'esperimento con lo stesso numero di lanci (basta premere un tasto per vedere simultaneamente come si modifica il grafico) sia di variare il numero di lanci.
L'insegnante guida gli studenti ad osservare che il grafico muta anche se il numero di lanci è identico. Perché? Quante sono le simulazioni di lanci che si possono fare, fissato il numero di lanci? Allora l'insieme di *n* lanci studiato ed osservato è forse un "campione casuale"?
Appare, inoltre, chiaro che all'aumentare del numero dei lanci la frequenza si avvicina alla valutazione teorica di probabilità.
Si può concludere che su un gran numero di prove ci si può attendere di avere frequenze di uscita sempre più vicine alla valutazione di probabilità?
E' un primo approccio alla legge dei grandi numeri.

<u>Seconda fase</u>
Nella fase successiva si propone il problema dei tre dadi .
Si invitano gli studenti ad eseguire l'esperimento lanciando i dadi ed effettuando un numero alto di prove. Si possono far lavorare gli studenti in gruppo e far loro costruire il grafico delle frequenze.
E' opportuno comunque far ripetere l'esperimento con una simulazione al computer.
Il foglio elettronico, riportato in Figura 3, è stato costruito riportando nella colonna B le uscite del primo dado, nella colonna E le uscite del secondo e nella colonna H le uscite del terzo (ottenute tramite la funzione casuale); per calcolare la somma dei dadi rimasti(colonna J), dopo aver scartato il punteggio maggiore, si è utilizzata, per ogni terna di valori, la funzione "Max( )", che restituisce il valore massimo.



I23 = =MAX(B23;E23;H23)

Lancia tre dadi e, ad ogni lancio, elimina il dado col punteggio maggiore annotando la somma dei due dadi rimasti (se quelli col punteggio maggiore sono due, eliminane uno qualsiasi). Avrai gli stessi risultati che nel lancio di due dadi? Eseguiamo l'esperimento con il foglio elettronico effettuando un numero alto di prove

	A	B	C	D	E	F	G	H	I	J
19		PRIMO DADO			SECONDO			TERZO		SOMMA DEI DADI RIMASTI
20	1,16996	2		1,46329	2		2,22954	3	3	4
21	3,5894	4		2,78112	3		2,62169	3	4	6
22	2,89431	3		4,26873	5		2,50929	3	5	6
23	2,96419	3		0,61913	1		3,73917	4	4	4
24	3,68677	4		5,05047	6		5,96392	6	6	10
25	0,91528	1		2,24784	3		3,33311	4	4	4
26	5,73208	6		4,00959	5		1,41102	2	6	7
27	4,91961	5		4,75386	5		2,96675	3	5	8
28	1,75621	2		5,03489	6		2,88963	3	6	5
29	0,31508	1		5,64307	6		5,84135	6	6	7
30	3,20051	4		5,46263	6		3,72739	4	6	8
31	1,91658	2		2,00728	3		4,7734	5	5	5

Figura 3

Dall'esame del grafico di Figura 4 gli allievi possono notare il diverso andamento ottenuto in questo esperimento rispetto a quello relativo al lancio di due dadi.

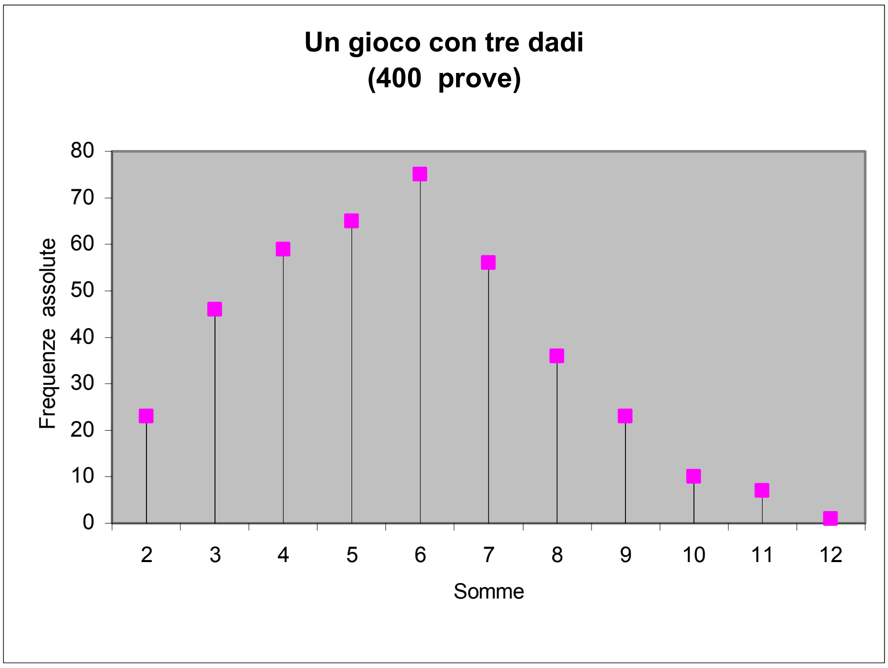
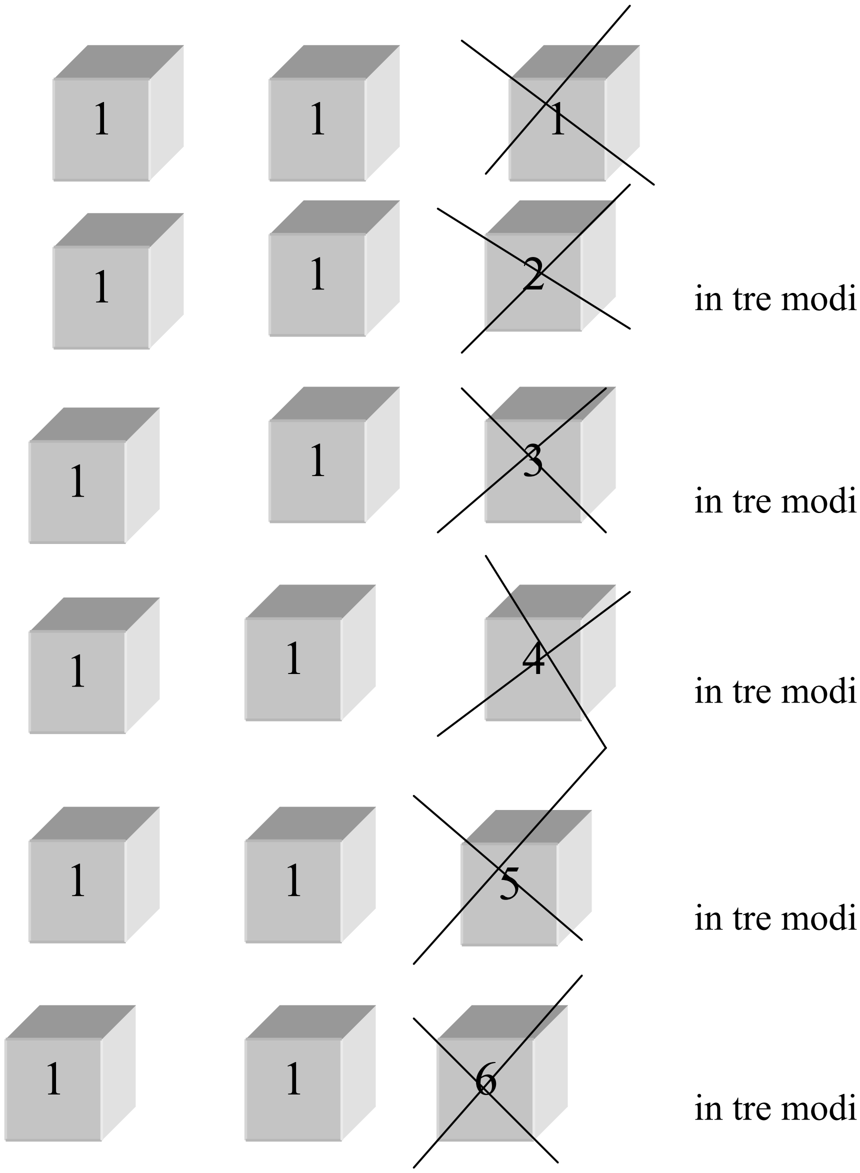
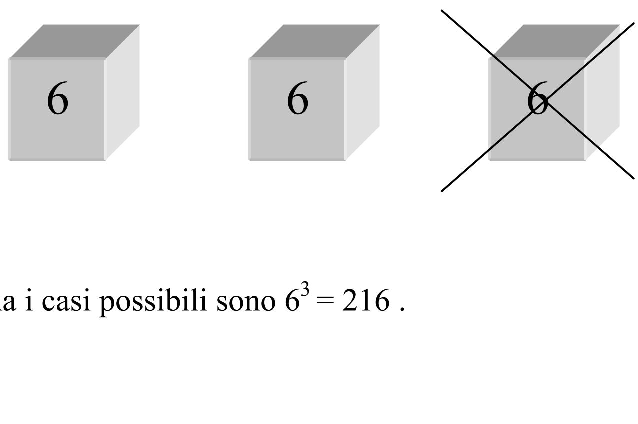


Figura 4

Si faranno riflettere gli allievi sul fatto che, anche se gli eventi sono sempre l'uscita dei numeri da 2 a 12, il fatto di aver eliminato l'esito del terzo dado, quello col punteggio maggiore, non può essere ignorato. Si invitano, poi, gli allievi a calcolare i diversi casi possibili per ogni evento.
Per esempio l'evento E_1 : "uscita del 2" e l'evento E_2 "uscita del 12" non hanno in questo caso la stessa probabilità .
Infatti, l' evento E_1 si realizza in 16 modi:



mentre l'evento E_2 si realizza in un sol modo:



In questo problema i casi possibili sono $6^3 = 216$.

A questo punto si invitano gli studenti a calcolare la probabilità dei diversi eventi, dividendoli in gruppi perché il calcolo seppure semplice è un po' laborioso, e si ottiene:

P(uscita del 2) = $16/ 6^3$
P(uscita del 3) = $27/ 6^3$
P(uscita del 4) = $34/ 6^3$
P(uscita del 5) = $36/ 6^3$
P(uscita del 6) = $34/ 6^3$
P(uscita del 7) = $27/ 6^3$
P(uscita del 8) = $19/ 6^3$
P(uscita del 9) = $12/ 6^3$
P(uscita del 10) = $7/ 6^3$
P(uscita del 11) = $3/ 6^3$
P(uscita del 12) = $1/ 6^3$

Le valutazioni teoriche si accordano con i risultati ottenuti ?

Il problema delle parti

Livello scolare: 1° biennio.

Abilità interessate	**Conoscenze**	**Nuclei coinvolti**	**Collegamenti esterni**
Valutare la probabilità in diversi contesti problematici. Distinguere tra eventi dipendenti e indipendenti.	Eventi e operazioni con gli eventi. Eventi incompatibili; eventi esaustivi. Significato di probabilità e sue valutazioni. Probabilità condizionata, probabilità composta; probabilità totale.	<u>Dati e previsioni</u> Argomentare, congetturare, dimostrare Risolvere e porsi problemi Laboratorio di matematica	Storia

Contesto
Giochi, probabilità.

Il contesto è di tipo matematico, in particolare riguarda l'ambito probabilistico; ha anche aspetti collegati ai giochi.
Questa attività può essere introdotta, nella forma che qui viene proposta, in una prima o seconda classe del primo biennio; non è richiesto alcun prerequisito di calcolo delle probabilità. L'attività ha come obiettivo primario proprio quello di introdurre al pensiero probabilistico e va quindi svolta secondo i tempi della didattica lunga, tipici del contesto laboratoriale, dando agli studenti il tempo necessario per appropriarsi dei concetti e delle prime tecniche del calcolo delle probabilità. L'attività è suscettibile di ulteriori sviluppi nel secondo biennio, quando si voglia introdurre o applicare la distribuzione binomiale.
Il contesto è quello dei giochi, con la storia della matematica che dovrebbe favorire la costruzione di significati, presentando genesi ed evoluzione del concetto di probabilità. In questo caso la storia della matematica funge anche da elemento motivante, soprattutto quando evidenzia che le soluzioni non adeguate degli studenti al problema proposto sono state date, nel corso della storia, dai matematici che hanno affrontato lo stesso problema: questa considerazione dà dignità agli "errori" degli studenti e fa capire che, talvolta, una risoluzione adeguata e soddisfacente a un problema può essere determinata solo con un cambio di prospettiva reso possibile dallo sviluppo di nuovi concetti, come, in questo caso, quello di probabilità.

Descrizione dell'attività
Il problema che viene proposto in questa attività è noto in letteratura come "problema della divisione della posta in gioco" o "problema delle parti". La prima versione che ci è nota è presente in un manoscritto di anonimo del 1400 circa, anche se il problema deve la sua fama a Luca Pacioli, che lo propose nel 1494 nella *Summa de arithmetica, geometria, proportioni et proportionalità*.
Il problema riguarda la suddivisione della posta fra due giocatori di "uguale valore" (ossia che hanno la stessa probabilità di guadagnare un punto), che sono costretti a interrompere la partita prima che uno di essi abbia totalizzato il numero di punti necessario per vincere la partita.
L'attività viene strutturata in quattro fasi. Nella prima si propone di risolvere il problema in piccoli gruppi di lavoro (ovviamente) collaborativi; nella seconda, in una discussione matematica rivolta all'intera classe, si evidenziano i dati del problema e alcuni modi di rappresentarli e si discutono le

soluzioni proposte dai vari gruppi; nella terza fase si ricostruiscono i gruppi con la consegna di sistemare il procedimento risolutivo prima proposto, alla luce dei nuovi elementi emersi nella discussione collettiva. Nella quarta fase, infine, si passa alla discussione di bilancio, nella quale si costruiscono, gradualmente, le tecniche necessarie alla risoluzione del problema, eventualmente traendo spunto dalla storia della matematica e, in particolare, dalle soluzioni proposte da Pascal e da Fermat.

<u>Prima fase</u>
L'insegnante propone la "situazione – problema" sotto riportata a gruppi collaborativi formati da tre – quattro studenti di livello di preparazione simile (gruppi omogenei al loro interno).
Situazione – problema:
In una taverna del piccolo paese di Matelandia, Ariele e Calibano giocano a testa e croce con una moneta a due facce non truccata. A ogni lancio viene assegnato 1 punto al giocatore che indovina l'esito. Vince tutta la posta di 24 denari (12 dei quali sono di Ariele e 12 di Calibano) chi per primo totalizza 6 punti.
I giochi d'azzardo sono però proibiti a Matelandia e il gendarme Prospero, venuto a conoscenza della partita che si sta giocando, si avvia verso la taverna per arrestare Ariele e Calibano. Informati del pericolo, i due giocatori interrompono la partita sul 5 a 3 per Ariele e fuggono, ciascuno con i 12 denari messi per formare la posta, concordando di ritrovarsi il giorno dopo senza finire la partita ma solo per dividere equamente la posta in gioco.
Problema:
Come dovrebbero, Ariele e Calibano, dividersi i 24 denari in modo tale che la suddivisione sia equa, ossia in modo tale da tenere conto del fatto che avevano contribuito alla posta con 12 denari ciascuno e che quando la partita è stata interrotta il punteggio era 5 a 3 per Ariele?
Tutte le volte che il problema è stato proposto a studenti del primo biennio che non avevano ancora affrontato sistematicamente il calcolo delle probabilità, la prima reazione è stata di stupore, perché il problema è ritenuto troppo facile, banale. Ci si può aspettare che, in breve tempo, i vari gruppi di studenti giungano a proporre la seguente risoluzione (o una equivalente): dividere la posta, ossia i 24 denari, per il numero totale di partite giocate, e moltiplicare il risultato prima per 5, per ottenere i denari spettanti ad Ariele, poi per 3, in modo da ottenere quelli di Calibano.
L'insegnante deve verificare, prima di passare alla seconda fase, che gli tutti studenti abbiano raggiunto un buon livello di fiducia nella risoluzione proposta.

<u>Seconda fase</u>
L'insegnante avvia una discussione matematica alla presenza dell'intera classe, invitando i rappresentanti di alcuni gruppi a presentare la propria strategia risolutiva[1]. L'insegnante deve aver cura di predisporre la classe alla discussione collettiva, sia lodando il lavoro fino ad allora svolto, sia ricordando che tra gli elementi di valutazione della discussione vi è la capacità di analisi critica delle idee altrui.
In seguito l'insegnante fa notare che, in generale, i dati del problema sono il numero n di punti necessari per vincere la partita e i numeri a e b di punti che hanno totalizzato, rispettivamente, i due giocatori al momento dell'interruzione della partita. Tali dati possono essere rappresentati con la notazione $[n: a ; b]$ ove a e b sono numeri naturali minori di n, ma anche con la notazione $[-p; -q]$, dove p e q sono i punti che mancano, al momento dell'interruzione della partita, rispettivamente al primo e al secondo giocatore per vincere l'intera posta. In questo momento l'insegnante si limita a

[1] Non è necessario che tutti gli studenti presentino la soluzione proposta, soprattutto se le strategie risolutive sono fortemente simili, come avviene spesso in questo caso. L'insegnante invita un gruppo a presentare la propria strategia risolutiva e a chiedere agli altri gruppi di limitarsi a interventi che servano per precisare, aggiungere qualcosa o, eventualmente, per criticare e contestare le strategie risolutive proposte dal gruppo che ha iniziato le presentazioni.

presentare le due differenti modalità di rappresentazione dei dati a disposizione, senza evidenziare che la seconda modalità è enormemente più suggestiva della prima[2].
A questo punto l'insegnante ha il compito di mettere in crisi la fiducia degli studenti nella soluzione proposta. Allo scopo può chiedere a qualche studente di giocare a testa e croce, contro di lui. Se il numero degli studenti è sufficientemente elevato, è abbastanza probabile che si verifichi la situazione che vede l'insegnante, dopo il primo lancio della moneta, vincere per 1 a 0 su uno studente. Si tratta di una situazione critica che consente all'insegnante di mettere in crisi la fiducia degli studenti nella soluzione proposta (quella di dividere la posta in parti direttamente proporzionali al punteggio dei due giocatori). Infatti, se la partita viene interrotta sull'1 a 0 per l'insegnante, con la strategia di suddividere la posta in parti direttamente proporzionali ai punteggi dei due giocatori, tutta la posta andrebbe all'insegnante e nulla allo studente. Ma quale studente è disposto ad accettare una suddivisione di questo tipo visto che sta perdendo solo per 1 a 0 e avrebbe ancora ottime possibilità di recuperare? La proposta di suddividere in parti proporzionali al punteggio appare, in questo caso, manifestamente non equa.
L'insegnante, per far capire agli studenti che la strategia risolutiva da loro proposta ha comunque una propria dignità e non è indice di ingenuità e superficialità da parte loro, può far presente che la stessa soluzione è stata proposta nel 1494 da Luca Pacioli, un grande matematico, secondo cui la suddivisione equa sarebbe stata di 15 denari per Ariele e di 9 per Calibano, ossia la stessa suddivisione proposta da alcuni degli studenti. Inoltre il problema, nonostante i tentativi di risoluzione di altri matematici, come Cardano, Tartaglia e di Pietro Cataneo, non fu risolto in modo soddisfacente fino alla metà del diciassettesimo secolo, quando fu data una risposta adeguata con Pascal e, indipendentemente, con Fermat.
L'insegnante, in questa fase, può approfittare dell'occasione per fissare l'attenzione su alcuni termini come suddivisione in parti direttamente proporzionali al punteggio e suddivisione equa, iniziando a precisarne il significato.

