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.	<u>Dati e previsioni</u> 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

Prima fase

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

Seconda fase

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 espolrare 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.

Prima fase
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:

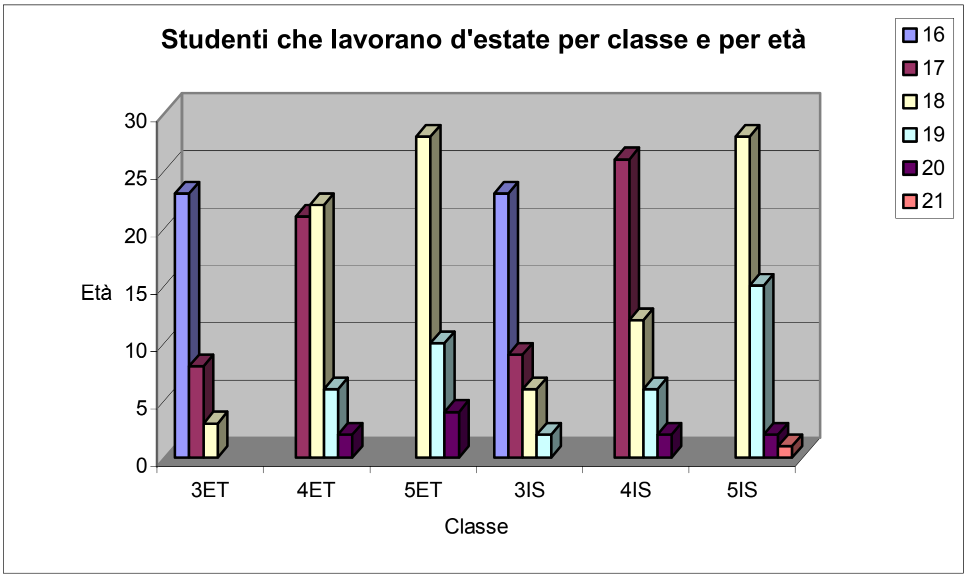


Figura 1

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.

