Università di Torino

QUADERNI DIDATTICI
del
Dipartimento di Matematica "G. Peano"

MARIA GARETTO

Laboratorio di Statistica con Excel

Soluzioni

Corso di Laurea in Biotecnologie
A.A. 2009/2010

Quaderno # 46 – Dicembre 2009

INDICE

Prefazione

3 DISTRIBUZIONI DI FREQUENZA, STATISTICHE

4 CORRELAZIONE E REGRESSIONE

5 DISTRIBUZIONI DI PROBABILITA'

6 STIMA DEI PARAMETRI

7 TEST DI IPOTESI

8 TEST CHI-QUADRO

Torna su

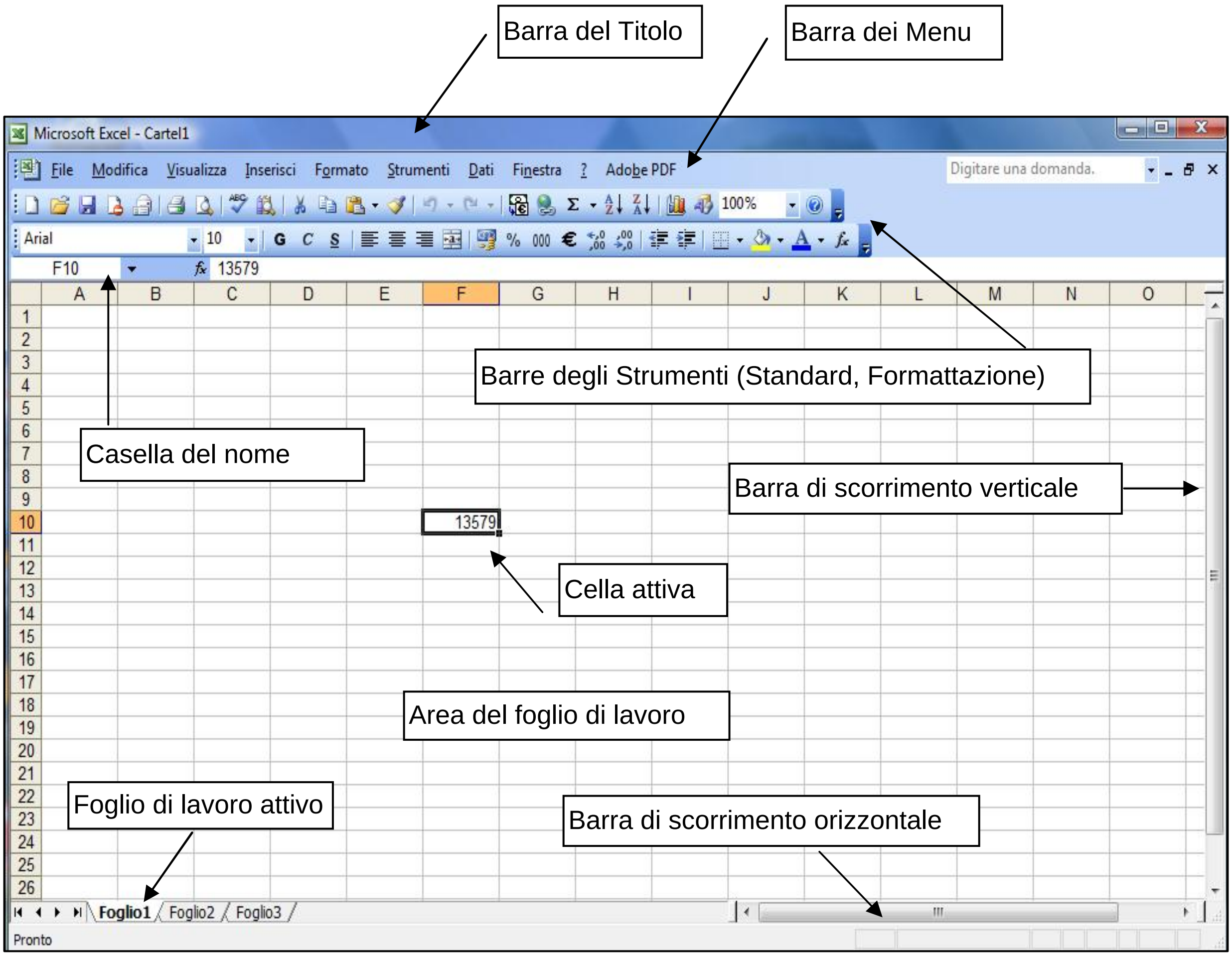
1. ESERCIZI INTRODUTTIVI

Introduzione
Il foglio di lavoro

Indice

Aprire Excel

Cliccare sul menu **Start** e nell'elenco **Programmi**, cliccare su **Microsoft Office**, poi su **Microsoft Office Excel 2003**



- Una **cartella di lavoro** è il file in cui si elaborano e si memorizzano i dati Excel
  Il file è caratterizzato dall'estensione .xls
  Ciascuna cartella di lavoro può contenere uno o più **fogli di lavoro**
- Il **foglio di lavoro** è il documento principale utilizzato in Excel per memorizzare e elaborare dati.
  Un **foglio di lavoro** è costituito da celle disposte in righe e colonne (65536 righe e 256 colonne)
  I **nomi dei fogli** sono visualizzati sulle schede poste nella parte inferiore della finestra di lavoro
  Per spostarsi da un foglio all'altro cliccare con il tasto sinistro del mouse sul nome del foglio
  Per operare sui fogli (inserire, eliminare, rinominare, spostare, ecc.) cliccare con il tasto destro del mouse sul nome del foglio
  Foglio attivo è il foglio su cui si sta lavorando; il nome è visualizzato in grassetto
- La **cella** è l'elemento fondamentale del foglio: ogni operazione fa sempre riferimento a una cella.
  La cella è individuata dall'incrocio di una riga, numerata da 1 a 65536, e di una colonna, indicata da una o due lettere dell'alfabeto

- La **casella del nome** 

  visualizza l'indirizzo (i riferimenti di colonna e riga) della cella attiva, in questo caso F10
- Le **celle di input** contengono delle costanti (numeri o testi); le **celle di output** contengono delle formule, che iniziano con il segno =, e visualizzano il risultato delle operazioni indicate