Terza fase

Per trovare una strategia risolutiva alternativa, che possa soddisfare tutti i componenti della classe l'insegnante può ricostituire i gruppi di studenti invitandoli a ripensare alla strategia risolutiva, cercando di trovarne una che sia adeguata anche alla trattazione di casi limite come quello dell' 1 a 0 per uno dei giocatori.
L'insegnante, in questa fase, deve intervenire sistematicamente nei gruppi di lavoro per evitare che si sclerotizzino posizioni semplificatrici del tipo "la partita può essere ripresa in seguito a partire dal punteggio sul quale è stata interrotta" oppure "si divide la posta a metà indipendentemente dal punteggio" o, ancora, "se il gioco d'azzardo è proibito, allora i giocatori devono essere multati e non spetta loro alcuna somma". Si dovrebbe far presente agli studenti che le risoluzioni proposte non devono far perdere senso al problema: la strategia di cambiare le richieste di un problema quando non lo si sa risolvere porta a non soddisfare la richiesta iniziale. Se Ariele e Calibano hanno chiesto aiuto per risolvere il loro problema, l'obiettivo dell'insegnante e quello degli studenti è cercare di risolverlo con i vincoli che loro hanno imposto e non quello di cambiare il problema posto.
In questa terza fase ci si aspetta che da qualche gruppo di lavoro emergano soluzioni che affermino che una suddivisione equa deve tenere conto non solo delle partite giocate, ma anche di quelle che rimangono da giocare per raggiungere i 6 punti necessari a vincere l'intera posta (è stato osservato, in alcuni gruppi di lavoro, l'attenzione esclusiva, con tipico pensiero probabilistico, alle sole partite che ancora rimangono da giocare).

[2] Con la seconda modalità, infatti, si fissa l'attenzione sul numero di punti che ancora servono per vincere e non sul punteggio ottenuto al momento dell'interruzione: è il cambio di orizzonte richiesto dal pensiero probabilistico (non guardare a quanto è già accaduto, ma prendere in considerazione ciò che deve ancora accadere o, meglio, il mondo degli eventi possibili, per effettuare previsioni).

Può essere interessante osservare se gli studenti, inconsapevolmente, ripropongano alcune strategie risolutive che i matematici che si sono occupati del problema hanno fatto pervenire attraverso testi a stampa[3]. Se questa situazione si verifica, è possibile utilizzare le fonti storiche in una specie di gioco "voci – eco", nel quale le voci della storia, ossia le soluzioni proposte dai matematici, fanno eco alle voci della classe, ossia alle soluzioni proposte dagli studenti. In alcuni casi le voci della storia danno dignità agli errori commessi dagli studenti; altre volte contribuiscono a dar forza a idee espresse da alcuni studenti, consentendo che esse vengano riconosciute da tutta la classe diventando, a tutti gli effetti, oggetto di discussione. Ciò, oltre a consentire di creare un atteggiamento riflessivo nei confronti dei concetti messi in gioco nel problema, favorisce l'instaurarsi e lo svilupparsi della discussione matematica con inevitabili benefici per la socializzazione e la condivisione del sapere.

Quarta fase

L'insegnante si incarica di proporre una sistemazione delle varie soluzioni proposte e della loro evoluzione verso la soluzione ottenuta utilizzando il calcolo delle probabilità. Sottolinea il fatto che in questa soluzione la prospettiva, rispetto alla suddivisione in parti direttamente proporzionali al punteggio al momento dell'interruzione del gioco, è profondamente cambiata: per suddividere la posta l'attenzione è rivolta ai punti che mancano per vincere e non a quelli già ottenuti dai due giocatori (può essere utile a questo punto ritornare sulle due modalità di rappresentazione dei dati del problema proposto, commentando la maggiore utilità della notazione [- p ; - q]). L'idea è quindi quella di suddividere la posta in parti proporzionali alle possibilità che i due giocatori avrebbero di vincere l'intera posta al momento dell'interruzione, se il gioco potesse continuare. Il problema diventa quindi quello di misurare tale possibilità.

A questo punto l'insegnante può utilizzare le voci della storia, presentando in classe una sintesi dell'approccio di Pascal e di quello di Fermat alla risoluzione del problema, oppure può limitarsi a presentare alcuni elementi di calcolo delle probabilità, in particolare la legge delle probabilità totali e delle probabilità composte, magari con qualche cenno all'uso dei diagrammi ad albero come strumento di efficace rappresentazione di situazioni probabilistiche. Sia in un caso che nell'altro, la presentazione non può essere confinata in tempi e spazi angusti, per la delicatezza dei concetti coinvolti e quindi deve avere i ritmi e i tempi della didattica lunga.

Qui di seguito si dà un primissimo cenno dell'approccio alla risoluzione del problema da parte di Pascal e si propone poi una soluzione con il moderno linguaggio dei diagrammi ad albero.

Soluzione "alla Pascal"	tradotta in termini moderni
Sul punteggio di 5 a 3 per Ariele, se si gioca un'altra partita e se Ariele vince, allora ad Ariele va l'intera posta, mentre se vince Calibano vanno sul 5 a 4. Allora ad Ariele spetta almeno metà della posta, ossia 12 denari. Sul 5 a 4 per Ariele, se si gioca un'altra partita e Ariele vince, allora ritira tutta la posta rimanente, mentre se vince Calibano vanno sul 5 a 5. Allora ad Ariele vanno, oltre ai 12 denari già stabiliti, almeno la metà dei 12 rimanenti, ossia 12+6= 18. Sul 5 a 5 si può giocare al più un'altra partita. Chi fra Ariele e Calibano vince ritira tutta la stessa posta rimanente. Quindi, se interrompono sul 5 a 5	La speranza di vittoria di Ariele è legata al verificarsi di almeno una fra le seguenti successioni di eventi: E_1 ; E_2E_1 ; $E_2E_2E_1$ Ove: E_1 è l'evento "Ariele guadagna un punto"; E_2 è l'evento "Calibano guadagna un punto". Poiché E_1 ed E_2 hanno probabilità ½ , per la regola della probabilità composta di eventi indipendenti, si ha che E_1, E_2E_1, $E_2E_2E_1$ hanno, rispettivamente, probabilità uguali a 1/2 , 1/4, 1/8. Quindi la probabilità che Ariele ha di vincere è, per la regola sulla probabilità totale

[3] A questo proposito si può far riferimento a uno dei lavori riportati nei riferimenti bibliografici.

devono dividersi la posta rimanente, ossia 3 denari a testa. Quindi, se il gioco viene interrotto sul 5 a 3 per Ariele, ad Ariele vanno 12+6+3=21 denari, mentre a Calibano ne spettano 24-21=3.	per eventi incompatibili, 1/2 + 1/4 + 1/8 = 7/8. Un'equa ripartizione dei 24 denari può essere effettuata suddividendo la posta in parti proporzionali alle probabilità di vittoria dei due giocatori: 21 denari ad Ariele e 3 a Calibano.

Risoluzione con l'aiuto dei diagrammi ad albero
Il seguente diagramma ad albero prende in considerazione tutte le possibili situazioni che potrebbero verificarsi se la partita continuasse, invece di essere interrotta sul 5 a 3 per Ariele (A indica la vittoria di Ariele e C quella di Calibano).

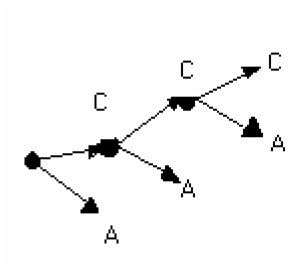


Figura 1

Il diagramma ad albero di Figura 1 consente di visualizzare lo spazio delle situazioni possibili: in questo caso è uno strumento particolarmente indicato, in quanto la partita è stata interrotta sul 5 a 3 e quindi lo spazio delle situazioni possibili non è eccessivamente complesso. Si fa notare agli studenti che Calibano, per arrivare a sei punti prima di Ariele, deve vincere tre partite consecutive senza che Ariele ne vinca alcuna. Ciò vuol dire che la probabilità di vincita di Calibano è $\frac{1}{8}$, mentre quella di Ariele è $\frac{7}{8}$. La suddivisione della posta in parti direttamente proporzionali alla probabilità che i due giocatori avrebbero di vincere il gioco al momento dell'interruzione, se questo fosse continuato è quindi 21 denari per Ariele e 3 per Calibano.

Possibili sviluppi
- Introduzione del concetto di gioco equo e relative applicazioni.
- La generalizzazione del problema della parti al caso [n: a ; b] e la distribuzione binomiale.
- Lettura e analisi del carteggio Pascal – Fermat sul problema delle parti.
- Il problema dei compleanni.

Elementi di prove di verifica

1. Un'urna contiene 6 palline nere e 5 bianche. Si eseguono due estrazioni successive, rimettendo, dopo la prima estrazione, la pallina estratta nell'urna. Qual è la probabilità di ottenere
a) due palline nere
b) una pallina bianca e una nera

2. Sia data un'urna contenente 5 palline, di cui 2 azzurre e 3 bianche. Effettuando due estrazioni successive in ciascuna delle quali si estrae una sola pallina che poi si inserisce nuovamente nell'urna, calcolare la probabilità di ottenere i seguenti eventi:
a) due palline azzurre
b) una pallina azzurra e una bianca

3. Un'urna contiene palline bianche, nere e rosse; sapendo che la probabilità di estrarre una pallina nera è 1/2, che quella di estrarre una pallina nera o bianca è 2/3 e che vi sono 2 palline bianche, è possibile determinare il numero di palline rosse e quello di palline nere? In caso di risposta negativa spiega perché; in caso di risposta affermativa, determina tali numeri.

4. Qual è la probabilità che, lanciando due volte un dado cubico regolare, con facce numerate da 1 a 6, escano due numeri multipli di 3?

5. Sapendo che, al gioco del Lotto, sulla ruota di Napoli il 12 non è uscito per 80 settimane e che il 20 è stato estratto nelle due ultime settimane, possiamo dire che alla prossima estrazione è più probabile che esca il 12 rispetto al 20? Perché?

6. Una classe è composta da 9 maschi e 11 femmine. Per partecipare a una rappresentazione musicale studentesca vengono estratti tre nominativi. Qual è la probabilità che almeno uno di essi sia quello di una studentessa?

7. Un numero di 5 cifre viene scritto nel sistema binario con la seguente procedure:
a) si scrive il numero 1
b) si estraggono da un'urna (contenente cento cartellini con la cifra 0 e cento cartellini con la cifra 1) quattro cartellini in successione, rimettendo ogni volta i cartellini estratti nell'urna e scrivendo, a ogni estrazione, la cifra estratta di seguito a quelle già scritte.

Qual è la probabilità che si ottenga, in questo modo, un numero minore del numero che, in base dieci, si rappresenta con la scrittura 30?

8. In un sacchetto vi sono palline indistinguibili fra loro al tatto, ma colorate in modo diverso. Alcune sono blu, altre rosse, altre verdi. Non ci sono palline di altri colori. La probabilità di estrarre a caso una pallina rossa è 1/2; la probabilità di estrarre una pallina che non sia blu è 4/5. Qual è la probabilità di estrarre una pallina verde? Che cosa sai dire sul numero totale di palline contenute nel sacchetto? Secondo te, quale somma dovrebbero scommettere tre giocatori, ciascuno su un colore diverso, per incassare 350 Euro in caso di vincita? Giustifica la risposta.

9. L'estrazione di due carte da un mazzo di 40 carte può essere condotta in due modi differenti. Nel primo, si estrae a caso dal mazzo una carta e poi si estrae una seconda carta senza rimettere nel mazzo la prima. Nel secondo modo, dopo aver estratto la prima carta, questa viene rimessa nel mazzo e quindi ne viene estratta una seconda. Per avere la maggior probabilità di estrarre due carte di cuori, quale modalità di estrazione sceglieresti? Giustifica la risposta.

10. Due sacchetti contengono cinque bussolotti ciascuno, all'interno dei quali è contenuta una lettera dell'alfabeto. Nel sacchetto A vi sono le cinque lettere della parola "pappa" e nel sacchetto B le cinque lettere della parola "posta". Viene estratta a caso una lettera dal sacchetto A e introdotta nel sacchetto B. In seguito viene estratta una lettera dal sacchetto B che viene introdotta in A. Qual è la probabilità che, effettuata l'esperienza a due prove descritta, la composizione dei due sacchetti sia uguale a quella originaria (ossia quella che essi possedevano prima che venisse effettuata l'esperienza a due prove)?

Griglia di correzione

1. 36/121 e 60/121
2. 4/25 e 12/25
3. 2 bianche, 4 rosse, 6 nere
4. 1/9
5. No, le successive estrazioni settimanali al gioco del lotto sono eventi fra loro indipendenti.
6. 1 – (9/20)(8/19)(7/18)
7. 7/8
8. *p(R)* = 1/2 ; *P(V)* = 3/10; *p(B)* = 1/5 . Possiamo però dire che le palline rosse sono il 50% del totale, che le blu sono il 20% e le verdi il 30%. Un giocatore che decidesse di puntare sull'evento "esce la pallina verde" dovrà quindi pagare un premio pari al 30% della somma che vincerebbe. Il 30% di 350 euro dà 105 euro. Il 50% di 350 euro dà 175 euro. Il 20% di 350 euro dà: 70 euro. Ovviamente 70+105+175=350. Si può anche dire che le palline sono almeno 10.
9. È più conveniente la seconda modalità.
10. 1/3

Elementi di prove di verifica

Valutazione delle probabilità

Approccio frequentista

1. In una confezione sono contenute 50 lampadine di cui 5 difettose. Se ne prende una a caso: qual è la probabilità che sia difettosa? E se se ne prendono due, contemporaneamente, qual è la probabilità che siano entrambe difettose?

2. Da una indagine su una popolazione risulta che una persona su 20 è mancina. Qual è la probabilità che una persona, presa a caso in quella popolazione, non sia mancina?

3. Lancia 100 volte un dado a sei facce (puoi effettuare l'esperimento oppure fare una simulazione al computer).
Qual è la frequenza relativa dell'evento *uscita del 6*? Qual è la probabilità dell'evento *uscita del 6* nel lancio di un dado? (Spiega le eventuali discordanze tra i due valori ottenuti).

Approccio classico

1. Un professore di matematica, in una classe composta da 25 alunni, per interrogare decide di lasciarsi guidare dal caso. Ha a disposizione un sacchetto che contiene 30 palline numerate da 1 a 30. Decide di usarlo così come è ed estrae una pallina; se il numero non supera 25, interroga l'alunno che ha quel numero nell'elenco; altrimenti fa il prodotto delle cifre e interroga l'alunno che ha nell'elenco il numero corrispondente. Tutti gli alunni hanno la stessa probabilità di essere interrogati? Il professore interroga in ogni caso qualcuno? Qual è la probabilità che non interroghi nessuno?

2. Tre signori lasciano il loro cappello al guardaroba di un ristorante. Se, all'uscita, riprendono i loro cappelli a caso qual è la probabilità che nessuno riprenda il suo cappello?
(Per rispondere considera che tale evento è quello contrario all'evento unione dei tre eventi: *il primo o il secondo o il terzo signore riprende il suo cappello*).
Se i signori sono quattro la probabilità che nessuno riprenda il suo cappello aumenta o diminuisce? E se sono cinque ?

Approccio soggettivista

1. Considera gli eventi: *la squadra A vince la partita di calcio di domenica* e *la squadra A termina in vantaggio il primo tempo*. Tali eventi sono indipendenti o no? Se no, sono correlati positivamente o negativamente?

2. Supponi che in un *questionario* presentato alla tua classe si chieda di indicare l'anno di nascita di Dante Alighieri scegliendo tra 5 possibili risposte:

- ❑ 753 a.C.
- ❑ 825 d.C.
- ❑ 1265 d.C.
- ❑ 1321 d.C.
- ❑ 1920 d.C.

Per ogni risposta indica qual è, secondo te, la probabilità che i tuoi compagni indichino quello come anno di nascita di Dante Alighieri. Giustifica le tue valutazioni secondo un tuo criterio relativo al livello medio di cultura dei tuoi compagni.

Griglia di correzione

Approccio frequentista

1. 1/10; 4/490
2. 19/20

Approccio classico

1. 1/30
2. 1/3 ; 3/8; 11/30

A proposito di valutazione scolastica

Livello scolare: 2° biennio

Abilità interessate	**Conoscenze**	**Nuclei coinvolti**	**Collegamenti esterni**
Identificare situazioni che richiedono di rilevare lo stesso carattere su una unità statistica formata da 2 elementi, o 2 caratteri diversi sulla stessa unità statistica. Impostare una tabella a doppia entrata; classificare i dati secondo due caratteri e riconoscere in essa i diversi elementi individuabili.	Distribuzione doppia di frequenze e tabella a doppia entrata. Distribuzioni condizionate e marginali.	Dati e previsioni Numeri e Algoritmi Relazioni e funzioni Risolvere e porsi problemi Laboratorio di matematica	Organizzazione di attività

Contesto
Extramatematico, sociale, distribuzioni doppie.