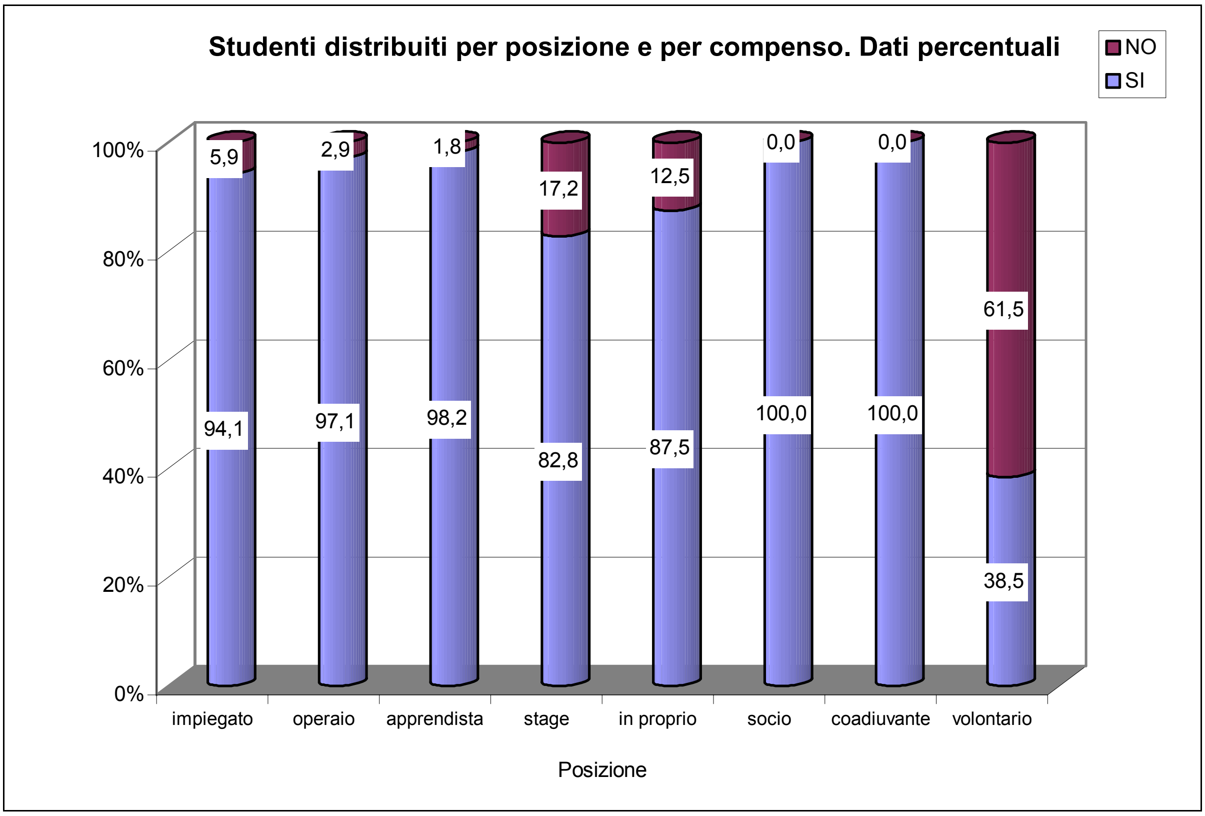


Figura 2

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!

Terza fase

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.

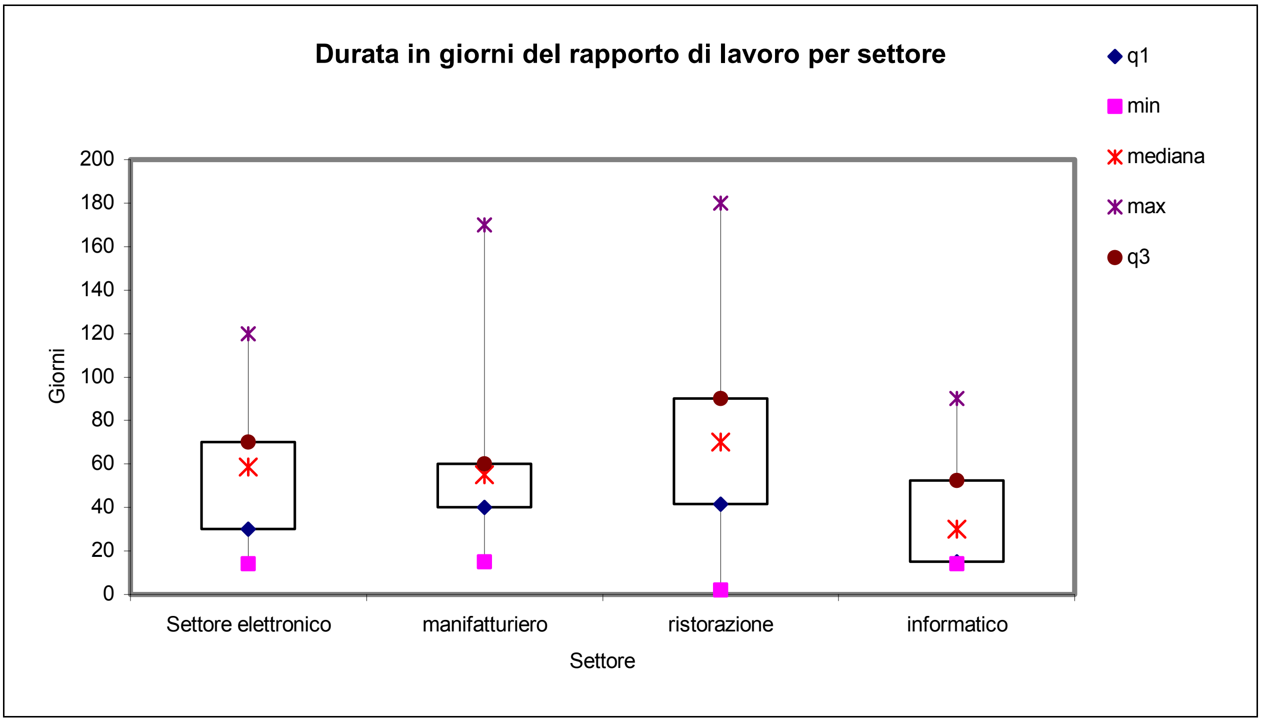


Figura 3

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.	Dati e previsioni 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).
Nel foglio elettronico Excel allegato (si veda simulazione) sono riportati i valori ottenuti e i grafici relativi alle frequenze di Testa per 20, 120 e 300 lanci.
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).