Impostazioni di base di Excel 2003

Prima di svolgere gli esercizi è opportuno predisporre lo **standard** con cui si presenta **Excel 2003**
La figura precedente mostra le **Barre degli strumenti Standard e Formattazione**, entrambe essenziali. Per visualizzarle:
Aprire il menu **Visualizza** > **Barre degli Strumenti** e nell'elenco che appare cliccare sulle due barre principali (**Standard** e **Formattazione**); l'operazione deve essere ripetuta per ciascuna barra.
In alcuni casi può essere utile la barra **Disegno**, che si attiva con la stessa procedura.
Può inoltre essere necessario modificare le impostazioni riguardanti il modo in cui vengono visualizzati numeri, date, ora, valuta
Per impostare tali opzioni:
Aprire il menu **Start**, cliccare su **Pannello di controllo**, cliccare due volte su **Opzioni Internazionali**, nella scheda Formato scegliere Italiano; con questa scelta i numeri vengono visualizzati usando come separatore decimale la virgola (se si sceglie Inglese, viene usato il punto)

Utilizzo degli esercizi

Ogni foglio di lavoro denominato Esercizio contiene uno o più esercizi proposti; nel foglio di lavoro è presente un collegamento ipertestuale, che rinvia al corrispondente foglio di lavoro Soluzione, contenente la soluzione completa dell'esercizio.
In ogni foglio di lavoro è presente un collegamento ipertestuale che rimanda all'indice.
Nei fogli degli esercizi sono spesso presenti dei suggerimenti utili per la soluzione
Il foglio Esercizi 1 contiene i riferimenti a numerosi Esempi; in ciascun Esempio sono contenuti semplici esempi ed esercizi introduttivi, che devono essere svolti prima dei successivi esercizi
Si suggerisce vivamente di svolgere gli esercizi nell'ordine indicato dal loro rispettivo numero e, nello svolgimento dell'esercizio 1, di leggere attentamente tutti gli esempi, svolgendo le operazioni richieste in ciascun esempio

Torna su

Esercizio 1

Celle e fogli di lavoro. Formule. Riferimenti relativi e assoluti. Funzioni. Collegamenti ipertestuali. Stampa

Indice

Operazioni sulle celle e sui fogli di lavoro
Selezionare celle e intervalli di celle — Esempio 1
Riferimenti di celle e intervalli di celle

Inserire dati nelle celle — Esempio 2
Modificare il contenuto di una cella
Cancellare una cella o un intervallo di celle
Eliminare una cella o un intervallo di celle

Formato celle — Esempio 3

Copia di celle — Esempio 4
Copia e incolla speciale

Creare fogli di lavoro e spostare dati — Esempio 5
Inserire, eliminare, modificare fogli di lavoro
Inserire, eliminare righe e colonne
Modificare larghezza delle colonne e altezza delle righe
Salvare il file

Formule — Esempio 6
Formule, testo

Riferimenti relativi e assoluti — Esempio 7
Formule con riferimenti assoluti e relativi
Copia di formule con riferimenti assoluti e relativi
Riferimenti circolari

Funzioni — Esempio 8
Sintassi delle funzioni, inserimento delle funzioni
Funzioni nidificate
Strumenti Analisi Dati

Collegamenti ipertestuali — Esempio 9
Collegamenti a pagine web, fogli di lavoro,
documenti, indirizzi di posta elettronica

Stampa — Esempio 10
Stampa di un foglio di lavoro, di una cartella, di un grafico

Help on line — Esempio 11
Uso della guida in linea di Excel

Torna su

Esempio 1
Selezionare celle e intervalli di celle

Indice

Ritorna Esercizio 1

Selezionare celle
Per **selezionare una cella** e renderla attiva (per modificare il contenuto):
Con il mouse: cliccare sulla cella
Con la tastiera: usare i tasti di spostamento orizzontale/verticale (frecce verso sinistra, verso destra)

Per **selezionare una riga o una colonna**: cliccare sul numero della riga o colonna corrispondente

Per **selezionare un intervallo di celle** (area rettangolare di celle):
Cliccare sulla cella in alto a sinistra dell'intervallo e trascinare con il mouse tenendo premuto il tasto sinistro, fino a raggiungere un'altra cella dell'intervallo
(tutte le celle si colorano di grigio, tranne la prima)

1	5	45	6
32	434	5	56
3	4	0	6
3	43	6	7

Altro metodo:
1. Cliccare con il mouse sulla prima cella dell'intervallo
2. Spostarsi con il mouse su un'altra cella senza cliccare
3. Premere il tasto Maiuscolo e cliccare sulla cella scelta

Attenzione: Usare il **tasto Maiuscolo** (↑) e non il tasto Blocca maiuscolo (lucchetto)
Questo metodo si può usare anche per selezionare due o più righe o colonne adiacenti.

Per **selezionare più celle disgiunte**
Cliccare sulle celle da selezionare tenendo premuto il tasto Ctrl fino al termine della selezione
Selezionare per esercizio le celle della diagonale nella tabella precedente.

Per **selezionare tutto il foglio**
Cliccare sul rettangolo nell'angolo in alto a sinistra, all'incrocio fra le intestazioni di riga e di colonna.

Riferimenti di celle e intervalli di celle
Un riferimento identifica una cella o un intervallo di celle su un foglio di lavoro.
Le colonne sono etichettate con lettere maiuscole, le righe con numeri; ogni cella è identificata dall'etichetta della colonna e della riga al cui incrocio si trova la cella

Esempi

Cella posta nella colonna G e nella riga 44	G44
Intervallo di celle della colonna G fra le righe 40 e 50	G40:G50
Intervallo di celle della riga 40 fra le colonne F e H	F40:H40
Intervallo dalla cella A1 alla cella C10	A1:C10
Tutte le celle della riga 10	10:10
Tutte le celle della colonna B	B:B
Tutte le celle delle righe da 1 a 10	1:10
Tutte le colonne da A a D	A:D

Torna su

Esempio 2
Inserimento di dati nelle celle

Indice

Ritorna Esercizio 1

Inserire dei dati nelle celle
Selezionare la cella, scrivere il dato e dare Invio (oppure usare le frecce di spostamento, oppure cliccare con il mouse su un'altra cella)
Se la cella scelta contiene già un dato, premendo un tasto si cancella il contenuto precedente:
Attenzione a non perdere inavvertitamente dei dati!
Si può ripristinare il dato perso con il comando Modifica>Annulla, o premendo il pulsante Annulla nella barra degli strumenti