Il contesto è di tipo matematico ed extramatematico ed è connesso all'ambito sociale; nel contesto matematico in particolare riguarda l'ambito statistico (distribuzioni doppie).
Questa attività è consigliata nel 2° biennio, come uno dei possibili approcci iniziali al nucleo. Il contesto è extramatematico e si basa sui dati di ingresso e di uscita dell'insieme degli studenti della classe prima di un istituto superiore e pone gli studenti di fronte all'analisi di un problema concreto che li coinvolge: trovare se esiste una relazione tra i giudizi di ingresso e i giudizi finali.

Descrizione dell'attività
Per affrontare il problema proposto è necessario rilevare per ogni studente il giudizio finale della scuola media (carattere qualitativo ordinato che si esprime con le modalità: Sufficiente, Buono, Distinto, Ottimo) e l'esito finale (carattere qualitativo ordinato che si esprime con le modalità decrescenti: Promosso, Promosso con 1 debito, Promosso con 2 debiti, Promosso con 3 debiti, Respinto). Con questi dati è possibile costruire una distribuzione doppia di frequenze che si presenta come nella Tabella 1: i dati ai quali la tabella si riferisce derivano da una rilevazione effettuata presso L'I.T.I.S. "C. Zuccante" di Mestre (Ve), nell'anno scolastico 2001/2002.

Studenti iscritti nelle classi prime dell'a.s. 2001/02
per giudizio della scuola media ed esito finale (frequenze assolute)

Giudizio scuola media	Esito finale					Totale
	Promosso	Prom1Deb	Prom2Deb	Prom3Deb	Respinto	
Sufficiente	8	8	15	34	41	106
Buono	40	**21**	15	17	13	106
Distinto	64	6	5	5		80
Ottimo	32	1				33
Totale	144	36	35	56	54	325

Tabella 1

<u>Prima fase</u>

L'insegnante stimola e guida gli studenti alla scoperta del significato degli elementi che compaiono nella Tabella 1.
Quante sono le combinazioni possibili delle modalità delle due variabili?
In generale sono tante quanto è il prodotto del numero delle modalità in riga per il numero delle modalità in colonna. C'è qualche relazione con il prodotto cartesiano di due insiemi ?.
Cosa indica il numero 21 in grassetto nella tabella? E' il numero degli studenti che hanno avuto "contemporaneamente" Buono nella scuola media e che sono stati Promossi con 1 debito formativo.
L'insegnante suggerisce di fissare l'attenzione sulla modalità Distinto tra quelle che appartengono al carattere "Giudizio della scuola media". Quanti sono gli studenti che hanno avuto tale valutazione? Quale distribuzione di "Esito finale" hanno conseguito? Si tratta di una distribuzione semplice? Da chi dipende? Ha senso dire che si tratta di una distribuzione semplice dell'"Esito finale" condizionatamente al fatto di avere conseguito Distinto alla scuola media?
L'insegnante conduce gli studenti a scrivere la distribuzione condizionata:

	Promosso	Prom1Deb	Prom2Deb	Prom3Deb	Respinto	Totale
Distinto	64	6	5	5		80

Quante sono le distribuzioni parziali condizionate alle modalità del giudizio della scuola media?
In modo analogo e come rinforzo del concetto in discussione, l'insegnante può opportunamente lavorare anche sulle 5 distribuzioni condizionate di colonna.
Si possono fare altre considerazioni sulla Tabella 1. Cosa si ottiene sommando le righe o le colonne? L'insegnante guida gli studenti a rendersi conto che da una distribuzione congiunta di frequenze relativa a due caratteri si ottengono in modo univoco le due distribuzioni semplici dei due caratteri rispetto ai quali si è classificato congiuntamente.

Distribuzione degli studenti rispetto all'Esito finale:

Promosso	Prom1Deb	Prom2Deb	Prom3Deb	Respinto	Totale
144	36	35	56	54	325

Distribuzione degli studenti rispetto al Giudizio della scuola media:

Sufficiente	106
Buono	106
Distinto	80
Ottimo	33
Totale	325

<u>Seconda fase</u>

Si può costruire una distribuzione doppia di frequenze relative? Quali informazioni fornisce?

L'insegnante guida gli studente ad esprimersi con le percentuali. Qual è la percentuale di studenti che hanno conseguito Distinto alla scuola media e sono stati Promossi in seconda? Qual è la percentuale di coloro che hanno conseguito Sufficiente alla scuola media e sono stati Respinti?
L'insegnante può stimolare la discussione circa l'esistenza di un legame tra il "Giudizio della scuola media" e l'"Esito finale".

Studenti iscritti nelle classi prime dell'a.s. 2001/02
per giudizio della scuola media ed esito finale (percentuali sul totale)

Giudizio scuola media	Esito finale					Totale
	Promosso	Prom1Deb	Prom2Deb	Prom3Deb	Respinto	
Sufficiente	2,5	2,5	4,6	10,5	12,6	32,6
Buono	12,3	6,5	4,6	5,2	4,0	32,6
Distinto	19,7	1,8	1,5	1,5	0	24,6
Ottimo	9,8	0,3	0	0	0	10,2
Totale	44,3	11,1	10,8	17,2	16,6	100,0

Tabella 2

Come si può "vedere" ed, eventualmente, misurare l'intensità di questo legame?

Per capire come si analizza l'esistenza di un eventuale legame (**connessione**) tra due caratteri, l'insegnante propone all'attenzione degli studenti la seguente tabella:

	Promosso	Prom1Deb	Prom2Deb	Prom3Deb	Respinto	Totale
Sufficiente	*P^S*	*1D^S*	*2D^S*	*3D^S*	*R^S*	*S*
Buono	*P^B*	*1D^B*	*2D^B*	*3D^B*	*R^B*	*B*
Distinto	*P^D*	*1D^D*	*2D^D*	*3D^D*	*R^D*	*D*
Ottimo	*P^O*	*1D^O*	*2D^O*	*3D^O*	*R^O*	*O*
Totale	*P*	*1D*	*2D*	*3D*	*R*	*n*

Tabella 3

Ogni riga della tabella precedente rappresenta, in simboli, la distribuzione condizionata, del carattere "Esito finale" rispetto alle diverse modalità del carattere "Giudizio delle scuola media", con riferimento alle frequenze assolute.
I simboli di Tabella 3 fanno riferimento ai corrispondenti valori della Tabella 1.
Per rispondere alla domanda posta, però, occorre fare riferimento ai valori percentuali. Perché? L'insegnante dopo aver fatto osservare che la distribuzione marginale e le condizionate hanno diverse numerosità, ricorda agli studenti che per poter confrontare distribuzioni con diverse numerosità occorre ricorrere alle frequenze relative o a quelle percentuali.
Se le modalità di un carattere non avessero influenza sulla distribuzione dell'altro, tutte le distribuzioni percentuali condizionate di riga o di colonna dovrebbero essere uguali e coincidere con la corrispondente distribuzione marginale.
Ad esempio, se il carattere "Giudizio della scuola media" non avesse influenza sull'"Esito finale" si avrebbero per i "Promossi"le seguenti situazioni:

$$\frac{P \wedge S}{S} = \frac{P \wedge B}{B} = \frac{P \wedge D}{D} = \frac{P \wedge O}{O}$$

Applicando tra queste proporzioni la proprietà del comporre si ottiene ad esempio:

$$\frac{P \wedge S + P \wedge B + P \wedge D + P \wedge O}{S + B + D + O} = \frac{P \wedge D}{D}$$

Il numeratore del primo membro è l'insieme dei "Promossi" mentre il denominatore rappresenta l'insieme degli studenti, quindi:

$$\frac{P}{n} = \frac{P \wedge D}{D}$$

Questo implica che i valori delle frequenze assolute congiunte dovrebbero essere dati da:

$$P \wedge D = \frac{P * D}{n}$$

Questo valore rappresenta la frequenza teorica associata alla coppia di eventi (Distinto, Promosso) nell'ipotesi di indipendenza associativa e viene perciò chiamata frequenza teorica in condizione di indipendenza (l'insegnante suggerisce agli alunni di rivedere la regola per il calcolo della probabilità composta per eventi indipendenti).
Questo ragionamento vale per ogni coppia di modalità e quindi è possibile costruire, eventualmente con l'uso del computer, la tabella di indipendenza in cui ogni cella soddisfa alla condizione precedente; tale tabella si dice anche Tabella di connessione nulla.

Tabella di connessione "nulla" fra giudizio della scuola media ed esito finale

	Promosso	Prom1Deb	Prom2Deb	Prom3Deb	Respinto	Totale
Sufficiente	46,966	11,742	11,415	18,265	17,612	106
Buono	46,966	11,742	11,415	18,265	17,612	106
Distinto	35,446	8,862	8,615	13,785	13,292	80
Ottimo	14,622	3,655	3,554	5,686	5,483	33
Totale	144	36	35	56	54	325

Tabella 4

L'insegnante domanda come sono le distribuzioni marginali della Tabella 1 (tabella osservata) e quelle della Tabella 4 (tabella teorica)? Come mai la Tabella 1 contiene solo numeri interi mentre la Tabella 4 contiene tutti numeri decimali?
Come è possibile valutare la "distanza" tra dati osservati e dati teorici in ipotesi di indipendenza? Dopo aver lasciato dibattere gli studenti su questo problema, l'insegnante propone una possibile quantificazione di tale "distanza" costruendo la differenza in modulo tra le frequenze delle caselle corrispondenti nelle due tabelle.

Tabella delle differenze assolute tra i dati della Tabella 1 e quelli della Tabella 4

	Promosso	Prom1Deb	Prom2Deb	Prom3Deb	Respinto
Sufficiente	**38,97**	3,74	3,58	**15,74**	**23,39**
Buono	6,97	9,26	3,58	1,26	4,61
Distinto	**28,55**	2,86	3,62	8,78	**13,29**
Ottimo	**17,38**	2,66	3,55	5,69	5,48

Tabella 5

I dati della Tabella 5 mettono in luce la presenza di alcune distanze (in grassetto) piuttosto elevate rispetto alle altre. Esse segnalano un allontanamento "consistente" dalla condizione teorica di indipendenza.

L'insegnante guida gli studenti a descrivere i dati della Tabella 5. Nel corso del dibattito si osservano alcuni legami interessanti: gli ottimi promossi sono più dell'atteso; i distinti promossi sono più del previsto e non esistono distinti respinti, pur esistendo la corrispondente frequenza teorica; i sufficienti-promossi sono meno dell'atteso mentre i sufficienti-respinti e i sufficienti-promossi con tre debiti formativi sono più del previsto.

Elementi di prova di verifica

1. Si conosce la distribuzione di cento studenti secondo il sesso e il voto in una prova di matematica:

Sesso	Voto in matematica		Totale
	insufficiente	sufficiente	
Maschi	4	6	10
Femmine	36	54	90
Totale	40	60	100

a) Secondo te, i caratteri "Sesso" e "Voto in matematica" sono connessi? Sì No

b) Data una qualunque distribuzione di studenti secondo il "Sesso" e il "Voto in matematica", si dice che tra i due caratteri in questione non c'è connessione se:

1. Al variare delle modalità del carattere "Sesso" le distribuzioni percentuali secondo il carattere "Voto" sono uguali. Vero Falso
2. Rispetto al carattere "Sesso", la distribuzione percentuale di quanti hanno un voto insufficiente è uguale alla distribuzione percentuale di quanti hanno un voto sufficiente. Vero Falso
3. Tra i maschi, la percentuale di quelli che hanno ottenuto sufficiente, sul totale dei maschi, è uguale alla percentuale di femmine con voto sufficiente, sul totale delle femmine. Vero Falso
4. Rispetto al totale degli alunni, la percentuale dei maschi con voto insufficiente è uguale alla percentuale dei maschi con voto sufficiente. Vero Falso

2. La tabella seguente riporta i risultati di un sondaggio rivolto ad un campione di 400 italiani adulti che riguarda il numero delle volte che si sono fatti visitare dal medico negli ultimi tre mesi:

Fumatore	Numero di visite mediche			Totale
	0 – 1	2 – 4	≥ 5	
Sì	20	60	80	160
No	110	90	40	240

a. In percentuale tra i fumatori quanti sono quelli che si fanno visitare non meno di 5 volte?
b. Il fatto di essere fumatori incide o non incide sulla distribuzione delle visite mediche degli ultimi tre mesi?

3. Il senato degli Stati Uniti nel 1994 era formato da 100 membri. Il 44% era repubblicano e tra i repubblicani il 95% era di sesso maschile. Si sa che le femmine erano in totale il 7% dei membri del senato.

a. Inserisci le frequenze assolute nella tabella a doppia entrata per sesso e partito politico.
b. Il partito politico incide sulla composizione per sesso?

Partito politico	Sesso		Totale
	Maschi	Femmine	
Repubblicani			
Democratici			
Totale			

Griglia di correzione

1.a. Sì
1.b. Vero; Vero, Vero, Falso
2. 50%; sì
3.

Partito politico	Sesso		Totale
	Maschi	Femmine	
Repubblicani	42	2	44
Democratici	51	5	56
Totale	93	7	100

I grafici parlano…

Livello scolare: 2° biennio

Abilità interessate	Conoscenze	Nuclei coinvolti	Collegamenti esterni
Selezionare, produrre ed usare appropriate rappresentazioni grafiche delle distribuzioni doppie.	Principali rappresentazioni grafiche per le distribuzioni doppie rispetto a caratteri di qualsiasi natura.	Dati e previsioni Risolvere e porsi problemi Laboratorio di matematica	

Contesto
Sociale.

Il contesto è di tipo matematico ed extramatematico; in particolare, per il contesto matematico, l'attività riguarda l'ambito statistico (distribuzioni doppie e rappresentazioni grafiche).
L'attività prende spunto da un'indagine effettuata dagli studenti, nell'ambito dell'area di progetto, dell'I.T.I.S. "C. Zuccante" di Mestre (Ve) sui lavori svolti dagli stessi studenti nel periodo estivo. La zona offre possibilità di lavoro estivo in diversi settori, grazie ad un buono sviluppo industriale e commerciale. L'istituto stesso, inoltre, organizza stages di formazione nel periodo estivo.
La Presidenza dell'Istituto avvertiva l'esigenza di esplorare più a fondo il fenomeno, di sapere se e quanti studenti erano interessati, a conoscere i settori produttivi di maggiore interesse. Per tale motivo ha commissionato un'area di progetto ad una classe terza dello stesso Istituto. Gli studenti hanno provveduto a predisporre un opportuno questionario, a farlo compilare, a codificarlo e a costruire il corrispondente data base relativo a 499 studenti (unità statistiche) rispetto ai 13 caratteri rilevati.

Descrizione dell'attività
Questa unità utilizza alcuni dei risultati dell'analisi per mostrare rappresentazioni grafiche particolarmente utili per sintetizzare le tabelle e coglierne il significato. Ciò consente quindi sia avere una idea più immediata e generale dell'andamento di un carattere, sia la possibilità di confrontare distribuzioni di caratteri diversi. Con l'avvertenza, sempre valida nella rappresentazione per immagini, che una stessa distribuzione presentata in un modo o in un altro sembra cambiare profondamente.

<u>Prima fase</u>
Una prima domanda riguarda l'"Età" degli studenti che lavorano, le "Classi" e i "Corsi di appartenenza": elettronica e telecomunicazioni (ET) ed informatica (IS).

L'insegnante mette in evidenza come il modo in cui le frequenze si dispongono segnala l'esistenza di una connessione fra le età e le classi. Qual è l'età minima e l'età massima degli studenti per ogni classe? La distribuzione per età degli "elettronici" e la distribuzione per età degli "informatici" sono analoghe nelle classi corrispondenti?

Studenti che lavorano d’estate per età e classe di appartenenza (dati assoluti)

Classe	Età in anni compiuti						Totale
	16	17	18	19	20	21	
3ET	23	8	3				34
4ET		21	22	6	2		51
5ET			28	10	4		42
3IS	23	9	6	2			40
4IS		26	12	6	2		46
5IS			28	15	2	1	46
Totale	46	64	99	39	10	1	259

Tabella 1

L’insegnante, a questo punto, invita gli studenti a rappresentare graficamente la tabella. (*In questa attività faremo riferimento ai colori anche se, per motivi tipografici, i grafici sono riportati in bianco e nero.)* Ecco l’elaborato degli studenti:

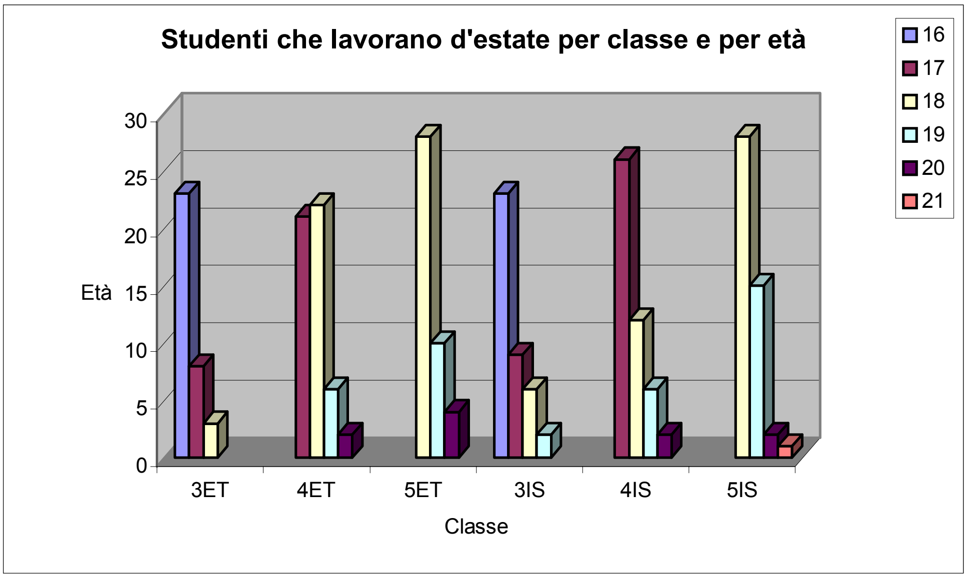


Figura 1

L’insegnante conduce gli studenti ad interpretare il grafico a colonne accostate invitandoli ad interpretare l’andamento delle età al variare delle classi, così come è evidenziato dai colori dei prismi. Un grafico a colonne giustapposte, anziché a prismi giustapposti sarebbe stato più o meno corretto?