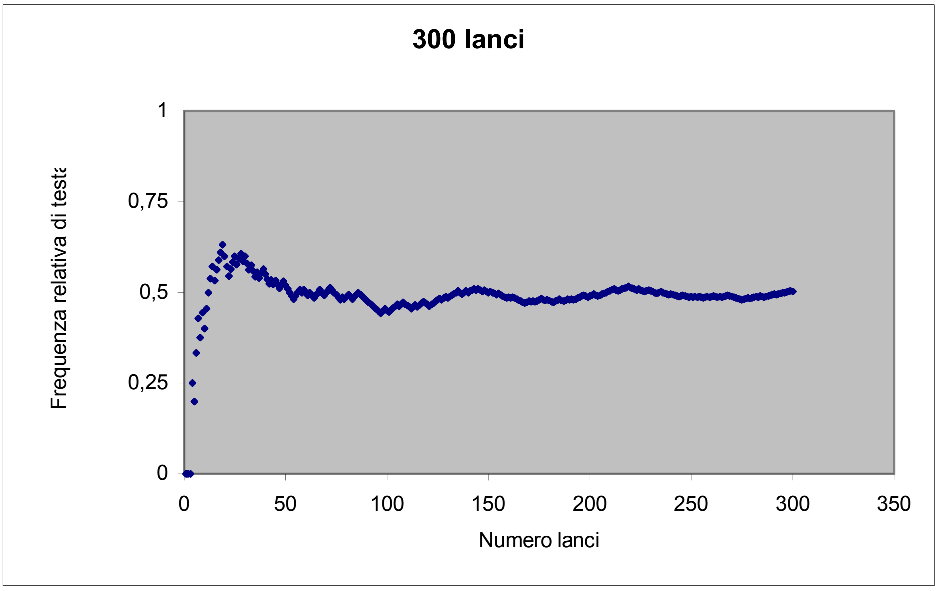


Figura 2

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 (vedi foglio differenze).

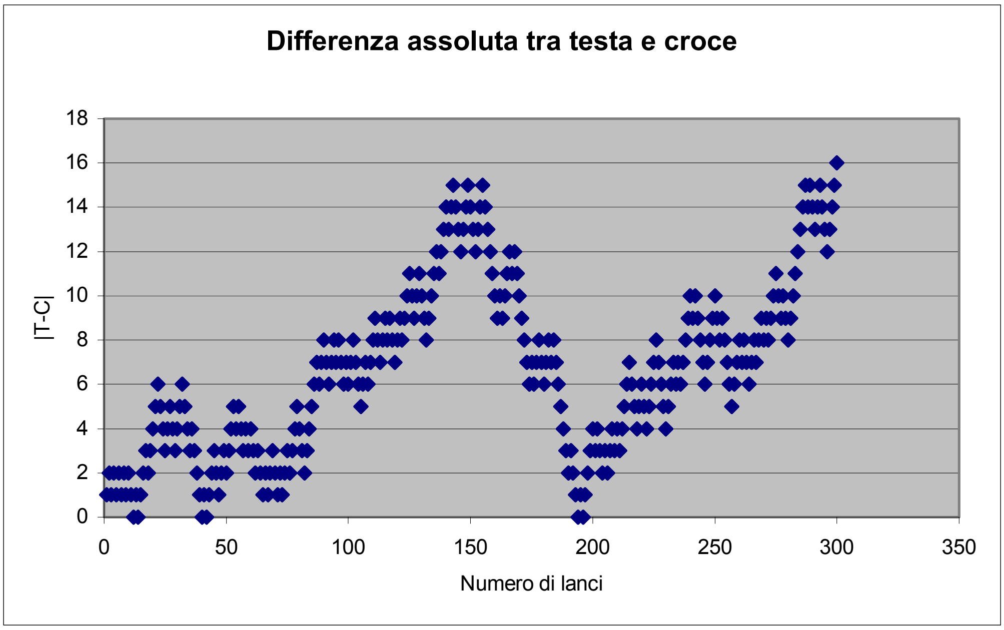


Figura 3

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 $|2T\text{-}n|$ va tenuto presente che la varianza del numero aleatorio ($2T\text{-}n$) 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 $|2T\text{-}n|$ (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.	Dati e previsioni 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.