Modificare il contenuto di una cella
Selezionare la cella, cliccare sulla barra della formula e scrivere le correzioni, oppure fare doppio clic sulla cella da modificare e fare le correzioni; premere Invio sulla tastiera per introdurre le modifiche, oppure il pulsante di Invio a sinistra della barra della formula
Per annullare le modifiche prima di averle inserite con Invio, premere il tasto Esc

Esempio 2.1
Inserire il numero 314 nella cella H21
Correggere il dato inserito nella cella H21, inserendo il numero 315

Una cella può contenere **Costanti** (cella di input) o **Formule** (cella di output)
Costanti:
Numeriche (automaticamente allineate e destra) o di tipo **Testo** (allineate a sinistra)
Per trattare un numero come testo: Menu Formato>Celle>Testo (vedi anche Esempio 3)
Formule: per calcolare e visualizzare risultati ottenuti con operazioni logico-matematiche su numeri; le formule devono cominciare con il simbolo = , altrimenti il contenuto è interpretato come testo

Esempio 2.2
Inserire le seguenti costanti nelle celle indicate
il numero 3,1415 nella cella H33
il testo Programma nella cella H34
il numero 250 nella cella H35 con formato testo
Inserire la formula =15+3 nella cella H36

Cancellare una cella o un intervallo di celle
Selezionare le celle da cancellare
Menu Modifica>Cancella, scegliere Tutto, Formati, ecc.
Per cancellare solo il contenuto (e non il formato o i commenti) premere il tasto Canc

Esempio 2.3
Cancellare il contenuto della cella H33

Eliminare una cella o un intervallo di celle
Selezionare le celle da eliminare
Menu Modifica>Elimina, scegliere l'opzione voluta nella finestra Elimina
Le celle circostanti (o le righe/colonne) vengono spostate secondo la scelta fatta.
Altro metodo: selezionare le celle da eliminare, premere il tasto destro e scegliere Elimina

Esempio 2.4
Eliminare la cella F52, spostando le celle a sinistra

12	13	14

Torna su

Esempio 3
Formato celle

Indice

Ritorna Esercizio 1

Dal Menu Formato>Celle si può modificare l'aspetto della cella e del suo contenuto (non il contenuto della cella, solo l'aspetto!); Il menu Formato Celle può essere anche attivato premendo Ctrl+1
La finestra di dialogo Formato>Celle presenta varie schede, ognuna contenente più opzioni:
Numero, allineamento, carattere bordo, motivo, protezione
Attivando il Menu Formato>Cella>Numero si cambia la visualizzazione del dato numerico

Formato celle: numero

Formato numero con ... cifre decimali
Nella cella G15 inserire il numero 12,345678; scegliere il formato con 2 decimali
Formato scientifico con ... cifre decimali
Nella cella G17 inserire il numero 12,345678; scegliere il formato scientifico con 4 cifre decimali
Formato valuta
Nella cella G19 inserire una cifra e scegliere il formato Valuta>Euro
Formato percentuale
Nella cella G21 inserire il numero 0,25; scegliere il formato Percentuale
(si possono anche usare i decimali)
Formato testo
Nella cella G24 inserire il testo: Esercizi con Excel; scegliere il formato Testo
Ogni stringa, anche un numero, è interpretata come testo se è preceduta da un apice ' , come nella cella G27 1245

'1245

Formato data e ora
Se si scrive 10/4 oppure 10-4 oppure 10-4-2007 in una cella, Excel usa **automaticamente** il formato data e visualizza 10-apr oppure 10/4/2007
Lo stesso accade per l'ora: scrivendo 13.5 in una cella, si visualizza 13.05 in formato data
Se per sbaglio si inserisce in una cella un valore nelle forme sopra indicate si possono ottenere **sgradevoli conseguenze**: cancellando il dato errato e inserendo un nuovo valore, non si ottiene il risultato desiderato perché si mantiene il formato data/ora della cella e ogni inserimento viene interpretato in quel formato: bisogna cambiare il formato della cella prima di inserire il nuovo dato (Menu Formato>Celle>Numero>Generale)

Formato celle: allineamento, larghezza, testo a capo, ecc.

Esempi
(Vedere suggerimenti sotto)

3.1 Adattare la la *
3.2 Adattare il contenuto alla larghezza della cella G44 permettendo che il testo vada a capo.

adattare il contenuto

3.3 Unire le celle in modo che questo testo risulti contenuto in un'unica cella

3.4 Provare i vari tipi di allineamento del testo contenuto nella cella H46 **Ciao**

Formato celle: carattere, tipo, stile, dimensione

3.5 Nella cella H50 trasformare in dimensione carattere 8, carattere grassetto | 25

Usare i pulsanti nella barra degli strumenti

Formato celle: bordi e motivo

3.6 **Inserire i bordi alla seguente tabella e riempire con diverso colore le celle contenenti numeri** pari e dispari. Bordi spessi all'esterno e sottili all'interno

Pari	1	3	5	7	9	11
Dispari	2	4	6	8	10	12

SUGGERIMENTI

Esempio 3.1: la cella C42 contiene un carattere (un asterisco), il testo inserito nella cella B42 non viene completamente visualizzato; per visualizzarlo tutto cancellare il carattere nella cella C42 oppure selezionare un numero opportuno (4 o più) di celle, poi usare il Menu Formato Formato>Celle>Allineamento>Unione celle

Esempio 3.2: Selezionare la cella G44, Menu Formato>Celle>Allineamento>Testo a capo

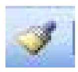
Esempio 3.3: selezionare le celle B45:G45, Menu Formato > Celle > Allineamento > Unione celle

Esempio 3.4: selezionare la cella H46, Menu Formato>Celle>Allineamento (orizzontale e verticale: provare le varie opzioni)

Esempio 3.6: Selezionare le celle B57:H58, Menu Formato>Celle>Bordi (scegliere i bordi) Selezionare le celle C57:H57; Menu Formato>Celle>Motivo, scegliere i colori Ripetere con le celle C58:H58

Copia Formato

Il formato di una cella o di un intervallo di celle può essere applicato ad altre parti del foglio con il pulsante Copia Formato nella barra degli strumenti.