<u>Seconda fase</u>

Studenti e insegnanti sono curiosi di sapere se i lavori estivi sono retribuiti e se ciò è condizionato dalla posizione lavorativa. Il data base consente di estrarre la seguente tabella a doppia entrata:

Studenti del triennio che hanno svolto attività
classificati per posizione e per compenso ricevuto (dati assoluti)

Posizione	Compenso		Totale
	Sì	No	
Impiegato	16	1	17
Operaio	67	2	69
Apprendista	109	2	111
Stage	24	5	29
In Proprio	14	2	16
Socio	2		2
Coadiuvante	2		2
Volontario	5	8	13
Totale	239	20	259

Tabella 2

Sono confrontabili fra loro le otto distribuzioni del compenso rispetto alla posizione?
La risposta è ovviamente negativa ed è perciò necessario passare alla tabella con i dati percentuali, nella quale tutte le righe hanno per somma cento:

Studenti del triennio che hanno svolto attività
suddivisi per posizione e per compenso ricevuto (dati percentuali).

Posizione	Compenso		Totale
	Sì	No	
Impiegato	94,1	5,9	100
Operaio	97,1	2,9	100
Apprendista	98,2	1,8	100
Stage	82,8	17,2	100
In Proprio	87,5	12,5	100
Socio	100		100
Coadiuvante	100		100
Volontario	38,5	61,5	100
Totale	92,3	7,7	100

Tabella 3

L'insegnante chiede agli studenti il significato della distribuzione che compare nell'ultima riga. Il valore 92,3 come si giustifica? Dopo aver condotto ad osservare che esso è compreso tra 82,8 e 100, l'insegnante fa riflettere gli studenti sull'uguaglianza:

$$\frac{0{,}941\cdot 17+0{,}971\cdot 69+0{,}982\cdot 111+0{,}828\cdot 29+0{,}875\cdot 16+1\cdot 2+1\cdot 2+0{,}385\cdot 13}{259}=\frac{239{,}015}{259}\cong 0{,}923$$

Ritornando al problema delle rappresentazioni grafiche, qual è il grafico più opportuno per rappresentare contemporaneamente tutte le distribuzioni condizionate del "Compenso" rispetto alla "Posizione"? Ha influenza il fatto che entrambi i caratteri siano qualitativi?

L'insegnante conduce gli studenti a costruire il grafico della Figura 2.
Come ha influenzato il grafico il fatto che tutte le distribuzioni condizionate danno per somma cento?
Come è stato possibile evidenziare la percentuale di studenti che ha ricevuto un compenso condizionatamente alla posizione lavorativa?
L'insegnante guida gli studenti verso le risposte: dal grafico è immediatamente visibile il fatto che il totale è cento e il doppio colore permette di visualizzare le differenti percentuali di coloro che hanno o non hanno ricevuto un compenso.
Naturalmente non era necessario ricorrere ad un grafico tridimensionale che tuttavia è quello scelto dagli studenti.

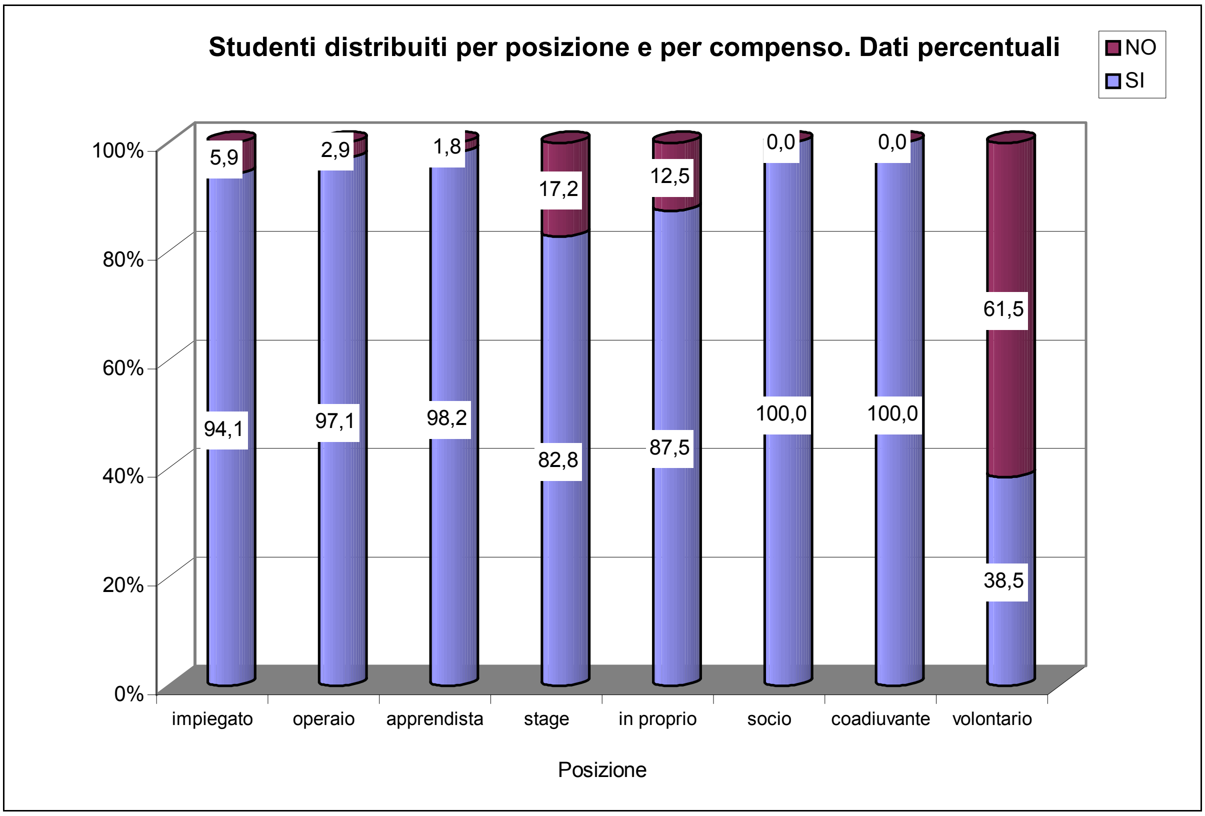


Figura 2

Come si poteva prevedere la percentuale più alta di studenti senza compenso ha operato nel volontariato, mentre chi lavora come socio o come coadiuvante riceve in ogni caso una retribuzione. Attenzione a trarre conclusioni affrettate: in ciascuna di queste due posizioni hanno operato soltanto due studenti!

<u>Terza fase</u>

L'attività proposta costituisce un approfondimento opzionale seppur molto suggestivo per l'interpretazione delle relazioni interne ad una distribuzione doppia. Essa fa perno sui concetti di primo quartile, mediana e terzo quartile di una distribuzione di un carattere almeno ordinato rettilineo.
Una sintesi grafica della tabella doppia, quando almeno uno dei caratteri è ordinato, è fornita dall'insieme di più diagrammi a scatola ("Boxplot"). Nel "boxplot" l'escursione del carattere si valuta in ordinata. Il primo ed il terzo quartile sono i valori in ordinata che delimitano la scatola stessa, mentre la mediana è l'altezza relativa al simbolo corrispondente interno alla scatola. A

partire da ciascun lato orizzontale della scatola sino all'ultima osservazione di coda (valore minimo, valore massimo) si disegna poi un segmento verticale.
Quando, come in una distribuzione doppia con almeno un carattere ordinale, si confrontano giustapponendoli più "boxplot" tutti relativi a tale carattere, in ascissa si pone un asse categorico che serve per spaziare i grafici corrispondenti alle diverse modalità dell'altro carattere.
Per rispondere alla domanda se il tempo lavorativo in giorni si differenzia a seconda del settore d'impiego, una soluzione potrebbe essere quella di costruire un insieme di "boxplot" utilizzando i dati riportati nella Tabella 4. In essa sono indicati, per ciascuno dei settori più importanti, il minimo (min), il primo quartile (q1), la mediana, il terzo quartile (q3) ed il massimo (max).

Alcuni valori caratteristici per principali settori d'impiego

Valori caratteristici	Settore			
	elettronico	manifatturiero	ristorazione	informatico
min	14	15	2	14
q1	30	40	42	15
mediana	59	55	70	30
q3	70	60	90	53
max	120	170	180	90

Tabella 4

Il "boxplot" della durata in giorni del "Rapporto di lavoro" condizionato al "Settore d'impiego" è riportato nella Figura 3.

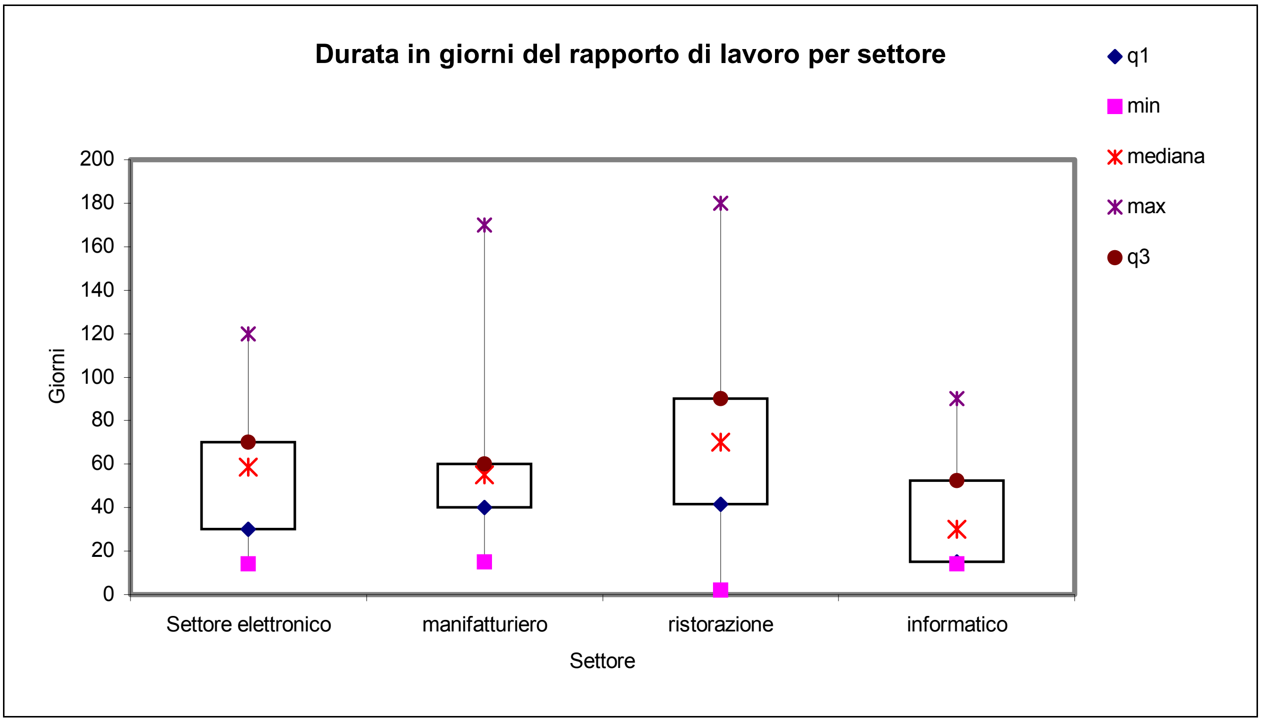


Figura 3

Qual è il settore in cui si manifesta il minor campo di variazione rispetto alla durata del rapporto di lavoro? E quello con maggiore variabilità? Può incidere il contesto economico del territorio ad alta densità turistica e manifatturiera?
Poiché all'interno della scatola cade circa il 50% del collettivo, l'insegnante conduce gli studenti ad osservare che, per esempio, il 50% circa degli studenti del settore manifatturiero ha "resistito" fra 40 e 60 giorni; nel settore elettronico la durata del rapporto di lavoro del 50% circa degli studenti che occupano la parte centrale della distribuzione ordinata dei giorni di lavoro è compresa tra 30 e 70 giorni.
Cosa succede negli altri settori? Nel settore informatico, qual è il numero massimo dei giorni di lavoro del 75% degli studenti che lavorano di meno? L'insegnante ricorda il significato del terzo quartile e porta gli studenti alla risposta.

Promossi con una domanda sola?

Livello scolare: 2° biennio

Abilità interessate	Conoscenze	Nuclei coinvolti	Collegamenti esterni
Utilizzare la Formula di Bayes.	Eventi e operazioni con gli eventi. Eventi incompatibili; eventi esaustivi. Significato della probabilità e sue valutazioni. Formula di Bayes e suo significato.	<u>Dati e previsioni</u> Relazioni e funzioni Risolvere e porsi problemi	Lingua italiana

Contesto
Extramatematico, sociale, probabilità.

Il contesto è di tipo matematico ed extramatematico; in particolare, per il contesto matematico, si pone in evidenza l'ambito probabilistico.
Il problema può essere affrontato a livello di primo biennio quando si conosca la definizione di probabilità e si sia in grado di valutare la probabilità di eventi condizionati. Si vuole portare gli studenti a rendersi conto che l'esser "fortunati" o meno (ad esempio sapendo rispondere in modo esatto ad un compito del tipo quiz a risposta multipla) può essere un fatto quantificato. Nel suo futuro, nella ricerca di sbocchi occupazionali, ciascuno studente potrebbe trovarsi di fronte a strumenti di questo tipo, largamente usati dalle industrie dei giorni nostri nella ricerca-scelta di personale da inserire nelle proprie attività produttive. E quei quiz saranno molto più ardui, anche perché spazieranno ben oltre la pratica scolastica…

Descrizione dell'attività
Si propone in classe un quiz a risposta multipla e si chiede: qual è la probabilità che uno studente che ha risposto esattamente alle domande abbia effettivamente studiato ?
Prendendo spunto da questo fatto si può poi rispondere ad una domanda che può emergere spesso in ambito scolastico.
Si supponga che uno studente abbia, magari a fine anno, una media bassa. E' "possibile", con una sola interrogazione, recuperare il terreno perduto ed essere promosso?
Il problema si presenta molto accattivante e sicuramente stimola la fantasia e la scaltrezza degli allievi molto interessati a questo tipo di argomento.

<u>Prima fase</u>
Si tratta di lavorare con probabilità di eventi condizionati. È stato considerato dalla classe un solo risultato, anche se importante, concernente tale tematica ovvero la Formula di Bayes.
Per questo motivo è necessario fare, con la classe, alcune ipotesi di lavoro; senza ovviamente farne una trattazione teorica, che potrebbe esulare dal contesto della realtà scolastica, si introdurrà operativamente la necessità delle probabilità "a priori"; questo fatto contribuirà non poco ad indirizzare l'attenzione degli studenti sul fatto che trattare l'incertezza non è mai un fatto del tutto "automatico" come invece potrebbe apparire da uno studio superficiale di altre parti della matematica.
Valutare la probabilità che uno studente che ha studiato sappia rispondere alle domande è la prima ipotesi di lavoro da affrontare.
La risposta a tale questione può (e deve!) sollevare discussioni … al termine delle quali potrebbe esserci un accordo del tipo di quello presentato qui di seguito.

Di uno studente si potrebbe arrivare a dire che:

a. Se studia con impegno, allora risponde bene alle domande del quiz con probabilità 1 (ovvero per lui saper rispondere correttamente è un evento certo).
b. Se studia così così, sa rispondere bene alle domande del quiz con probabilità *p* (che lo studente stesso, più di ogni altro, sa quantificare !).
c. Se non ha studiato, può rispondere bene ad una domanda con *m* risposte con probabilità *1/m* (ovvero per lui tutte le risposte sono "equivalenti": da qui si vede che l'equiprobabilità non è un valore intrinseco ma una valutazione razionale successiva alla "raccolta" di (tutte) le informazioni a disposizione.)
d. Se uno studente a fine anno ha la media del 4 vuol dire che, durante l'anno, su 10 domande ha risposto bene solo a 4 (oppure a 20 su 50 e così via). Anche su questo punto la valutazione potrebbe essere diversa: ad esempio, se l'insegnante è uso non dare più di 8 ad un compito o ad una interrogazione perfetta, la valutazione precedente va riscalata a….. 3.2 (cioè la media del 4 diventa praticamente media del 3). Nella discussione lo studente impara a riflettere sul fatto che, cambiando le informazioni, può anche cambiare la probabilità: d'altronde nessuno di loro giocherebbe mai con monete che sa, prima di iniziare a giocare, essere "non regolari".

Una volta d'accordo (!) con queste, necessarie, considerazioni – e ricordandosi ancora una volta che la discussione fatta su di esse ha una importante valenza concettuale, oltre che didattica e operativa – si può passare alla formalizzazione del problema.
Indichiamo con *A* l'evento "lo studente ha (davvero) studiato" e con *B* l'evento "lo studente risponde esattamente alla domanda".
Chiedersi qual è la probabilità che uno studente che ha risposto esattamente alle domande abbia effettivamente studiato vuol dire chiedersi, formalizzando la richiesta, qual è la probabilità dell'evento *A/B*.
Dalla sopra ricordata Formula di Bayes si ha:

$$P(A/B) = \frac{P(B/A)P(A)}{P(B/A)P(A) + P(B/\overline{A})P(\overline{A})}$$

dove con $\overline{A}$ si è indicato l'evento "complementare" dell'evento *A* ovvero l'evento che si verifica quando *non* si verifica l'evento *A*.
Calcoliamo adesso i vari valori di probabilità, supponendo, per semplicità, che il compito consista di una sola domanda. (L'estensione a più domande sarà poi un fatto successivo, anche se non completamente automatico).

$P(B/A) = 1$ (da quanto si è argomentato al punto a)

$P(A) = p$ (anche questo valore è stato quantificato in seguito alla discussione sviluppata al punto d)

$P(B/\overline{A}) = \frac{1}{m}$ (se, come si è detto al punto c, nel quiz ci sono *m* risposte)

Dunque si ha:

$$P(A/B) = \frac{1 \times p}{1 \times p + \frac{1-p}{m}} = \frac{mp}{1 + p(m-1)}$$

Nota: Si potrebbe naturalmente eccepire sulla decisione di porre $\frac{1}{m}$, come valore equiprobabile.

Infatti in molte domande a risposta multipla, spesso anche chi non ha studiato si accorge che le risposte non sono equivalenti. Molti preparatori di test infatti seguono la "regola" di due risposte molto simili, una abbastanza verosimile ed una assai lontana. Anche in questo caso comunque si tratta di dare valutazioni di probabilità in senso razionale-soggettivo, tenendo conto dell'esperienza (ovvero dello scopo del test, del preparatore stesso, ecc…).