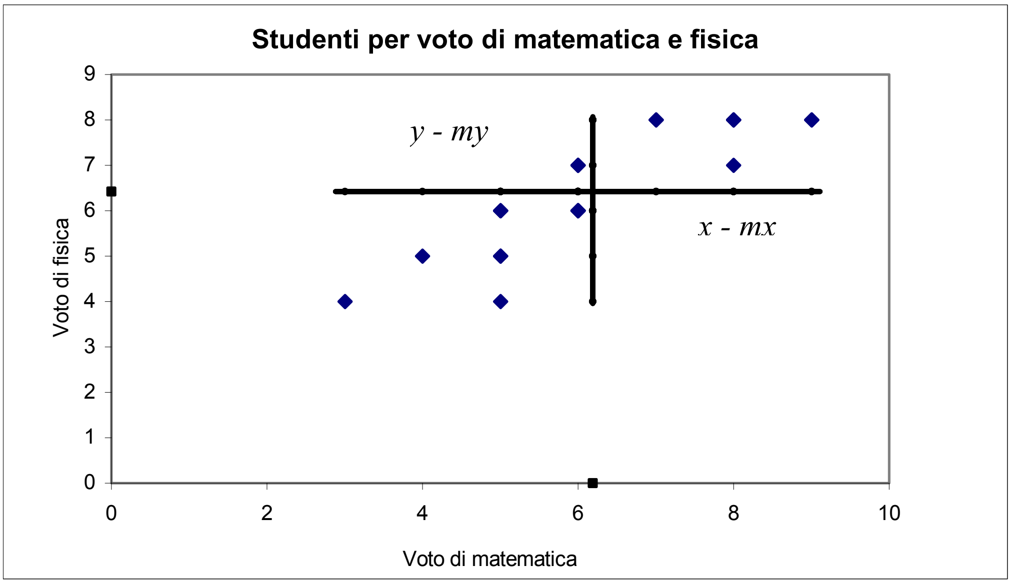


Figura 1

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?

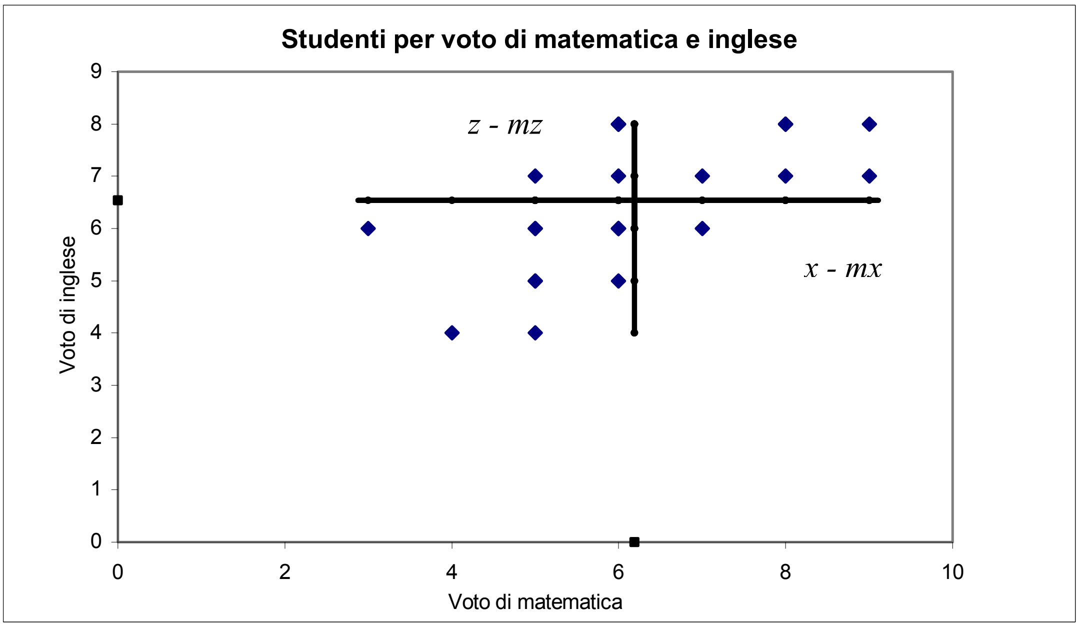


Figura 2

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:

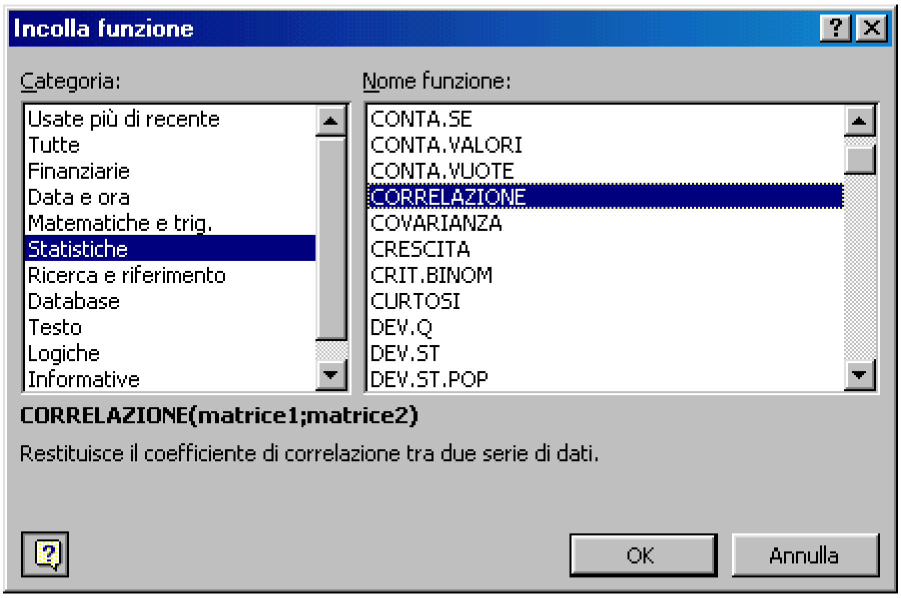


Figura 3

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

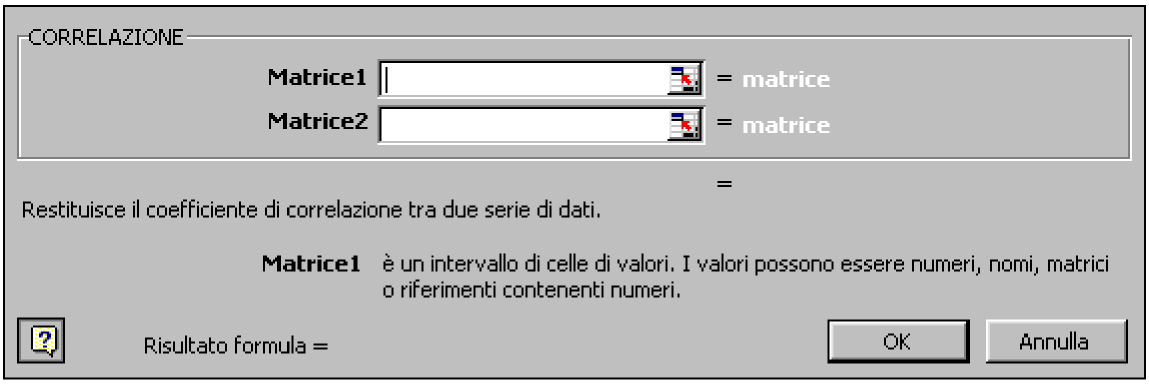


Figura 4

"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".

<u>Seconda fase</u>
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:

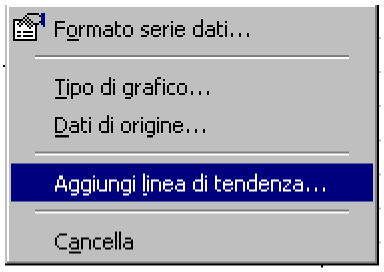


Figura 5

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.