Per applicare il formato della cella H46 alle celle C78:D78 selezionare la cella H46, premere il tasto Copia Formato, cliccare sulla cella C78 e trascinare fino alla cella D78

buon	giorno	buona	sera

Per copiare il formato di una cella e assegnarlo a più celle non contigue, selezionare la cella di cui si vuole copiare il formato, fare doppio clic sul pulsante Copia Formato, selezionare una per volta le celle a cui si vuole applicare il formato: la funzione Copia Formato resta attiva finchè non si preme un'altra volta il pulsante Copia Formato

Provare ad applicare di nuovo il formato della cella H46 alle celle non contigue F78 e H78

Torna su

Esempio 4
Copia di celle

Indice

Ritorna Esercizio 1

Copiare una cella o un intervallo di celle contigue.
Selezionare una cella o un intervallo di celle

Menu Modifica>Copia, oppure premere i tasti Ctrl+C, oppure usare il pulsante Copia
Selezionare la prima cella della zona in cui vuole copiare
Menu Modifica>Incolla oppure premere i tasti Ctrl+V oppure usare il pulsante Incolla

Esempio 4.1

Selezionare i seguenti numeri ed inserirli nel riquadro indicato:
1 utilizzare copia/incolla (nel riquadro A)
2 trasportare con il mouse (nel riquadro B), lasciando i dati iniziali nelle celle di partenza (Copia)
3 trasportando con il mouse (nel riquadro C) (Taglia)

1	5	45	6
32	434	5	56
3	4	0	6
3	43	6	7

A

B

C

SUGGERIMENTI

1 Selezionare le celle da copiare, premere il pulsante Copia; posizionarsi sulla prima cella in alto della zona A (cella G17), premere il pulsante Incolla
2 Selezionare le celle da copiare, posizionare il puntatore del mouse (frecce) sul bordo della selezione trascinare la selezione tenendo premuto il tasto Ctrl
3 Come al punto precedente, ma tenendo premuto il tasto ALT (oppure con Taglia e Incolla)

Esempio 4.2

Per copiare il contenuto della cella H35 nell'intervallo H37:J37
Selezionare la cella H35, Copia
Selezionare l'intervallo H37:J37, Incolla

100

Esempio 4.3

Per copiare il contenuto della cella H35 negli intervalli H40:J40 e H42:J42
Selezionare la cella H35, Copia
Premere il tasto Ctrl e selezionare le celle H40:J40 e H42:J42, premere il tasto Incolla

Esempio 4.4

Copiare il contenuto delle celle F48:F50 nell'intervallo H48:J50
Selezionare le celle F48:F50, Copia
Selezionare le celle H48:J50, Incolla

100
200
300

Copia e Incolla Speciale

Copiare la cella o l'intervallo di celle
Menu Modifica>Incolla speciale
Scegliere l'opzione fra quelle disponibili:

Tutto	incolla contenuto, formati e formule (fare attenzione alle formule!)
Formule	Incolla le formule (attenzione!)
Valori	incolla solo i valori, cioè costanti e risultati, e non le formule.
Formati	incolla solo il formato delle celle

Esempio 4.5
Incollare la cella D65 nella cella F65 con Incolla speciale, provando le varie opzioni

600	

Se i dati di una riga devono essere posti in colonna, o viceversa, si usa l'opzione **Trasponi**
Porre in riga i dati delle celle F48:F50
Selezionare le celle F48:F50, copiarle, selezionare la prima cella in cui incollarle (D72), Modifica>
Incolla speciale, scegliere l'opzione Trasponi

Torna su

Esempio 5
Pulsante Somma automatica
Creare fogli di lavoro, spostare da un foglio ad un altro

Indice

Ritorna Esercizio 1

Esempio 5.1

1 Trovare i totali entrate di ogni cassa e di ogni mese e aggiungere i bordi alla tabella
2 Creare un nuovo foglio di lavoro e chiamarlo PROVA
3 Copiare nel foglio PROVA la tabella costruita in questo foglio

		Entrate		
	Cassa 1	Cassa 2	Cassa 3	Totale entrate mese
Gennaio	€ 1.300	€ 225	€ 150	
Febbraio	€ 1.350	€ 125	€ 290	
Marzo	€ 1.279	€ 200	€ 189	
Aprile	€ 1.870	€ 158	€ 255	
Totale entrate cassa				

Totale complessivo entrate

SUGGERIMENTI

1 Per inserire i bordi usare il Menu Formato > Celle > Bordo

Pulsante somma automatica (barra Strumenti) Σ

Somme mese gennaio: selezionare le celle D16:F16 e premere il pulsante Somma automatica; procedere in modo analogo per gli altri mesi
Somme casse: selezionare le celle D20:F20 e premere il pulsante Somma automatica; procedere in modo analogo per le altre casse
Totale complessivo entrate: selezionare la cella G20 e premere il pulsante somma automatica; **controllare** che si sommino le celle G16:G19 (oppure le celle D20:F20) e dare Invio

2 Menu Inserisci > Foglio di lavoro
Per spostare il nuovo foglio di lavoro dopo il foglio Esempio 5 (o in altra posizione) premere sulla linguetta con il nome del foglio e premere il pulsante sinistro del mouse: compare il simbolo del foglio, spostarlo trascinando con il mouse nella posizione voluta
Per dare il nome Prova al foglio inserito fare doppio clic sulla linguetta e scrivere il nome Prova

3 Selezionare le celle in cui è contenuta la tabella (C14:G20), posizionare il puntatore del mouse (a forma di frecce) sul bordo delle celle, premere Alt e trascinare sulla linguetta del foglio Prova, posizionarsi sulle celle in cui si vuole copiare la tabella e rilasciare il pulsante del mouse e il tasto Alt
ATTENZIONE: in questo modo la tabella costruita in questo foglio viene **tagliata** e spostata nel foglio Prova.
Per non perdere la tabella su questo foglio, farne una copia in questo foglio prima di spostarla nel foglio Prova

Inserire, eliminare, modificare fogli di lavoro

Per inserire, eliminare, rinominare, ecc. un foglio di lavoro cliccare con il tasto destro sulla linguetta con il nome del foglio, attivare il menu di scelta rapida con il tasto destro, e scegliere l'opzione desiderata
In alternativa: Menu Inserisci>Foglio di lavoro

Inserire, eliminare righe e colonne

Inserire righe: selezionare una (o più righe) cliccando sul numero della riga (eventualmente trascinare con il mouse per selezionare più righe), cliccare con il tasto destro del mouse, Menu di scelta rapida, scegliere Inserisci
In alternativa: Menu Inserisci>Righe (inserisce una o più righe prima della/delle righe selezionate)
Procedere in modo analogo per inserire una o più colonne, e per eliminare righe e colonne

Modificare larghezza delle colonne e altezza delle righe

Per modificare la larghezza di una colonna selezionarla con il mouse, cliccare con il tasto destro, scegliere Larghezza colonne e inserire nella casella la larghezza scelta
Si può anche agire sulle intestazioni di colonna: posizionare il puntatore del mouse tra due colonne adiacenti (il puntatore assume la forma di doppia freccia), cliccare una volta e trascinare tenendo premuto il pulsante del mouse
Per adattare automaticamente la larghezza di una colonna al contenuto di una cella, cliccare due volte sul bordo destro dell'intestazione della colonna
In modo simile si modifica l'altezza di una riga.