Seconda fase

L'efficacia concreta di questa formula può essere vista subito rispondendo alla seconda questione che ci si è posti all'inizio.

Si supponga che uno studente abbia, magari a fine anno, una media bassa. E' "possibile", con una sola interrogazione recuperare il terreno perduto ed essere promosso?

In questo contesto "possibile" assume il significato di "ragionevole" se riferito ad una eventualità di questo genere; ci chiediamo cioè la liceità di fare questa proposta – sia da parte del professore che magari vuol aiutare lo studente, sia da parte dello studente che cerca, con un solo gettone, di far saltare il banco. Più concretamente ci stiamo chiedendo se ci sono dei margini entro i quali tale proposta può essere fatta senza eccessive indignazioni (da parte degli altri studenti, da parte degli altri professori, ecc..). E' peraltro evidente che se la proposta viene fatta a (o da) uno studente che ha media 5, nessuno si scandalizzerebbe, ma se lo studente avesse media 2…!

Sia, ad esempio, 4 il corrispondente di "media bassa"

Per le considerazioni successive alla discussione fatta in classe, possiamo dire che, per quanto ne sappiamo, il nostro studente è come se, durante l'anno, avesse "mediamente" risposto bene solo a 4 su 10 (oppure a 20 su 50 e così via).

Dobbiamo ancora fare una ulteriore osservazione di lavoro: rispondere bene ad una interrogazione, anche se finale, dovrà essere considerato come rispondere bene ad un quiz con due sole risposte. Porremo per questo $m = 2$.

Siamo dunque in un caso particolare della nostra formula finale: sarà sufficiente porre $p = 4/10$ per ottenere:

$$P(promozione) = \frac{4}{7}$$

Essendo $\frac{4}{7}$ circa 0,571 possiamo dire che siamo vicini alla sufficienza e dunque sarebbe lecito e corretto promuovere lo studente in quelle condizioni

Naturalmente ciò non sarebbe più vero nel caso di uno studente con "media bassa" uguale a 3, in quanto gli stessi calcoli ci portano ad una conclusione del tipo $P(promozione) = \frac{6}{13}$ e dunque non sarebbe più considerato ammissibile fare un simile "salto".

Nota: con questa attività lo studente, oltre a familiarizzarsi con il concetto di probabilità condizionata, inizia a decidere di dare proprie (e coerenti, e razionali, ed eque,….) valutazioni di alcuni eventi che ha davanti ai suoi occhi: inizia cioè a familiarizzarsi, magari anche senza conoscerne il nome, con le probabilità "a priori" che così tanta importanza rivestono nelle considerazioni statistiche e probabilistiche, e comprende che sono le "a priori" a condizionare il risultato finale, nel caso trattato la promozione o la bocciatura.

Nella discussione che apre la prima fase, è opportuno far emergere che nella valutazione della probabilità a priori dell'evento *A:* "lo studente ha (davvero) studiato", certamente porre $P(A) = 0,4$, la probabilità del voto 4, gioca un ruolo ambiguo ma solo "prima della discussione" e della conseguente assegnazione di probabilità. Si tratta infatti di una prima ipotesi di lavoro sulla quale si può discutere (in classe) per arrivare ad una formulazione che, dipendendo magari dal professore,

dal periodo dell'anno, dall'esame (orale o scritto o entrambi) può anche "cambiare". Questa, al di là della formalizzazione, è una consapevolezza acquisita importante in sé.
Nella discussione che chiude la seconda fase, poi, trovata la probabilità *"a posteriori",* l'insegnante porterà gli studenti a comprendere che il modo in cui ci si è espressi circa la *P(promozione*) fa passare implicitamente dalla probabilità al suo uso. Si può dire che il docente usa un criterio secondo il quale se la probabilità a posteriori è "di poco" inferiore a 0,6 allora lo studente viene promosso? E' argomentabile che, utilizzando il voto 4 come mezzo per assegnare il valore 0,4 alla probabilità di "Aver studiato da 4", allora avendo ottenuto 0,571 è come dire che lo studente ha studiato da 5,71, ossia da quasi 6 e, conseguentemente, la promozione viene data?

Elementi di prove di verifica

1. Siano dati gli eventi:
A = "essere il capocannoniere della Serie A di calcio"
B = "essere un uomo che abita a Milano"

Dopo avere valutato la probabilità di *A* e di *B* , calcolare *P(A/B)*, *P(B/A)* e fare commenti sul risultato.

Nota: Tale calcolo, oltre che far riflettere sul significato di evento condizionato e della relativa valutazione probabilistica, serve assai a mettere in guardia lo studente dal non confondere i due eventi *A/B* e *B/A*. Nel febbraio 2003 il capocannoniere della serie A era Vieri dell'Inter di Milano.

Se si insiste…*non* si vince

Livello scolare: 2° biennio

Abilità interessate	**Conoscenze**	**Nuclei coinvolti**	**Collegamenti esterni**
Valutare la probabilità in diversi contesti problematici. Distinguere tra eventi indipendenti e non. Valutare criticamente le informazioni fornite dai media relative ai giochi di sorte.	Significato della probabilità e sue valutazioni. Semplici distribuzioni di probabilità. Il concetto di gioco equo.	Dati e previsioni Risolvere e porsi problemi Laboratorio di matematica	Giochi Previsioni

Contesto
Giochi, probabilità.

Questa attività può essere introdotta nel secondo biennio, dopo aver introdotto le definizioni di probabilità e dopo aver trattato i principali valori medi e le principali misure di dispersione.
Per la simulazione al computer sono necessari alcuni prerequisiti di conoscenza del foglio elettronico: come si inseriscono i dati, come si inserisce una formula, come si copia una formula, riferimenti relativi e assoluti alle celle, come si crea un grafico. Le funzioni "Casuale( )" e "Se( )" possono essere introdotte anche in questo contesto.

Descrizione dell'attività
L'attività costituisce un primo approccio alla distribuzione binomiale e ha come obiettivo quello di far capire, fugando una credenza diffusa, che è negativa la risposta alla domanda: lanciando molte volte una moneta (bilanciata) il numero delle teste tende ad essere uguale al numero delle croci?
Per risolvere il problema conviene partire da semplici considerazioni.
Se si lancia una moneta si può ottenere testa o croce con probabilità rispettivamente $p = \frac{1}{2}$ e $q = \frac{1}{2}$.
Quale sarà la frequenza relativa di testa in 100 lanci di una moneta?
Per rispondere l'insegnante propone di simulare l'esperimento con il foglio elettronico usando la funzione "Casuale( )".
Questa funzione genera un numero compreso tra zero e uno. Si può associare ad un numero minore di 0,5 l'uscita di Testa e ad un numero maggiore o uguale di 0,5 l'uscita di Croce. Oppure si moltiplica la funzione "Casuale( )" per 2 (così si ottengono numeri compresi tra 0 e 2) e poi se ne prende la parte intera (funzione "Int( )"). Si ottiene 1 (Testa) o 0 (Croce).
Il foglio elettronico, di cui la Figura 1 mostra la parte iniziale, è stato costruito riportando: nella colonna D le sequenze di 0 e 1 ottenute tramite la funzione "Casuale( )", calcolata nella colonna C; nella colonna A il numero progressivo di lanci; nella B gli esiti dell'esperimento simulato, ossia le uscite di Testa (1) le uscite di Croce (0). Per ottenere la sequenza di T e C si usa la funzione "Se( )" che ha la seguente sintassi "=SE(condizione; azione nel caso sia vera la condizione; azione nel caso la condizione sia falsa)". In questo caso è stata digitata la formula "=SE(D2 = 1;"T";"C")". Il calcolo per il numero di teste (colonna E) si può fare in maniera agevole incrementando ogni volta il contenuto della cella in cui si trova la somma delle Teste e usando la funzione "Copia formula" e i riferimenti relativi alle celle. In questo caso si è inserita nella prima cella della colonna E (E2) il contenuto della prima cella della colonna D (D2), poi è stata digitata la formula E3 = E2+D3 per la cella successiva e infine si è copiato il comando nelle altre celle attive.

Nella colonna F sono riportate le frequenze relative di testa.

B7 = =SE(D7=1;"T";"C")

	A	B	C	D	E	F
1	**LANCI**	**MONETA**			n° **T**	FREQ.**T**
2	1	C	0,82011	0	0	0,00
3	2	C	0,543162	0	0	0,00
4	3	T	1,665016	1	1	0,33
5	4	C	0,085047	0	1	0,25
6	5	T	1,204324	1	2	0,40
7	6	T	1,876874	1	3	0,50
8	7	T	1,06786	1	4	0,57
9	8	T	1,662015	1	5	0,63
10	9	T	1,38247	1	6	0,67
11	10	C	0,990017	0	6	0,60
12	11	T	1,256956	1	7	0,64
13	12	C	0,14895	0	7	0,58
14	13	C	0,848766	0	7	0,54
15	14	C	0,09596	0	7	0,50
16	15	T	1,888573	1	8	0,53
17	16	T	1,354797	1	9	0,56
18	17	T	1,464304	1	10	0,59

Schema di Testa e Croce

Figura 1

Può essere interessante far ripetere più volte l'esperimento (basta premere un tasto!) e osservare come variano i grafici.
Si nota facilmente nella Figura 2 che, al crescere del numero dei lanci, la frequenza relativa oscilla, arrivando a stabilizzarsi attorno a 0,5; dunque la frequenza relativa "tende" alla probabilità, cioè a 0,5 (Legge dei grandi numeri).

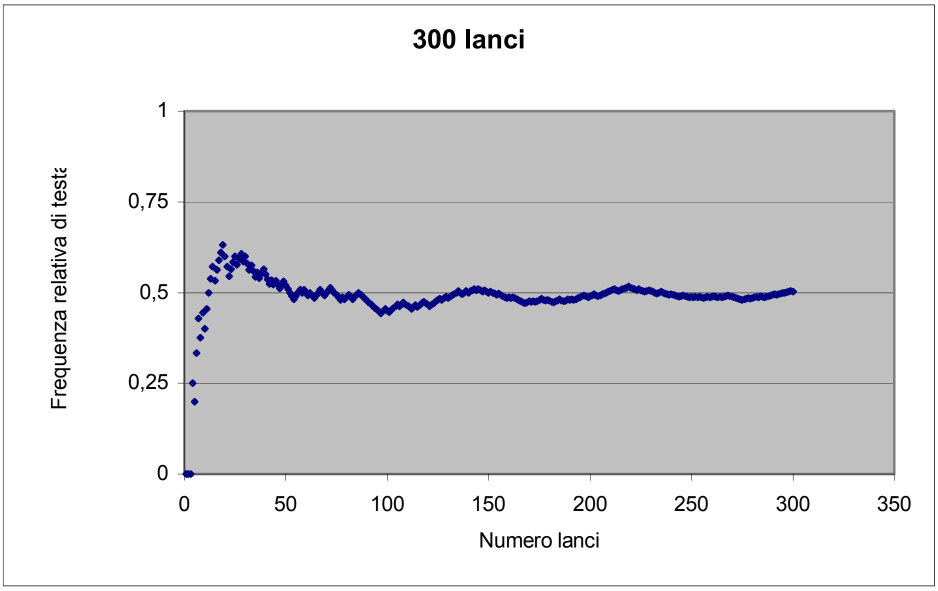


Figura 2

L'insegnante pone allora la seguente domanda: in un gran numero di lanci ci si può attendere che il numero delle teste sia uguale al numero delle croci?
In altre parole ci si può attendere che la differenza tra il numero delle teste e il numero delle croci tenda a zero al crescere del numero dei lanci?

Per aiutare gli studenti a rispondere, l'insegnante chiede se è importante sapere quale sia maggiore fra il numero delle teste e quello delle croci. Quando dal dibattito emerge che ciò che interessa è il modulo della differenza, l'insegnante propone di usare la simulazione con il foglio elettronico e di aggiungere un'altra colonna (colonna G) in cui si pone la differenza in valore assoluto tra il numero delle teste e il numero delle croci:

$$|T\text{-}C| = |\ T - (n\text{-}T)\ | = |\ 2T\text{-}n\ |$$

Nel foglio la formula è "=Ass(2*E2-A2)" che viene, come al solito, copiata in tutte le celle attive. L'insegnante invita gli studenti a costruire il grafico corrispondente.

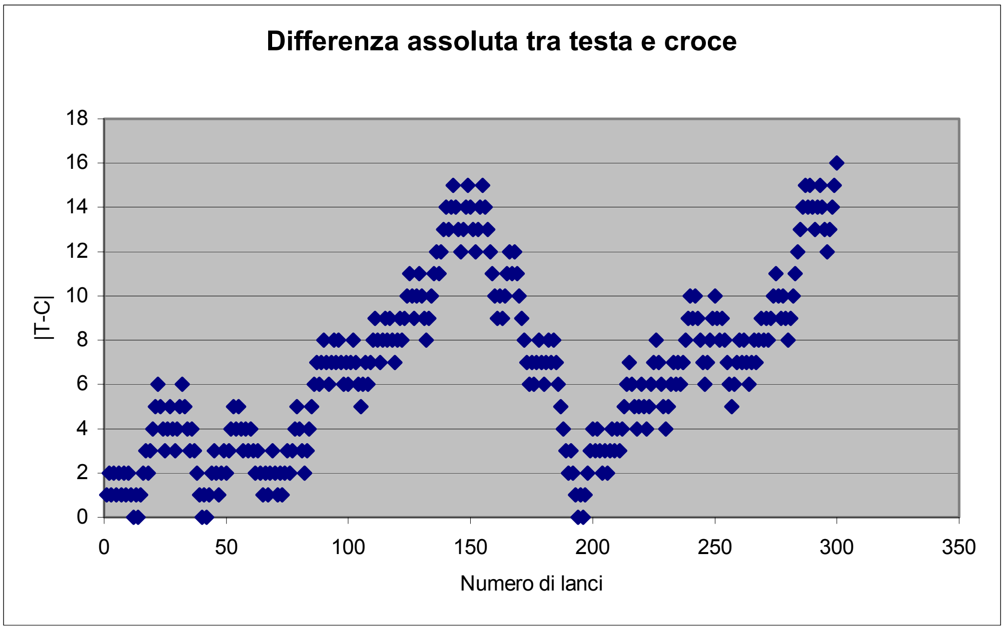


Figura 3

Il grafico in Figura 3 mostra che la differenza tra le teste e le croci tende ad aumentare.
Può essere interessante far ripetere più volte l'esperimento (basta premere un tasto!) e si osserva che la differenza non solo non si stabilizza attorno allo zero, ma anzi aumenta…
Perché?
Se gli studenti conoscono già la media e lo scarto quadratico medio per una distribuzione statistica, si possono introdurre questi concetti per la distribuzione di probabilità nel caso di Testa e Croce (distribuzione binomiale).

Nota: L'insegnante fa notare la caratteristica del foglio Excel di rappresentare le frazioni improprie, separando la parte intera dalla frazione propria. Così, ad esempio, 1 1/2 = 3/2.

Partendo dallo schema:

		C **T**		
2 monete	**CC**	**CT**	**TT**	
n° teste	0	1	2	
Probabilità	1/4	1/2	1/4	media = ½ ·**2**
3 monete	**CCC**	**CCT** **CTT**	**TTT**	
n° teste	0	1 2	3	
	1/8	3/8 3/8	1/8	media = ½ ·**3**
4 monete	**CCCC** **CCCT**	**CCTT**	**CTTT** **TTTT**	
n° teste	0 1	2	3 4	
Probabilità	1/16 4/16	6/16	4/16 1/16	media = ½ ·**4**

si arriva per gradi alle formule:

la media del numero di Teste in n lanci è $\frac{n}{2}$, la varianza è $\frac{n}{4}$ e $\sigma = \frac{\sqrt{n}}{2}$.

L'insegnante può guidare la classe ad approfondire ulteriormente la comprensione di come n e p (parametri) "governano" la distribuzione binomiale, proponendo, ad esempio, lo schema seguente.

2 monete	n. teste (xi)	Prob. P(xi)	xi*P(xi)	xi^2	xi^2*P(xi)
CC	0	1/4	0	0	0
CT	1	1/2	1/2	1	1/2
TT	2	1/4	1/2	4	1
			1		1 1/2

n	2
p	1/2
q	1/2

Media	1

varianza	1/2

media 1 = n*p
varianza 1/2 = n*p*q

3 monete	n. teste (xi)	Prob. P(xi)	xi*P(xi)	xi^2	xi^2*P(xi)
CCC	0	1/8	0	0	0
CCT	1	3/8	3/8	1	3/8
CTT	2	3/8	3/4	4	1 1/2
TTT	3	1/8	3/8	9	1 1/8
			1 1/2		3

n	3
p	1/2
q	1/2

media	1 1/2

varianza	3/4

media 1 1/2 = n*p
varianza 3/4 = n*p*q

4 monete	n. teste (xi)	Prob. P(xi)	xi*P(xi)	xi^2	xi^2*P(xi)
CCCC	0	1/16	0	0	0
CCCT	1	4/16	4/16	1	4/16
CCTT	2	6/16	12/16	4	1 8/16
CTTT	3	4/16	12/16	9	2 4/16
TTTT	4	1/16	4/16	16	1
			2		5

n	4
p	1/2
q	1/2

media	2

varianza	1

media 2 = n*p
varianza 1 =n*p*q

Tornando al problema del numero aleatorio |*2T-n*| va tenuto presente che la varianza del numero aleatorio (*2T-n*) è $4\cdot\frac{n}{4}$ e lo scarto quadratico medio è $\sigma = \sqrt{n}$. Quest'ultima quantità è una misura per eccesso della media del numero aleatorio |*2T-n*| (cfr.: Barra, 2000).
Dunque, contrariamente a quello che si poteva pensare, la differenza in modulo tra il numero delle teste e il numero delle croci non tende a zero ma cresce come $\sqrt{n}$, perciò, per esempio, su 100 lanci ci si può aspettare in media che tale differenza assuma il valore della radice di 100, cioè 10; così su 400 lanci possiamo aspettarci in media che tale differenza assuma il valore 20. Questo dipende dal fatto che stiamo considerando le frequenze assolute e non quelle relative.
Se non si ritiene opportuno introdurre i concetti di media e scarto quadratico per questa distribuzione si può dare una spiegazione intuitiva del fatto che dopo un eccesso di teste non è necessario un recupero delle croci; si tratta infatti di eventi indipendenti, la moneta non ha memoria!
Se si gioca a Testa e Croce con una scommessa alla pari il gioco è equo, ma ciò non significa che dopo 100 partite si possa essere certi di non aver perso né guadagnato nulla; infatti la probabilità che su 100 lanci vi siano esattamente 50 Teste e 50 Croci è:

$$p_{50} = \binom{100}{50} \cdot \frac{1}{2^{100}} \cong 0{,}08$$

cioè un numero molto piccolo! In caso generale di fronte a $n = 2m$ prove si otterrà sempre un risultato del tipo $\frac{1}{\sqrt{\pi m}}$ [1] e dunque più prove si fanno e minore è la probabilità di avere metà successi e metà fallimenti.
Può essere questa l'opportunità di introdurre il concetto di gioco equo.