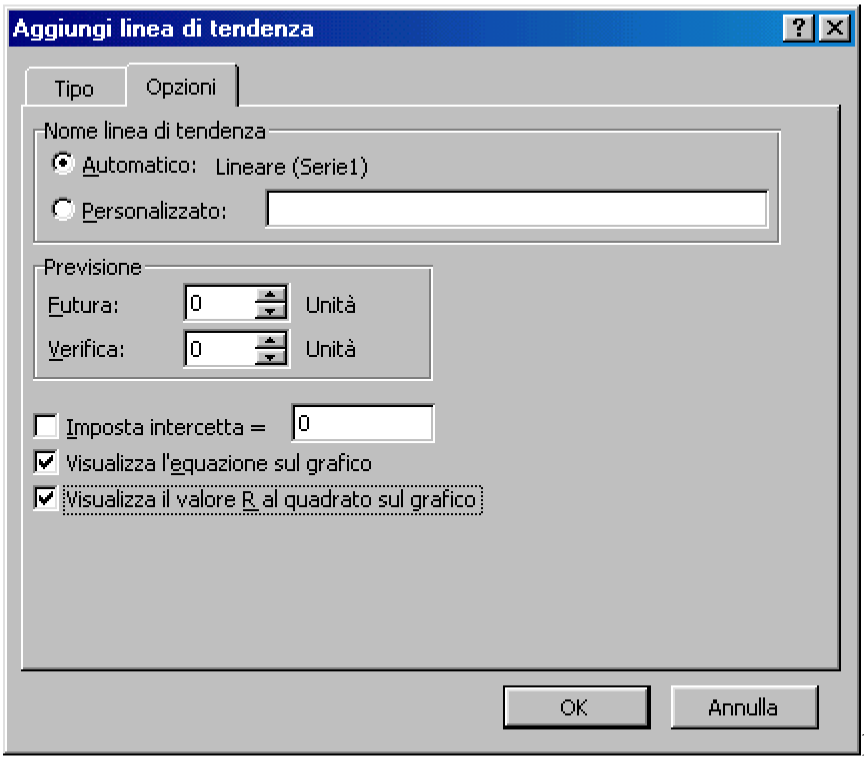


Figura 6

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.

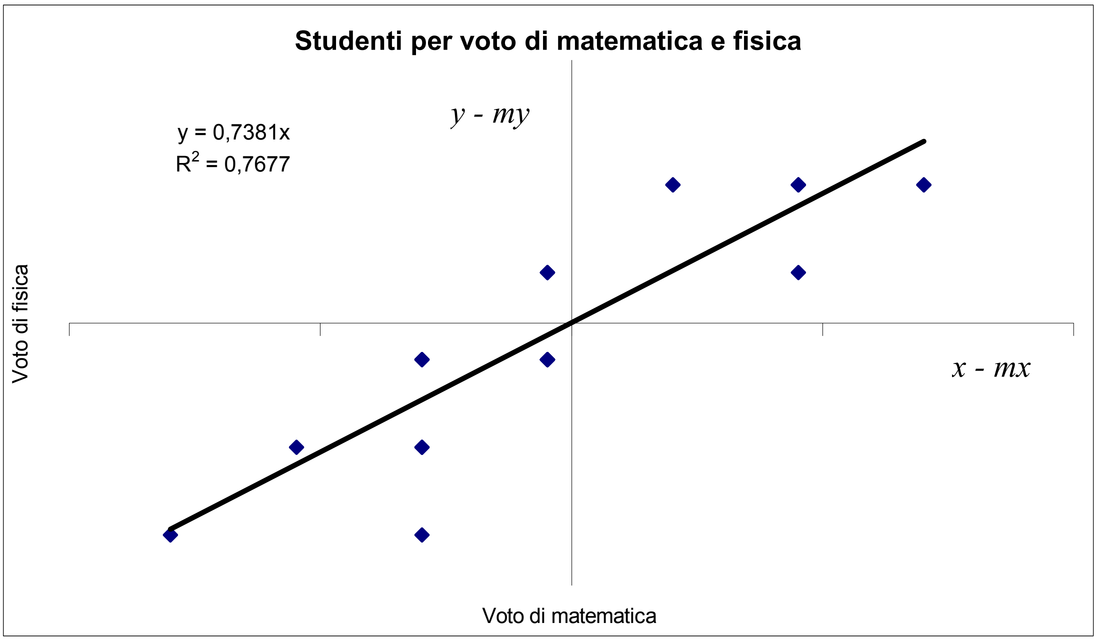


Figura 7

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:

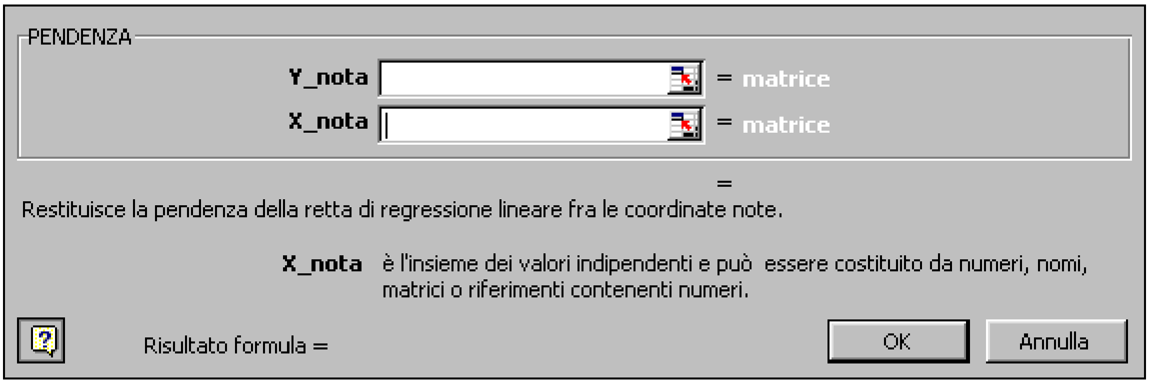


Figura 8

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 4^e e 5^e 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

Acquisto Auto (Femmine, Maschi, 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)

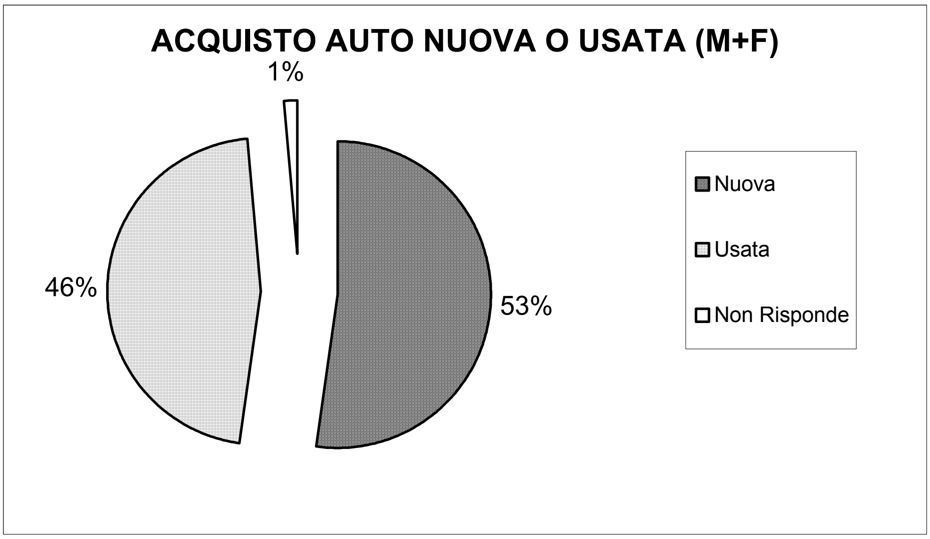


Tabella 1

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.

Somma da destinare all’acquisto di un veicolo nuovo

Somma acquisto nuovo (Femmine, Maschi, 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)

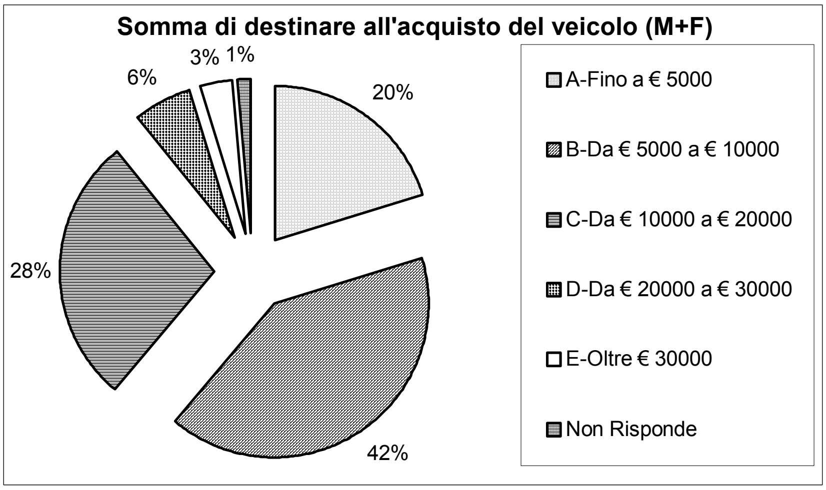
C-Da € 10000 a € 20000 (59, 22, 81, 27,70, 28,95, 28,03)