Salvare il file

Al primo salvataggio: Menu File>Salva con nome, scegliere il nome per il file con estensione .xls e la cartella in cui salvare il file.
Per i salvataggi successivi: Menu File>Salva

Torna su

Esempio 6
Formule e testi

Indice

Ritorna Esercizio 1

Formule
Una formula inizia con il simbolo = e può contenere numeri, operatori aritmetici, riferimenti ad altre celle, funzioni di Excel.
Si possono utilizzare anche parentesi tonde, per stabilire la priorità nell'esecuzione delle operazioni.

Esempio 6.1

=7+12*2

formula contenente numeri e operatori aritmetici 31

28	14	2	8	2

formule contenenti riferimenti alle celle D15:H15 2 =D15/E15

=D15*((E15-F15)/(G15-H15)) 56

formula contenente una funzione Excel (SOMMA) 54

=SOMMA(D15:H15)

Testo
Ogni elemento che non sia una formula (che inizia con il segno =) o un numero o una data viene inserito come testo
Se si inserisce il simbolo ' (apice) per primo nella cella, tutto il contenuto viene interpretato come testo

Esempio 6.2

Inserendo 14/7 nella cella G30 si ottiene automaticamente la data
Inserendo invece '14/7 nella cella H30 si ottiene un testo

Messaggi di errore
Inserendo una formula in una cella si possono commettere errori di sintassi; l'utente viene avvisato con messaggi di errore della presenza degli errori che il sistema non è in grado di correggere in modo automatico.
I messaggi di errore cominciano con il simbolo # e finiscono con il punto esclamativo o interrogativo. I messaggi predefiniti sono i seguenti

#DIV/0!	divisione per zero
#NUM!	valore numerico non valido
#NOME?	non riconosce il testo in una formula
#RIF!	riferimento di cella non valido o cella eliminata
#VALORE!	operando o argomento di una funzione errato
#N/D?	dato per una funzione o per una formula non disponibile
#########	spazio non sufficiente per rappresentare il valore: allargare la cella

Torna su

Esempio 7
Riferimenti relativi e assoluti

Indice

Ritorna Esercizio 1

I **riferimenti relativi e assoluti** sono fondamentali per la creazione e la copia delle formule in Excel.
Un **riferimento** identifica una cella o un intervallo di celle su un foglio di lavoro.
Quando si scrivono delle formule, si fa riferimento a dati contenuti in altre celle; di solito i riferimenti alle celle o agli intervalli si basano sulla posizione relativa di questi dati rispetto alla cella contenente la formula (riferimenti relativi).

Riferimenti relativi e assoluti

Riferimento relativo: è il riferimento a una cella la cui posizione viene definita riferendosi alla cella in cui si trova la formula, ad esempio:
Somma il contenuto della cella che sta nella stessa colonna, due righe sopra, con il contenuto della cella che sta nella cella due righe sopra e due colonne a destra

12		10
22		

=C21+E21

Esempio 7.1

Formule con riferimenti relativi

La formula nella cella C37 calcola la somma dei dati contenuti nelle due celle poste nelle due righe sopra; la formula nella cella G37 calcola l'importo totale come somma dell'imponibile contenuto nella cella posta due righe sopra più l'imposta, calcolata moltiplicando l'imponibile contenuto nella cella posta due righe sopra per l'aliquota IVA contenuta nella cella posta nella riga sopra.

	prezzo
prodotto A	100
prodotto B	200
totale	300

=C35+C36

imponibile	1000
aliquata IVA	0,2
totale	1200

=G35+(G35*G36)

Copia di formule con riferimenti relativi

Quando si copia una formula, i riferimenti di cella vengono copiati di solito in modo relativo: questo significa che i riferimenti della formula incollata verranno modificati per adeguarsi alla nuova posizione della formula.

Esempio 7.2

Copia delle formule dell'esempio precedente: selezionare le celle B34:G37 e fare Copia/Incolla nella cella B50

	prezzo
prodotto A	1000
prodotto B	2000
totale	3000

=C51+C52

imponibile	3000
aliquata IVA	0,2
totale	3600

=G51+(G51*G52
)

I nomi delle celle utilizzati fino ad ora nelle formule sono chiamati **riferimenti relativi**:
riferimenti perché si riferiscono al contenuto delle celle, **relativi** perché non indicano una cella fissata del foglio, ma la distanza dalla cella in cui si scrive la formula.
Cambiando la cella della formula cambia il nome del riferimento relativo, ma non la distanza

Se si desidera che alcuni riferimenti non si modifichino con l'operazione di copia occorre usare i riferimenti assoluti.
Riferimento assoluto: è il riferimento a una cella la cui posizione è fissata e indipendente dalla cella in cui si scrive la formula
Per creare un riferimento assoluto si antepone il **simbolo dollaro $** davanti a riga e colonna del riferimento che si desidera lasciare inalterato

=D70*H70

=D70+E70

Esempio 7.3

imponibile	IVA	totale
1000	200	1200
2000	400	2400
3000	600	3600

aliquota
20%

La formula della cella F70 può essere trascinata verso il basso senza problemi: i riferimenti di riga devono cambiare; invece trascinando verso il basso la formula della cella E70 il riferimento alla cella H70 non deve cambiare (l'aliquota IVA non cambia): in questo caso si usa il riferimento assoluto anteponendo il simbolo di dollaro davanti a colonna e riga.
Si noti che è opportuno inserire il valore dell'aliquota IVA in una cella e scrivere le formule facendo riferimento (assoluto) alla cella contenente il valore dell'aliquota: in questo modo, se l'aliquota IVA viene modificata, non si devono riscrivere tutte le formule, che vengono automaticamente aggiornate.