Da questa attività dovrebbe emergere il concetto espresso dalla seguente frase:

"*Giocare poche volte a Testa e Croce non è più ma meno rischioso che giocare molte volte…La legge dei grandi numeri non giustifica alcuna speranza che chi è in perdita debba "rifarsi". L'illusorietà e perniciosità di tale fiducia nel"rifarsi"appare sancita anche in una battuta popolare (sembra siciliana), notevole perché in generale le preferenze popolari sembra vadano alla tesi sbagliata. Si tratta della risposta di una donna ad un'amica che le aveva chiesto se era vero che suo figlio aveva perduto una forte somma al gioco: <<Sì, ma questo è niente: il peggio è che vuole rifarsi>>(cfr.:* de Finetti, 1970, vol. 1, pag.384).

[1] Tale risultato si ottiene considerando che, per l'approssimazione di De Moivre – Stirling, è circa $n! = n^n e^{-n} \sqrt{2\pi n}$

Anche le rette raccontano

Livello scolare: 2° biennio

Abilità interessate	**Conoscenze**	**Nuclei coinvolti**	**Collegamenti esterni**
Identificare situazioni che richiedono di rilevare lo stesso carattere su una unità statistica formata da due elementi, o 2 caratteri diversi sulla stessa unità statistica.	Concetto e significato di modello: correlazione e regressione.	<u>Dati e previsioni</u> Spazio e figure Argomentare, congetturare, dimostrare Risolvere e porsi problemi Laboratorio di matematica	Organizzazione delle attività

Contesto
Extramatematico, sociale.

Il contesto è di tipo matematico ed extramatematico; in particolare, per il contesto matematico, si pone in evidenza l'ambito statistico (rappresentazioni grafiche; correlazione e regressione).
L'attività riguarda lo studio di una distribuzione doppia rispetto a due caratteri entrambi quantitativi e può essere introdotta, nella forma che qui viene proposta, nel secondo biennio quando gli studenti hanno acquisito sicurezza sulle principali abilità relative all'uso del foglio elettronico. Si richiede l'uso dei filtri in una tabella, la padronanza delle diverse tipologie di indirizzamento e l'uso di alcune funzioni statistiche. Se non si dispone del foglio di calcolo, le proposte indicate di seguito sono ugualmente valide, ma richiedono più tempo per l'elaborazione.

Descrizione dell'attività
L'esempio presenta l'esito finale di una classe prima in tre materie. La tabella è stata tratta dal database dell'unità "A proposito di valutazione scolastica" con l' uso del filtro appropriato.

Alunno	Matematica	Fisica	Inglese	Alunno	Matematica	Fisica	Inglese
	x	y	z		x	y	z
1	5	4	5	14	8	8	8
2	5	6	4	15	9	8	8
3	4	5	4	16	3	4	6
4	6	6	5	17	5	6	6
5	8	8	8	18	6	7	8
6	6	6	6	19	6	6	8
7	6	6	6	20	6	7	8
8	7	8	7	21	6	6	8
9	9	8	7	22	6	6	6
10	6	6	7	23	5	6	7
11	8	7	7	24	6	7	6
12	5	5	6	25	7	8	6
13	8	8	8	26	5	5	5

Tabella 1

Prima fase

L'insegnante propone all'osservazione della classe la Tabella 1 e formula alcune domande. Chi è bravo in matematica lo è anche in fisica o in inglese? Come si può stabilire se uno studente è bravo in una determinata materia? In questa classe chi ha 7 in matematica è "bravo"?
L'insegnante propone la Figura 1 e stimola gli studenti a riflettere sul grafico in essa contenuto.

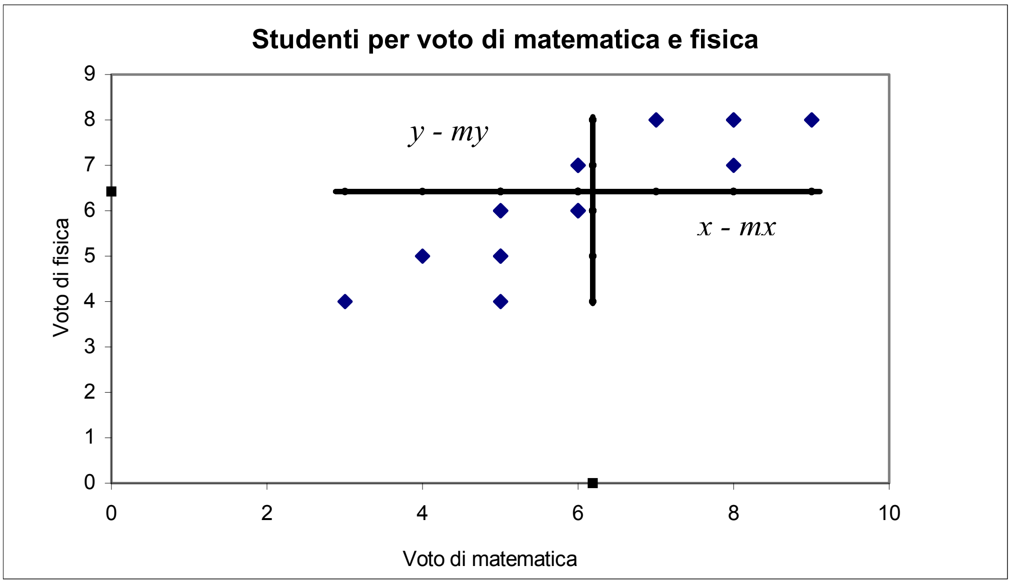


Figura 1

Come mai i punti sono meno di 26 (numero degli studenti)? La dislocazione dei punti sul piano cartesiano dà una informazione su un eventuale legame tra il voto di matematica (x) e il voto di fisica (y)? Perché sono stati rappresentati gli scarti dalla media aritmetica ($x - mx$) e ($y - my$)? Che informazioni aggiuntive offrono?

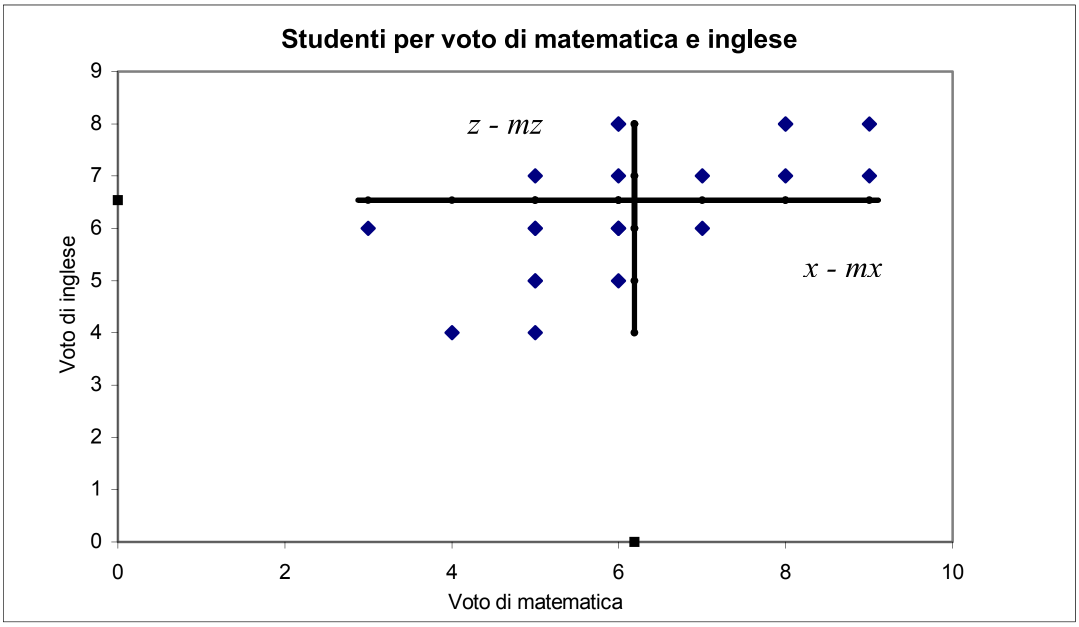


Figura 2

Successivamente l'insegnante stimola gli studenti a riflettere sul grafico di Figura 1.
Gli studenti più bravi in matematica sono più bravi anche in fisica? E' vero anche il viceversa? Ci sono studenti più bravi in matematica, ma meno bravi in fisica? E' vero anche il viceversa? Rispetto al sistema traslato nel baricentro [($x-mx$); ($y-my$)] quali segni hanno le coordinate di queste unità statistiche? L'insegnante sottolinea che a questo punto è possibile sfruttare le note relazione tra i segni di due numeri relativi per cui il prodotto tra due concordi è positivo e il prodotto tra due numeri discordi è negativo.
Rispetto al sistema degli scarti si può dunque dire che se i punti della nuvola si dispongono con maggiore frequenza nel primo e nel terzo quadrante fra le due variabili vi è "concordanza", mentre in caso contrario vi è "discordanza". In Figura 1 è evidente la concordanza, ossia al crescere del voto in matematica cresce mediamente anche il voto in fisica. A queste prime valutazioni soggettive è tuttavia indispensabile sostituire una misura oggettiva, che si fonda sul prodotto degli scarti che competono a ciascuna unità.
Per potere confrontare misure ottenute su distribuzioni diverse è generalmente indispensabile che esse varino fra 0 e 1. In questo particolare ambito concettuale però è necessario che si tenga conto che la correlazione ha un doppio aspetto: concordanza o discordanza. Questa duplicità si esprime con i segni: + esprime la concordanza, - esprime la discordanza, dunque ciò che si vuole è un indice che assuma valori fra -1 e $+1$. Se assume valori positivi è un indice di concordanza, se assume valori negativi è un indice di discordanza. Il valore 0 è indicatore di una situazione da esaminare con attenzione poiché si può realizzare o in ipotesi di indipendenza, connessione nulla, tra i caratteri oppure quando i punti, nel loro insieme, non "danno l'idea" di un andamento lineare.
Il "coefficiente di correlazione lineare di Bravais-Pearson" è la soluzione a questo problema concettuale. Esso ha la seguente espressione:

$$r=\frac{\sum_{i=1}^{n}(x_i-mx)(y_i-my)}{\sqrt{\sum_{i=1}^{n}(x_i-mx)^2\cdot\sum_{i=1}^{n}(y_i-my)^2}}$$

Da chi dipende il segno di r? r ha un'unità di misura? Perché?
Il coefficiente di correlazione lineare di Bravais–Pearson può essere agevolmente determinato utilizzando le potenzialità del foglio elettronico, evitando che gli studenti si concentrino eccessivamente sull'aspetto puramente di calcolo tralasciando quello interpretativo.

Alunno	Matematica	Fisica					
	x	*y*	*x - mx*	*y - my*	$(x-mx)^2$	$(y-my)^2$	$(x-mx)(y-my)$
1	5	4	-1,192	-2,423	1,422	5,871	2,88905
2	5	6	-1,192	-0,423	1,422	0,179	0,50444
3	4	5	-2,192	-1,423	4,806	2,025	3,11982
4	6	6	-0,192	-0,423	0,037	0,179	0,08136
5	8	8	1,808	1,577	3,268	2,487	2,85059
…	…	…	…	…	…	…	…
26	5	5	-1,192	-1,423	1,422	2,025	1,69675
Totali	161	167	0	0	54,0385	38,3462	39,8846
Media ar.	6,192	6,423					

Tabella 2

La Tabella 2 riporta soltanto i calcoli relativi alle prime cinque unità statistiche su 26 e fa riferimento al "Voto in matematica" e il "Voto in fisica" ed utilizza il foglio di calcolo Excel.

Perché la somma della 4° e 5° colonna è uguale a 0? Quale significato si può attribuire al segno della somma dell'ultima colonna? Un valore positivo come quello ottenuto indica che tutti gli scarti dalla media aritmetica sono positivi?
Il coefficiente di correlazione lineare r calcolato rispetto ai voti di matematica e fisica è:

$$r = \frac{39,88}{\sqrt{54,04 \cdot 38,35}} = 0,8762$$

Il valore ottenuto è coerente con le osservazioni fatte sulla Figura 1? Sapendo che $r = 1$ quando i punti della nuvola sono allineati su una retta crescente, l'insegnante chiede: dai dati a disposizione ci si poteva attendere un valore di r così alto?
E' un utile esercizio di verifica ricorrere alla funzione statistica messa a disposizione dal foglio elettronico Excel che consente di calcolare in modo automatico il coefficiente di correlazione lineare. Si indica di seguito il percorso da seguire per ottenerne il valore.
Attivare il pulsante "Incolla funzione" per aprire la finestra riportata in Figura 3:

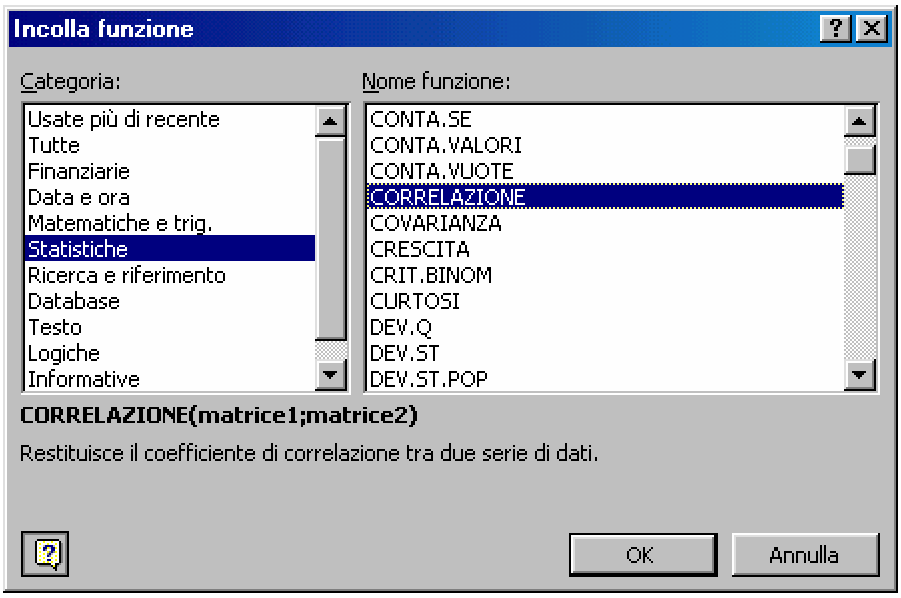


Figura 3

Dopo aver confermato con OK si ottiene la finestra della Figura 4:

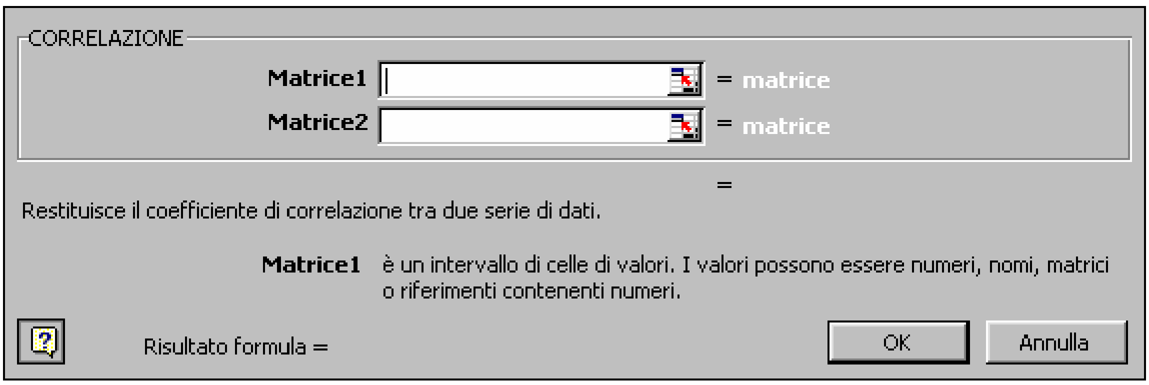


Figura 4

"Matrice1" è un intervallo di celle di valori e, in questo caso, l'intervallo dei valori dei voti di matematica, "Matrice2" è il secondo intervallo di celle di valori e, sempre in questo caso, l'intervallo dei valori dei voti di fisica.
L'indicazione riportata a fianco di Matrice 1 va intesa, nel calcolo del coefficiente di correlazione lineare, come limitata ai valori numerici.
La chiusura della finestra fornisce il valore del coefficiente cercato.
Il risultato ottenuto con la funzione "Correlazione( )" del foglio elettronico coincide con quello ottenuto dai calcoli effettuati seguendo tutti i passaggi?
Osservando la Figura 2, che si riferisce ai voti di matematica ed inglese, si può pensare che la somma dei prodotti degli scarti dalla media aritmetica sia positiva anche in questo caso? L'insegnante propone come esercizio il calcolo del coefficiente r per la distribuzione doppia dei caratteri "Voto in matematica" e "Voto in inglese".

Seconda fase

La forma della nuvola dei punti delle figure precedenti e l'alto valore di r, nel caso esaminato, giustificano la ricerca di una retta che interpoli l'insieme dei punti che rappresentano il fenomeno, passando il più vicino ad essi, secondo il "metodo dei minimi quadrati".
La determinazione di questa funzione può essere fatta, in modo agevole, usando le funzioni del foglio elettronico Excel, come evidenziato di seguito.
Occorre selezionare un punto qualsiasi sul grafico e, cliccando con il pulsante destro, far apparire il menu evidenziato nella Figura 5:

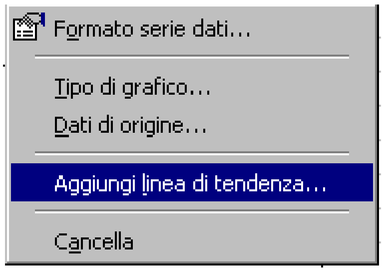


Figura 5

La voce evidenziata in Figura 5 fa apparire la finestra "Aggiungi linea di tendenza". La scheda "Tipo" consente di scegliere una funzione interpolatrice, lineare nel caso in esame[1], mentre la scheda "Opzioni" consente, selezionando le voci spuntate come in Figura 6, di tracciare sul grafico la retta interpolatrice, secondo il metodo dei minimi quadrati, come si vede in Figura 7.