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%[i] 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

Nota: 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

Tipo di auto (Femmine, Maschi, 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)

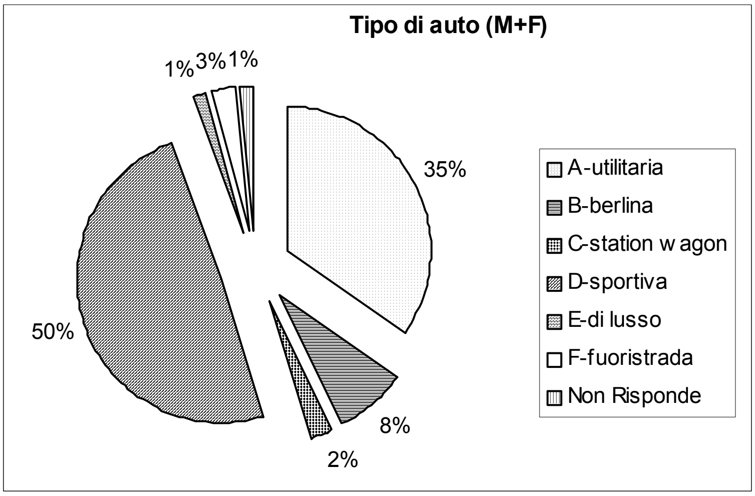


Tabella 3

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.

Studenti e collocazione di un salone multimarche. Dove è situato l'autosalone

Luogo Autosalone (Femmine, Maschi, 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, 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, 1,32, 0,35, 0,00)

Y-Altro (1, 0, 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)

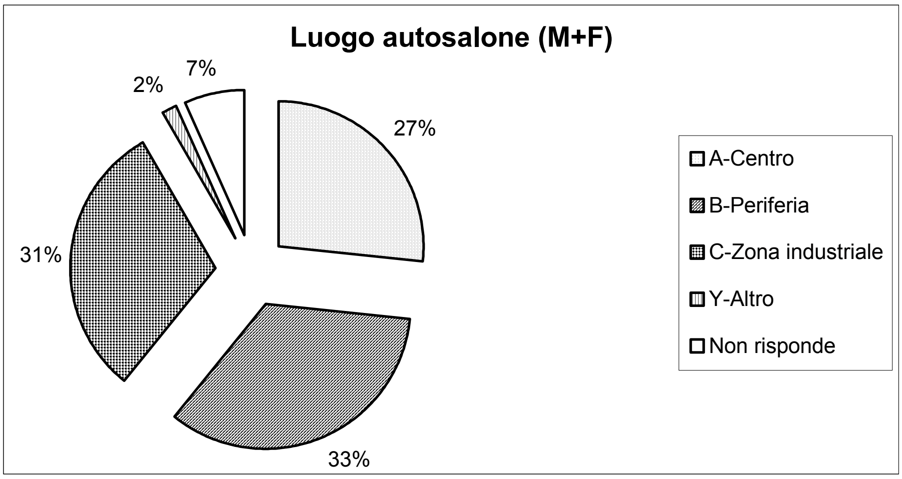


Tabella 4

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.

Studenti e servizi di un autosalone multimarche. Servizio offerto: velocità nella consegna

Servizi forniti:Velocità consegna (Femmine, Maschi, 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.

Servizio offerto: agevolazioni in officina (Femmine, Maschi, 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.

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 !!!!!!

L'elaborazione dei dati raccolti, i commenti , i grafici e le conclusioni della ricerca sono riportati nei file Excel allegati.

Database

Grafici e commenti

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.

Ottaviani, M.G., (2001), Strumenti per l'analisi dei dati, II^a parte, in *Induzioni*, 23..

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