La **differenza tra i due tipi di riferimenti** è la seguente:
I **riferimenti relativi** memorizzano la distanza dalla cella della formula, il nome della cella cambia a seconda di dove si copia o trascina la formula, sono cioè relativi alla posizione.
I **riferimenti assoluti** memorizzano il nome della cella e restano sempre uguali, non cambiano quando si copia o si trascina la formula in una nuova posizione.

Osservazione
Il problema di scegliere se usare i riferimenti assoluti e/o relativi va affrontato solo se si pensa di dover copiare la formula in un'altra posizione del foglio, ad esempio trascinando con il mouse, altrimenti può essere ignorato.
Prima di copiare una formula bisogna porsi la domanda se è necessario usare i riferimenti assoluti.

Riferimenti misti

È possibile rendere assoluto solo il riferimento della colonna o solo il riferimento della riga.
Questo serve per rendere fisso il nome della colonna e far variare il nome della riga o viceversa.
Per indicare un **riferimento misto** si deve digitare il simbolo **$** solo davanti all'intestazione della colonna o solo davanti all'intestazione della riga.
Ad esempio:

$B5 significa che la colonna B è fissa e la riga può cambiare
C$4 significa che riga 4 è fissa e la colonna può cambiare

Ricordare:
Nella copia (o trascinamento) in orizzontale cambiano solo i riferimenti di colonna
Nella copia (o trascinamento) in verticale cambiano solo i riferimenti di riga
Nella copia (o trascinamento) in diagonale cambiano sia i riferimenti di riga che di colonna

Riferimenti circolari

Si parla di **riferimento circolare** quando una formula fa riferimento, direttamente o indirettamente, alla cella che contiene il risultato della formula stessa

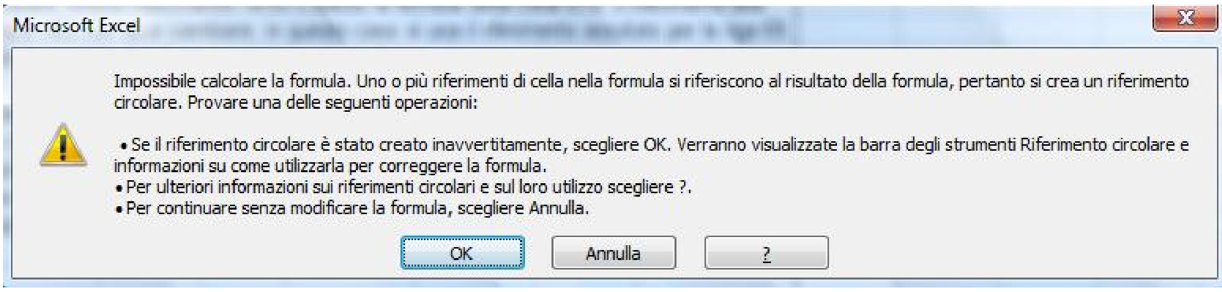
Esempi tipici

1 Inserire nella cella H113 la formula: =H113+10

2 Inserire nella cella H115 la formula: =H117/2
nella cella H116 la formula: =H115+10
nella cella H117 la formula: =H116*5

3 Inserire nella cella H119 la formula: =SOMMA(H119:H120)

Quando si instaura un riferimento circolare compare una finestra di avvertimento



Per evitare malfunzionamenti i riferimenti circolari vanno individuati e rimossi subito

Torna su

Esempio 8
Funzioni

Indice

Ritorna Esercizio 1

Excel mette a disposizione molti comandi e **funzioni** predefinite, che estendono le potenzialità del foglio elettronico.
Le funzioni sono formule predefinite, iniziano con il segno = , e seguono una sintassi.

Sintassi delle funzioni

Una funzione ha la seguente sintassi:
nome della funzione, parentesi aperta, argomenti separati da punti e virgola, parentesi chiusa

Esempio 8.1
Calcolare nella cella F16 la somma dei valori contenuti nelle celle D16:E16

4	5	9

=SOMMA(D16;E16)

Calcolare nella cella D20 la radice quadrata della somma contenuta nella cella F16

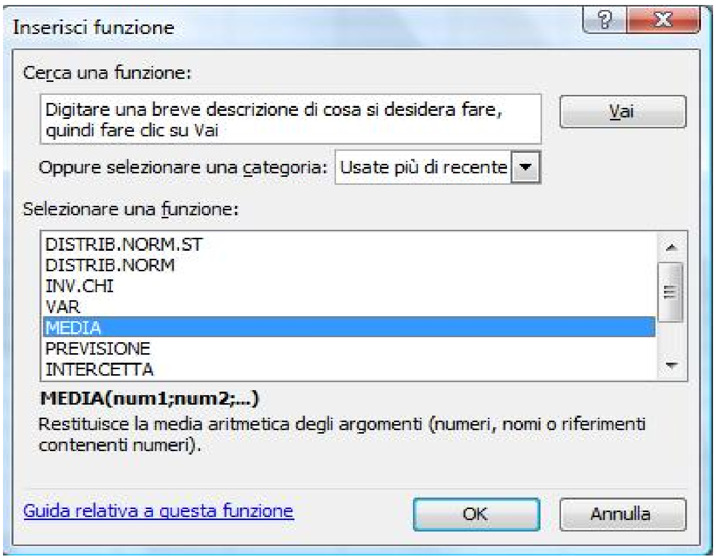
3

=RADQ(F16)

Le funzioni disponibili in Excel sono raggruppate in varie **categorie**:
Tutte, Usate più di recente, Matematiche, Statistiche, ecc.

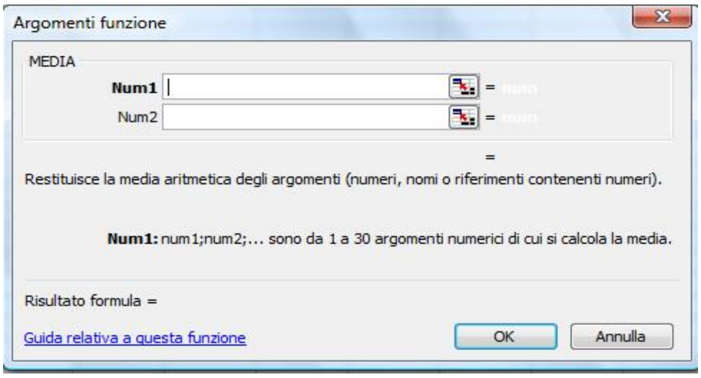
Inserimento di una funzione

Selezionare la cella di destinazione e procedere in uno dei modi seguenti:
1 **Menu Inserisci > Funzione**
2 **Barra degli strumenti**: cliccare sul pulsante Inserisci funzione 
3 **Barra della formula**: cliccare sul pulsante Inserisci funzione
Si attiva la Finestra Inserisci funzione, che guida nella scelta e composizione della funzione.