[1] Occorre porre attenzione al fatto che, pur scegliendo come funzione interpolatrice una funzione non lineare, il problema della ricerca dei parametri risolto con il metodo dei minimi quadrati è pur sempre lineare a patto che l'interpolatrice sia di tipo polinomiale o ad esso riconducibile. Si parla in tal caso di modello lineare nei parametri.

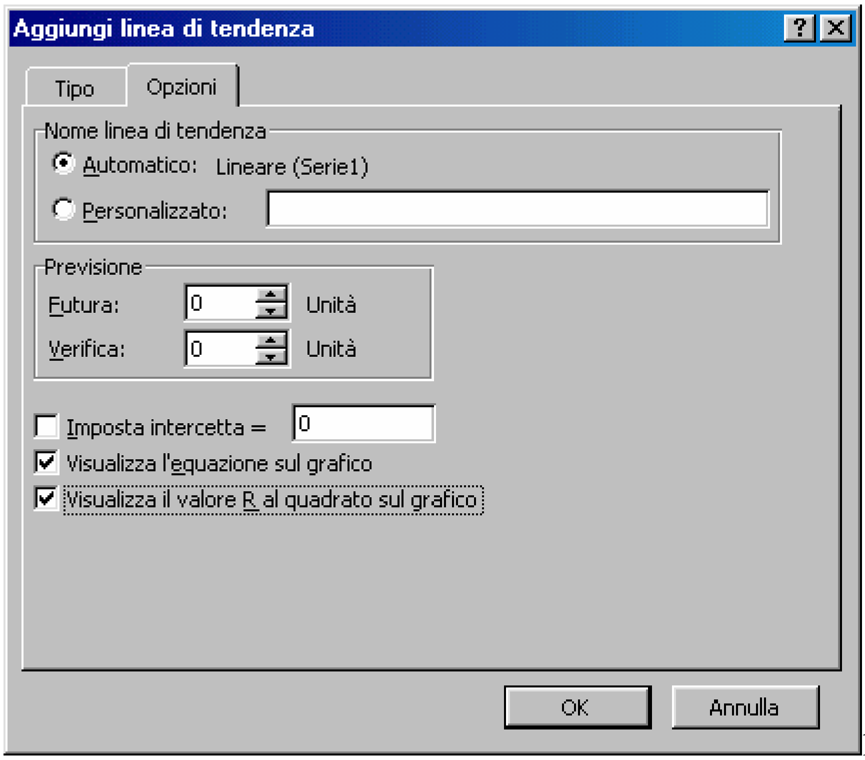


Figura 6

Nella Figura 7, oltre al grafico della retta interpolatrice, il foglio elettronico fornisce anche la sua equazione e l'"Indice di determinazione lineare" R^2 .
R^2 è un indice che può assumere valori compresi tra 0 ed 1. Esso vale 1 quando tutti i punti osservati appartengono ad una retta, crescente o decrescente. In tale caso quella retta esprime la dipendenza lineare che esiste tra X ed Y in base alla quale il valore di Y è determinato quando si conosce quello di X. R^2 vale 0 sia quando esiste indipendenza fra i caratteri, sia quando l'andamento della nuvola dei punti non consente di proporre l'interpolazione con una retta. In questo caso la retta interpolatrice ha equazione $y = 0$ e nota la variabile X non può dirsi nulla sul valore che assume l'altra variabile.
L'insegnante stimola l'attenzione della classe, ponendo una serie di domande sulla Figura 7.
Perché la retta passa per l'origine? C'è una relazione tra il coefficiente r di correlazione lineare di Bravais–Pearson e il coefficiente R^2 di determinazione lineare?
L'insegnante guida gli studenti ad osservare che, per una distribuzione doppia, R^2 altro non è che il quadrato di r.

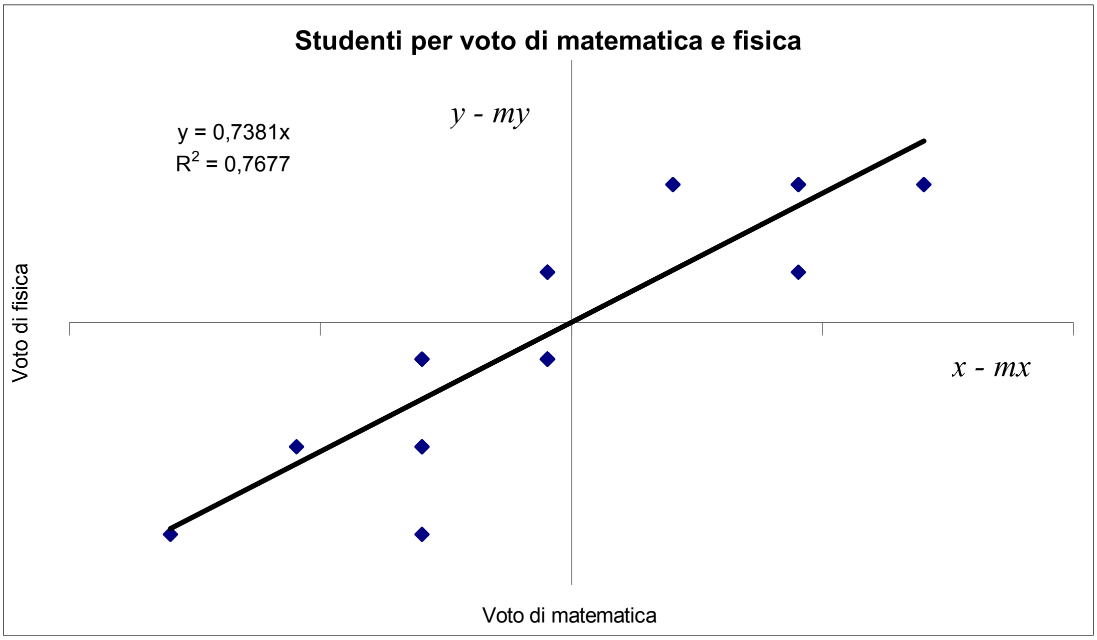


Figura 7

L'equazione della retta interpolatrice fornita dal foglio elettronico è coerente con quanto ci si attende, poiché manca il termine noto. Come si giustifica il valore del coefficiente angolare della retta? Nell'elaborazione dei dati del foglio elettronico usati in precedenza è possibile trovare quelli necessari al calcolo del "Coefficiente angolare (b) della retta interpolatrice"? L'insegnante mostra agli studenti la formula di b che risolve il problema [2]. Tale formula, applicata alla Tabella 2, consente arrivare alla soluzione anche quando non si ha a disposizione il computer.
Gli stessi risultati si possono ottenere utilizzando alcune funzioni di tipo statistico del foglio elettronico:

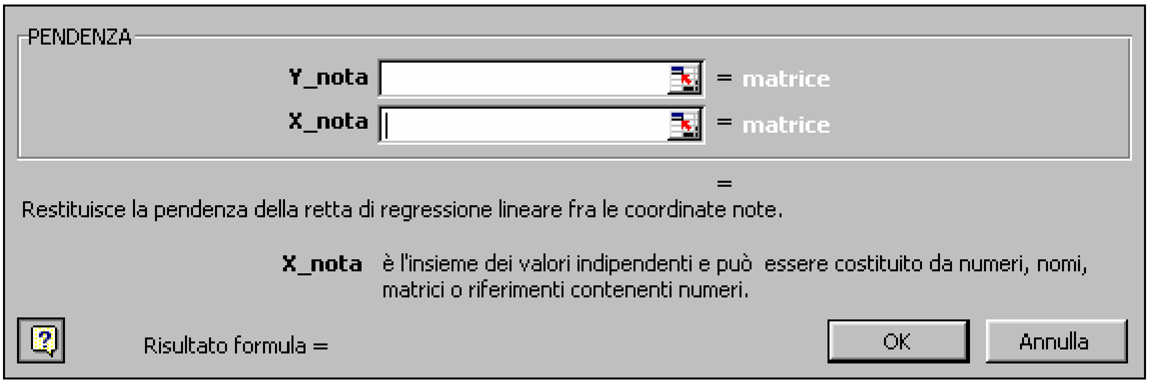


Figura 8

La funzione "Pendenza( )" restituisce il coefficiente angolare della retta di regressione lineare.

[2] La formula di b è:

$$b = \frac{\sum_{i=1}^{n} (x_i - mx)(y_i - my)}{\sum_{i=1}^{n} (x_i - mx)^2}$$

Qual è l'interpretazione statistica del coefficiente di regressione? Si può dire che ad ogni incremento unitario del voto di matematica (X) il voto in fisica (Y) aumenta di b? In quale relazione stanno tra loro r e b?
L'ordinata all'origine in questo caso vale 0, riportando il sistema di assi cartesiani nell'origine, come cambia il termine noto a? Quale interpretazione ha il valore trovato?
Quando il problema richiede di calcolare anche l'intercetta a, il foglio elettronico Excel mette a disposizione la funzione statistica "Intercetta( )" che ha una sintassi del tutto analoga alla funzione di Figura 8.

Nasce un'impresa!

Livello scolare: 2° biennio

Abilità interessate	Conoscenze	Nuclei coinvolti	Collegamenti esterni
Predisporre la struttura della matrice dei dati grezzi con riguardo ad una rilevazione pianificata ed inserire i dati in un foglio elettronico. Passare dai dati grezzi alle distribuzioni statistiche di frequenze ed alle corrispondenti rappresentazioni grafiche. Identificare situazioni che richiedono di rilevare lo stesso carattere su un'unità statistica formata da 2 elementi, o 2 caratteri diversi sulla stessa unità statistica. Impostare una tabella a doppia entrata; classificare i dati secondo 2 caratteri e riconoscere in essa i diversi elementi individuabili. Selezionare, produrre ed usare appropriate rappresentazioni grafiche. Valutare criticamente le informazioni fornite dai media, con riferimento particolare ai sondaggi ed ai giochi di sorte.	Distribuzioni doppie di frequenza e tabelle a doppia entrata. Distribuzioni condizionate e marginali. Principali rappresentazioni grafiche per le distribuzioni doppie rispetto a caratteri di qualsiasi natura.	<u>Dati e previsioni</u> Laboratorio di Matematica	Lingua italiana Economia e marketing

Contesto
Sociale: progettazione.

L'attività si sviluppa in un contesto extramatematico ed interdisciplinare, quello dell'area di progetto, e riguarda la simulazione della nascita di un'impresa che dovrà operare nel settore della vendita di automobili.
Tale attività si è sviluppata nell'ultimo biennio dell'I.T.C.S. "L. Einaudi" di Padova nell'a.s. 2001/2002. Lo scopo, attraverso la compilazione, la somministrazione e l'elaborazione di un questionario, è quello di creare la base informativa per compilare il cosiddetto BUSINESS PLAN, cioè quel documento che guida la nascita e lo sviluppo della nuova impresa, nei suoi primi anni di vita. L'esperienza viene riportata come è stata descritta dagli stessi alunni del gruppo matematico che l'ha condotta ed è un esempio di come la Statistica opera per creare basi informative e previsioni economiche.

Descrizione dell'attività

L'attività ha per scopo l'analisi dei problemi inerenti l'ubicazione dell'autosalone, il target della clientela e l'analisi del mercato automobilistico su cui indirizzare il settore operativo dell'impresa.
Per raggiungere gli obiettivi prefissati, all'inizio dell'anno scolastico, l'insegnante aiuta gli alunni a predisporre un questionario riguardante il mercato dell'auto, cercando di individuare le informazioni essenziali da raccogliere.
Il questionario è stato successivamente compilato dagli studenti delle classi 4^{e} e 5^{e} dell'istituto e, per la codifica e l'elaborazione delle risposte, è stato predisposto un datatabase con il foglio elettronico Excel.
Si sono create, poi, per l'analisi dei risultati, tante tabelle pivot quanti sono i quesiti posti.
Le elaborazioni contenute nelle tabelle, sono state differenziate rispetto al sesso degli intervistati.
I valori assoluti, per i confronti, sono stati trasformati in percentuale. Per rendere chiare e comprensibili le tabelle, esse sono state accompagnate da opportune rappresentazioni grafiche (quasi tutte con diagrammi a settori circolari) ed è stato affiancato, ad ogni elaborazione, un commento personale degli alunni sui risultati dell'indagine.
Ipotizzando anche di servire una clientela proveniente da paesi stranieri, il questionario è stato tradotto in tedesco ed in inglese.
Le informazioni ricavate hanno quindi costituito la base per rispondere ad altri quesiti che la simulazione dell'attività pone:

a) la determinazione del tipo di auto da vendere da parte del concessionario,
b) la scelta della zona più favorevole in cui far sorgere l'autosalone.

Le informazioni sulla fascia di prezzo che i potenziali acquirenti sono disposti a pagare, è servita al gruppo che si è occupato della parte giuridica e della pubblicità. Sono state elaborate le informazioni riguardanti la più opportuna localizzazione del concessionario.
I risultati salienti dell'indagine, tenuto conto delle preferenze espresse, si possono così riassumere:

1) gli intervistati preferiscono acquistare un'auto nuova;
2) la somma da destinare all'acquisto dell'auto è compresa fra € 5.000 ed € 10.000;
3) le auto preferite dai giovani sono quelle sportive;
4) la collocazione ideale dell'autosalone è in periferia;
5) i servizi più richiesti sono la velocità di consegna e l'appoggio di un'autofficina.

L'attività è terminata con l'esposizione dei risultati raggiunti dai vari gruppi ad un esperto del settore, che ha fornito una valutazione positiva dell'attività svolta.
A completamento del lavoro, si è predisposto un sito Web su cui far confluire i risultati dell'indagine e pubblicizzare l'impresa di nuova costituzione.

Analisi dei dati

La sintesi dei risultati esposti è il frutto di elaborazioni di un database di 289 righe o record (pari al numero degli studenti rispondenti) e di 37 colonne o campi (numero delle domande poste tenuto conto che alcune sono articolate in più quesiti), costruito in base alle risposte al questionario somministrato agli studenti di tutte le classi quarte e quinte dell'I.T.C.S. "L. Einaudi" di Padova, scelti come rappresentanti del segmento "giovani" di potenziali acquirenti dell'auto.
Si propongono di seguito alcune delle tabelle ricavate con le tabelle pivot del foglio Excel e i grafici prodotti nel corso della ricerca. A corredo di ogni elaborazione gli studenti hanno scritto i loro commenti che si riportano integralmente.
Tali tabelle giustificano i risultati sopra elencati, in particolare riguardano:

I. studenti e automobili;
II. studenti e collocazione di un salone multimarche;
III. studenti e servizi di un autosalone multimarche.

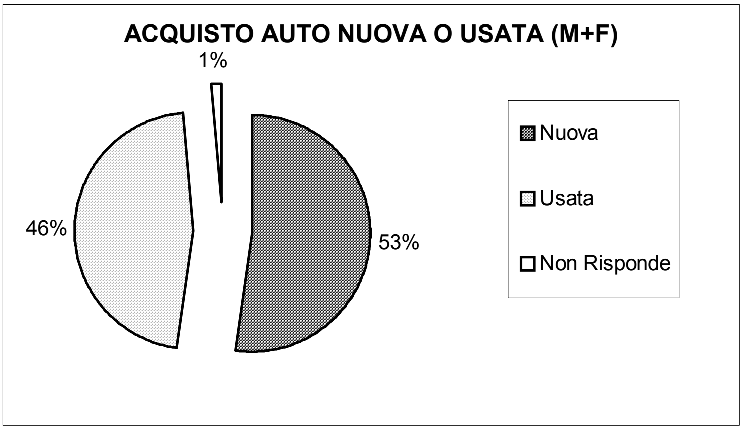
Gli studenti nella loro ricerca di esprimere in modo molto sintetico il titolo delle tabelle e dei grafici corrispondenti non hanno tenuto nel dovuto conto che occorre ben definire il collettivo osservato. In

questo caso si tratta sempre degli “Studenti rispondenti delle classi 4ᵉ e 5ᵉ dell’I.T.C.S. “L. Einaudi” di Padova, nell’anno scolastico 2001/2002”.
Le distribuzioni costruite sono sempre distribuzioni doppie nelle quali uno dei caratteri di classificazione è il “Sesso". Per quanto riguarda grafici e commenti ci si è riferiti alla distribuzione totale (M+F).

Gli studenti e l’automobile

Acquisto auto nuova o usata

	SESSO					
Acquisto Auto	F	M	Totale	%F	%M	%TOT
Nuova	116	35	151	54,46	46,05	52,25
Usata	95	39	134	44,60	51,32	46,37
Non Risponde	2	2	4	0,94	2,63	1,38
Totale	213	76	289	100,00	100,00	100,00

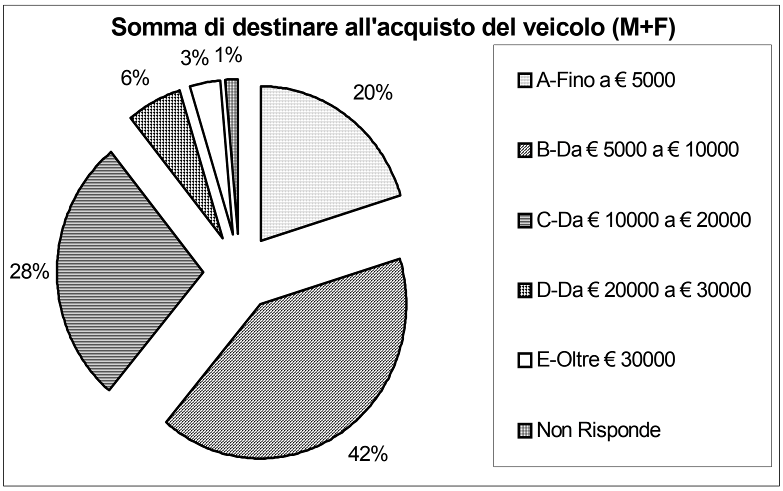


Il 53% degli intervistati preferirebbe acquistare un'auto nuova, mentre il 46% acquisterebbe un usato. Ciò significa che il nostro autosalone dovrebbe prendere in considerazione la possibilità di ritirare e di vendere usati.