In basso compare la sintassi della funzione e una sua breve descrizione.
Cliccando sul collegamento ipertestuale 

si apre la guida in linea della funzione scelta
Cliccando due volte sul nome della funzione scelta (ad esempio la funzione MEDIA) oppure cliccando una volta sul pulsante OK, si apre la nuova finestra Argomenti funzione nella quale si inseriscono gli argomenti



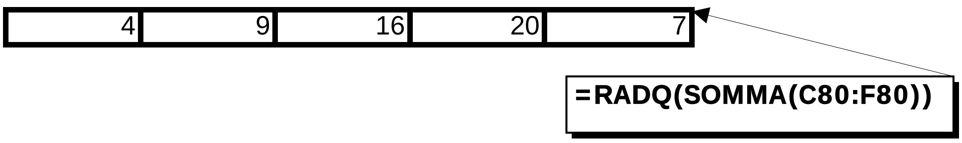
Gli argomenti possono essere costanti, valori logici come VERO o FALSO, riferimenti a una o più celle: in questo caso si selezionano con il mouse le celle dove si trovano i dati da utilizzare
I pulsanti Comprimi finestra  e Espandi finestra  posti accanto alle caselle per l'inserimento dei dati riducono e allargano la finestra Argomenti funzione

Funzioni nidificate

Si possono utilizzare le funzioni come argomento di altre funzioni
Una formula può contenere fino a 7 livelli di funzioni nidificate.

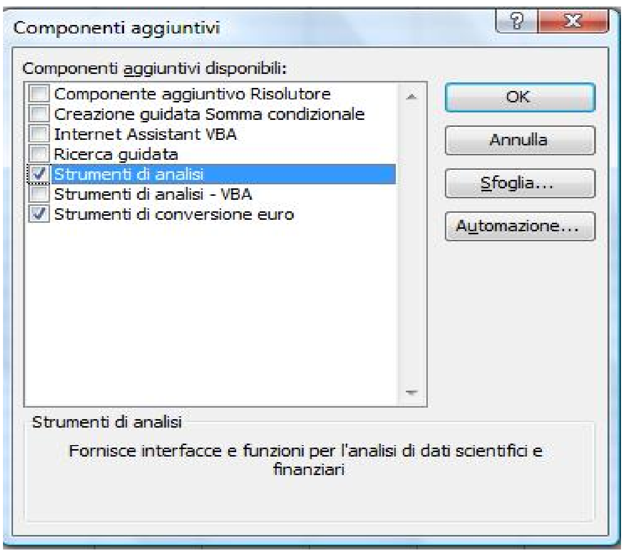
Esempio 8.2

Calcolare la radice quadrata della somma dei numeri contenuti nelle celle C80:F80



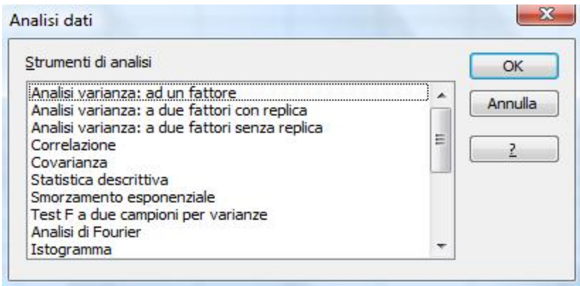
Strumenti Analisi Dati

Excel offre diverse funzioni e strumenti avanzati per l'analisi statistica dei dati.
Per verificare se sono già installati nel computer, aprire il menu Strumenti e controllare se l'opzione Analisi Dati è presente nel Menu. Se non è presente occorre installare tali strumenti.
Per installarli aprire il Menu Strumenti, cliccare su Componenti aggiuntivi, selezionare Strumenti di analisi e cliccare su OK



Nota. Se non è stata effettuata l'installazione completa del software, è necessario utilizzare il CD di installazione.

Cliccando ora su Menu Strumenti > Analisi dati, si apre la finestra Analisi dati, dove compare l'elenco di tutti gli strumenti disponibili.
L'uso di questi strumenti sarà illustrato in successivi esercizi



Torna su

Esempio 9
Collegamenti ipertestuali

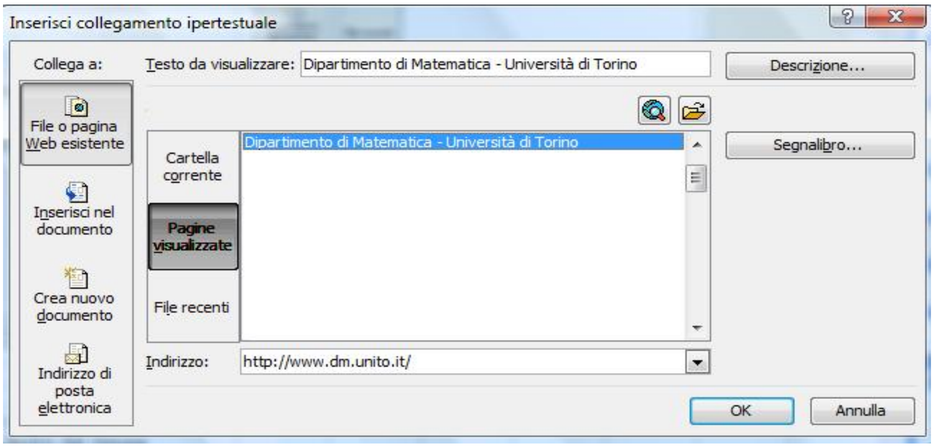
Indice

Ritorna Esercizio 1

Creare collegamenti ipertestuali
È possibile creare collegamenti ipertestuali ad altri file nel proprio sistema, sulla rete o su Internet, ad altri fogli di lavoro all'interno della cartella di lavoro corrente, a un indirizzo di posta elettronica.
La guida in linea di Excel, alla voce "Creare un collegamento ipertestuale" fornisce dettagliate spiegazioni. In sintesi:
Selezionare la cella dove comparirà il collegamento, scrivere un testo (non è indispensabile) e premere il tasto destro; scegliere Collegamento ipertestuale
Al termine il testo verrà visualizzato in blu e sottolineato.
In alternativa disegnare o inserire un oggetto grafico, ad esempio un pulsante, che potrà assumere una forma qualsiasi e impostarlo come collegamento ipertestuale.
In entrambi i casi, facendo clic sul testo o sull'oggetto grafico, si passerà al file o al percorso di destinazione.