Tabella 1

Somma da destinare all'acquisto di un veicolo nuovo

	SESSO					
Somma acquisto nuovo	F	M	Totale	%F	%M	%TOT
A-Fino a € 5000	39	20	59	18,31	26,32	20,42
B-Da € 5000 a € 10000	98	20	118	46,01	26,32	40,83
C-Da € 10000 a € 20000	59	22	81	27,70	28,95	28,03
D-Da € 20000 a € 30000	10	7	17	4,69	9,21	5,88
E-Oltre € 30000	5	5	10	2,35	6,58	3,46
Non Risponde	2	2	4	0,94	2,63	1,38
Totale	213	76	289	100,00	100,00	100,00



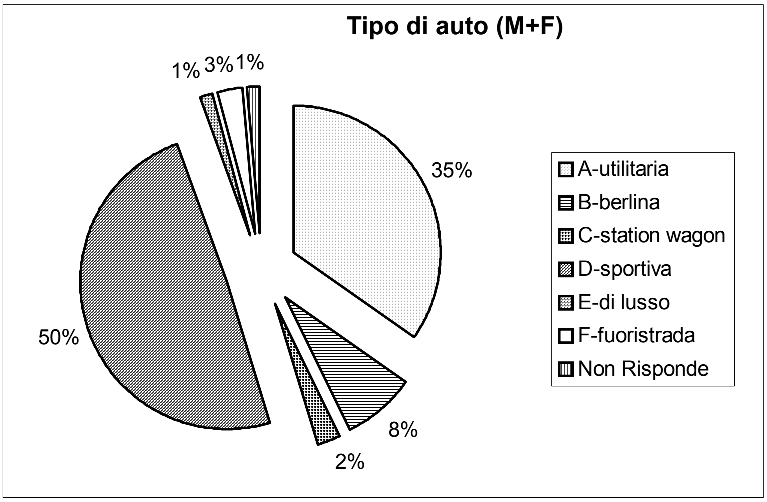
Dal risultato dell'intervista è possibile evidenziare che ben il 42%[1] della possibile clientela è disposta ad acquistare un auto spendendo da 5 a 10 mila euro. Un'altra percentuale, il 28%, è disposta a spendere da 10 a 20 mila euro. Altra percentuale da evidenziare è il 20% degli intervistati, i quali sono disposti ad acquistare un veicolo sborsando al massimo 5 mila euro e il veicolo potrà essere soltanto un usato o un veicolo nuovo ma dal modello vecchio.

Tabella 2

[1] Si osservi che, dopo aver arrotondato all'intero più vicino tutte le percentuali esclusa la maggiore, perché la somma risulti 100 è necessario arrotondare 40,83 a 42. In tal modo si produce l'errore relativo minimo.

Tipo di automobile

	Sesso					
9-c:tipo di auto	F	M	Totale	%F	%M	%TOT
A-utilitaria	83	17	100	38,97	22,37	34,60
B-berlina	18	6	24	8,45	7,89	8,30
C-station wagon	3	4	7	1,41	5,26	2,42
D-sportiva	102	40	142	47,89	52,63	49,13
E-di lusso	1	3	4	0,47	3,95	1,38
F-fuoristrada	4	4	8	1,88	5,26	2,77
Non Risponde	2	2	4	0,94	2,63	1,38
Totale	213	76	289	100,00	100,00	100,00



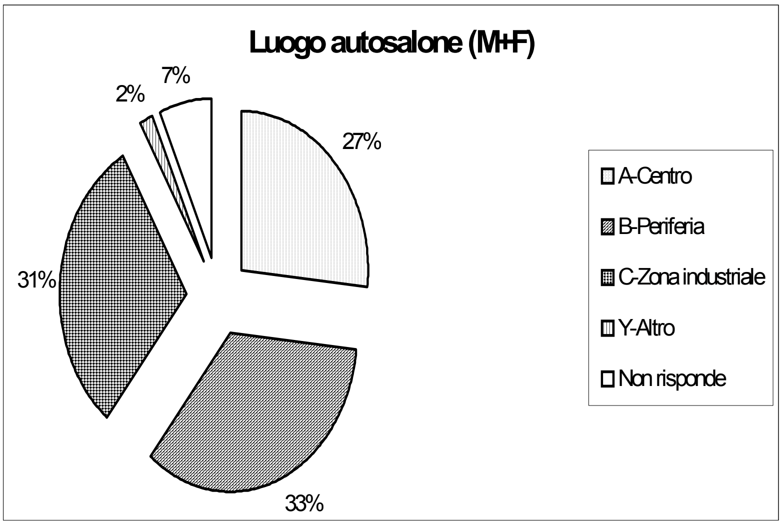
Dal grafico emerge che quasi la metà degli intervistati è interessata all'acquisto di un'auto sportiva e che un'altra fetta importante di preferenze è assegnata alle auto utilitarie. Ciò significa che il nostro bacino di utenza non è interessato all'acquisto di automobili di lusso o familiari né di fuoristrada. Il nostro autosalone dovrà quindi puntare sulla vendita di auto sportive o utilitarie.

Tabella 3

Studenti e collocazione di un salone multimarche:

Dove è situato l'autosalone

	Sesso					
Luogo Autosalone	F	M	Totale	%F	%M	%TOT
A-Centro	59	18	77	27,70	23,68	26,64
B-Periferia	77	22	99	36,15	28,95	34,26
C-Zona industriale	62	27	89	29,11	35,53	30,80
W ovunque		1	1	0,00	1,32	0,35
W-Altro: dappertutto	1	1	2	0,47	1,32	0,69
W-Altro: in internet		1	1	0,00	1,32	0,35
Y-Altro	1		1	0,47	0,00	0,35
Non risponde	13	6	19	6,10	7,89	6,57
Totale	213	76	289	100,00	100,00	100,00



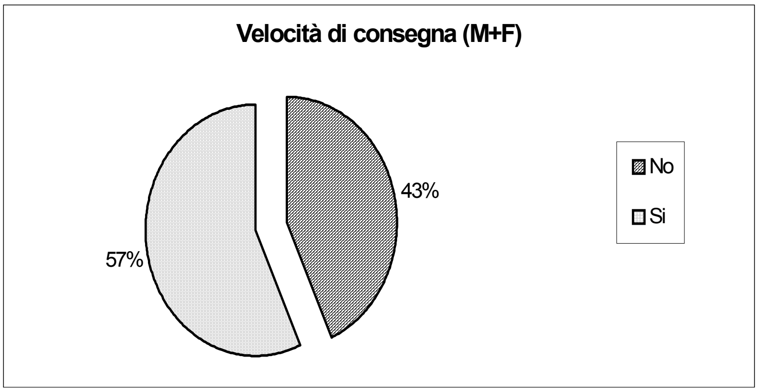
Dal grafico risulta che la collocazione ideale per l'autosalone è in periferia, anche se molto rilevanti sono le percentuali di chi lo preferirebbe in centro o in zona industriale. Oltre alle preferenze più o meno giustificabili degli studenti, potrebbe aver influito nella risposta anche la loro zona di residenza.

Tabella 4

Studenti e servizi di un autosalone multimarche:

Servizio offerto: velocità nella consegna

	Sesso					
Servizi forniti:Velocità consegna	F	M	Totale	%F	%M	%TOT
No	94	31	125	44,13	40,79	43,25
Si	119	45	164	55,87	59,21	56,75
Totale	213	76	289	100,00	100,00	100,00

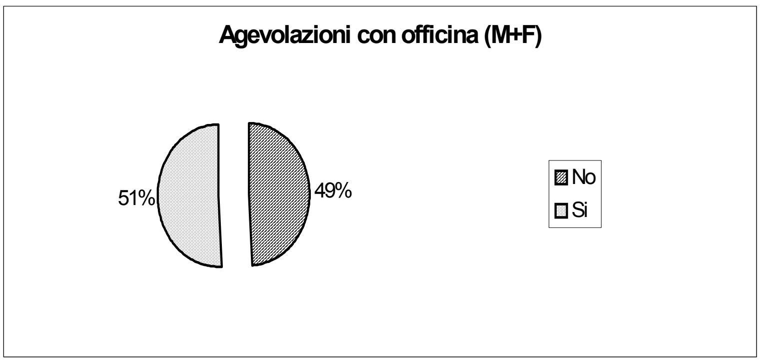


Tra gli ipotetici servizi che l'autosalone dovrebbe offrire, il 57% degli studenti considera più importante la velocità di consegna dell'automezzo acquistato.

Tabella 5

Servizio offerto: agevolazioni in officina

	Sesso					
Servizio fornito:Agevolazioni con officina	F	M	Totale	%F	%M	%TOT
No	103	39	142	48,36	51,32	49,13
Si	110	37	147	51,64	48,68	50,87
Totale	213	76	289	100,00	100,00	100,00



La disponibilità di un'officina guadagna appena la maggioranza delle preferenze tra i servizi che l'autosalone dovrebbe offrire. Il restante 49% probabilmente preferisce rivolgersi alla sua officina di fiducia.

Tabella 6

Questionario

AREA DI PROGETTO 5^E: QUESTIONARIO

RICERCA DI MERCATO RIGUARDANTE IL SETTORE AUTOMOBILISTICO

Rispondi alle seguenti domande secondo le tue preferenze e possibilità economiche

1- CLASSE SEZIONE CORSO
- [ ] Igea
- [ ] Brocca
- [ ] Iter

2- SESSO M F

3- ETA'

4- COMUNE DI RESIDENZA

5- PROFESSIONE DEL CAPOFAMIGLIA:
- [ ] lavoratore dipendente
- [ ] lavoratore professionista
- [ ] lavoratore autonomo
- [ ] imprenditore
- [ ] altro ______________________

6- SEI IN POSSESSO DELLA PATENTE DI GUIDA?
- [ ] Sì
- [ ] No

se la risposta è NO
- [ ] ho già richiesto il foglio rosa e/o sono iscritto ad una scuola guida
- [ ] penso di avviare le pratiche entro i prossimi 3 mesi
- [ ] non ho intenzione di conseguirla per il momento
- [ ] altro ______________________

7- DI QUALI MEZZI DI TRASPORTO LA TUA FAMIGLIA DISPONE?
- [ ] nessuno
- [ ] ciclomotore
- [ ] moto
- [ ] auto
- [ ] fuoristrada
- [ ] altro mezzo

Se possiede auto quante ne ha?
- [1] di quale marca_________________
  di quale cilindrata _______________
- [2] di quale marca_________________
  di quale cilindrata _______________
- [3 o +] di quale marca_________________
  di quale cilindrata _______________

8- NELL'AMBITO DEL NUCLEO FAMILIARE E' STATO GIA' ACQUISTATO UN VEICOLO NUOVO?
(negli ultimi 6 mesi)

- [ ] Sì
- [ ] No

9- NEL CASO TU POTESSI DESTINARE UNA SOMMA ALL'ACQUISTO DI UN'AUTOMOBILE (secondo le tue possibilità economiche) SARESTI DISPOSTO/A AD ACQUISTARLA

- [ ] Nuova
- [ ] Usata

E QUALE SOMMA SPENDERESTI? (una sola possibilità)

- [ ] fino a 5.000 Euro
- [ ] da 5 a 10 mila Euro
- [ ] da 10 a 20 mila Euro
- [ ] da 20 a 30 mila Euro
- [ ] oltre 30 mila Euro

NELLA SCELTA DEL TIPO DI AUTOMOBILE SEI ORIENTATO AD UNA

- [ ] utilitaria
- [ ] berlina
- [ ] station wagon/ familiare
- [ ] sportiva
- [ ] di lusso
- [ ] fuoristrada

10- PER LA SCELTA RIGUARDANTE L'ACQUISTO

- [ ] ti interessi personalmente
- [ ] ti informi su riviste specializzate
- [ ] ti rivolgi a parenti/ amici/ conoscenti
- [ ] ti fai consigliare da esperti
- [ ] altro ________________________

11- ATTUALMENTE LA DISTRIBUZIONE AVVIENE TRAMITE CONCESSIONARI MONOMARCA, TI INTERESSEREBBE TROVARE UN LUOGO DOVE PUOI ACQUISTARE E PROVARE PIU' DI UN MODELLO CON ADDETTO AUTOSALONE MULTIMARCHE?

- [ ] Sì
- [ ] No

DOVE VORRESTI TROVARLO?

- [ ] in centro
- [ ] in periferia
- [ ] nella zona industriale
- [ ] altro_______________________

12-QUALI SERVIZI O AGEVOLAZIONI VORRESTI AVERE RIVOLGENDOTI AD UN AUTOSALONE MULTIMARCHE ?
(è possibile più di una risposta)

- [ ] velocità nella consegna
- [ ] agevolazioni con officina
- [ ] altro ______________________

13- SEI A CONOSCENZA DELLE POSSIBILITA' DI ACQUISTO IN INTERNET DI AUTOMOBILI?

- [ ] Sì
- [ ] No

USERESTI LA RETE PER ACQUISTARLE?

- [ ] Sì
- [ ] No

14- RITIENI CHE L'ACQUISTO IN RETE DEBBA AVVENIRE:

- [ ] senza alcun servizio o agevolazioni
- [ ] con l'appoggio di un autosalone che offre servizi ed agevolazioni
- [ ] con l'appoggio di un'officina di assistenza

Grazie per la disponibilità che ci avete concesso !!!!!!

Elementi di prove di verifica

Formula di Bayes e gioco equo

1. In un torneo di calcio all'italiana si sa che la squadra A ha vinto il torneo. Ricordo che le squadre che hanno disputato l'altra semifinale erano B e C, ma non ricordo più quale delle due sia arrivata in finale. Se la squadra B aveva probabilità 2/5 di battere la squadra C in semifinale e la squadra A aveva probabilità ¼ di battere C e 2/3 di battere B, qual è la probabilità che l'altra squadra finalista sia stata la B?

2. Antonio vuole costruire un puzzle da 300 pezzi. La sorella ha messo i pezzi di due puzzle diversi, uno da 300 e uno da 100 pezzi, in due contenitori uguali. Antonio non sa qual è il contenitore che deve prendere, sa solo che la sorella, nel rimettere a posto i pezzi, ha cambiato di posto a 30 di essi. Antonio prende un pezzo, scegliendo a caso da uno dei due contenitori.
Che probabilità ha di pescare un pezzo utile?
Se Antonio ha preso un pezzo utile, che probabilità c'è che provenga dal contenitore giusto?

3. La probabilità che l'azienda A, produttrice di videogiochi abbia un incremento del fatturato è 2/3. In base ai dati delle gestioni precedenti, si sa che l'azienda ha un incremento del 2% con probabilità ½; un incremento del 5% con probabilità 1/3; un incremento del 10% con probabilità 1/6. Sapendo che l'azienda, a fine anno, non ha avuto un incremento del 10%, qual è la probabilità che abbia avuto un incremento del fatturato del 5% ?

4. In un'urna vi sono 3 palline rosse e una nera. Si gioca ad estrarre una coppia di palline alla volta. Saresti disposto a scommettere con un tuo amico alla pari in questo modo:
Vinci se esce la coppia "rossa-nera"
Vince il tuo amico se esce la coppia "rossa-rossa"?

5. In un'urna vi sono 8 palline rosse, 6 nere e 5 verdi. Si gioca in due ad estrarre una coppia di palline alla volta. Uno dei due giocatori vince se esce la coppia "rossa-rossa" e l'altro se esce la coppia "rossa-nera"; gli altri casi non vengono considerati.
Quale dei due giocatori ha maggiore probabilità di vincere? Come deve essere la ripartizione delle quote perché il gioco sia equo?

Griglia di correzione ed osservazioni

1. 16/25

2. 0,6 ; 0,75

3. 1/4

4. Nel caso dell'esercizio 4, gli studenti spesso rispondono che la scommessa alla pari non è accettabile, perché sono fuorviati dal fatto che le palline rosse sono il triplo di quelle nere. Il calcolo combinatorio può essere utile per calcolare il numero di coppie possibili nei vari casi. Si trova che la coppia "rossa-nera" si può avere in tre modi e così la coppia "rossa-rossa"; si ha *P(rossa-nera)* = *P(rossa-rossa)* = 1/2 per cui la scommessa alla pari è equa.

5. In questo caso la scommessa alla pari non è accettabile; infatti la coppia "rossa-rossa" si può avere in 28 modi, mentre la coppia "rossa-nera" si può avere in 48 modi. Il gioco sarà equo se il giocatore che scommette sulla coppia "rossa-nera" è disposto a pagare una somma che sia i 12/7 di quella che scommette il suo avversario.

Riferimenti bibliografici

Anichini, G. (1988), Quanto è probabile avere lo stesso compleanno, in *Archimede*, 1, pp. 19-29.

Barra, M. (2000), Probabilità e gioco d'azzardo, in *Le Scienze il loro insegnamento*, nn.5-6, pp.26-32.

Brunelli, L. (2000), e altri, Un'indagine in classe per apprendere la statistica. Guida per un corso di base di statistica descrittiva, in *Induzioni*, 21.

de Finetti, B., (1970) *Teoria delle probabilità*, vol.I e II, Einaudi.

Fraire, M. - Rizzi, A. , (2001), *Esercizi di Statistica*, Carocci Editore.

Galilei, G., (1993), Sopra le scoperte de i dadi, in *Induzioni*, 7, pp. 15-17

Ministero della Pubblica Istruzione, (1995), Direzione Classica, L'insegnamento di probabilità e statistica nella scuola liceale, Seminario di formazione per docenti, in *Quaderni*, 8.

Ministero della Pubblica Istruzione, (1995), Direzione Classica, Probabilità e statistica nella scuola liceale, Seminario di formazione per docenti, in *Quaderni*, 28.

Ottaviani, M.G., (2001), Strumenti per l'analisi dei dati, in *Induzioni*, 23, pp. 33-81.

Ottaviani, M.G., (2002), Strumenti per l'analisi dei dati, II^a^ parte, in *Induzioni*, 24, pp. 39 - 77.

Pezzulli, S. (2001), Le rappresentazioni grafiche dei dati statistici. Materiali e metodi introduttivi, in *Induzioni*, 23, pp. 83-127.