Esempio 9.1

1 **Collegamento a una Pagina Web**: la finestra seguente mostra il collegamento alla pagina web del Dipartimento di Matematica.

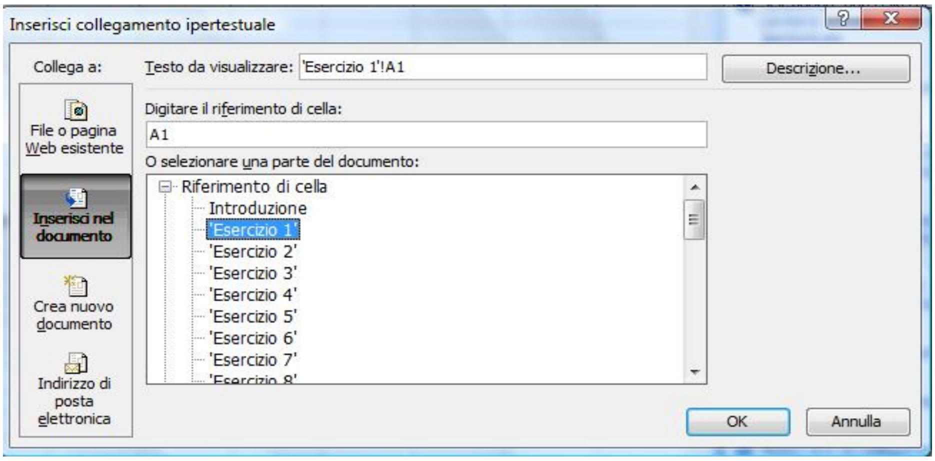


Dipartimento di Matematica - Università di Torino

Per realizzare il collegamento a questa pagina procedere nel modo seguente:
selezionare la cella in cui creare il collegamento e premere il tasto destro del mouse
scegliere fra le pagine visualizzate la pagina a cui ci si vuole collegare e premere Ok;
se la pagina non compare nell'elenco, premere il tasto Esplora il Web, 
nella Barra degli strumenti Web
Collegarsi alla pagina voluta, lasciare il browser aperto, tornare alla finestra Inserisci collegamento ipertestuale e premere Ok

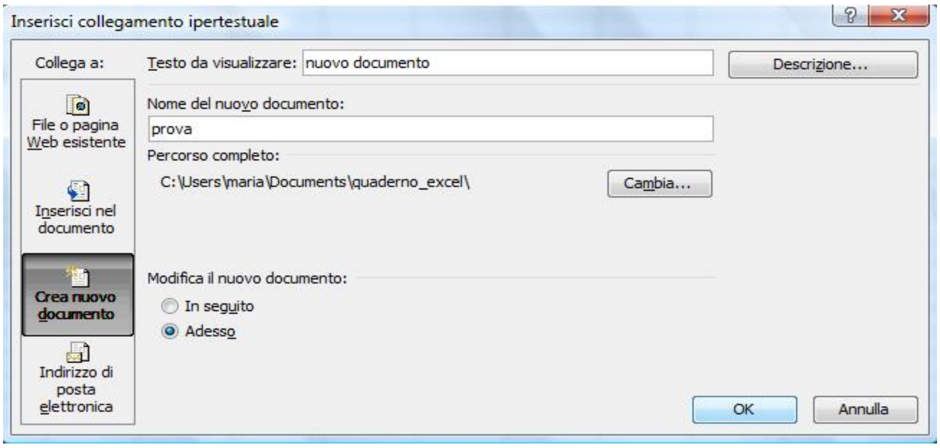
I collegamenti seguenti si realizzano in modo analogo.

2 **Collegamento a un foglio di lavoro nel documento**: la finestra seguente mostra il collegamento al foglio Esercizio 1 in questo documento; il collegamento è realizzato con due simboli diversi, il testo (Esercizio 1) e la freccia: cliccando sul simbolo si apre il collegamento.



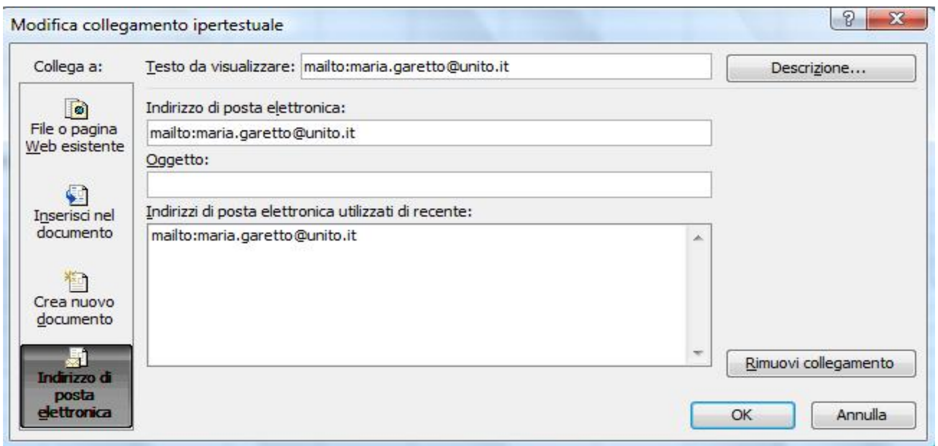
[Esercizio 1]

3 **Collegamento a un nuovo documento**: la finestra seguente mostra il collegamento al nuovo documento di nome prova



[nuovo documento]

4 **Collegamento a un indirizzo di posta elettronica**: la finestra seguente mostra il collegamento all'indirizzo e-mail indicato



[mailto:maria.garetto@unito.it]

Modificare o rimuovere collegamenti ipertestuali

Per modificare o rimuovere un collegamento, selezionare con il tasto destro il collegamento,
scegliere Modifica collegamento ipertestuale (per fare delle modifiche)
oppure Rimuovi collegamento ipertestuale (per rimuoverlo)

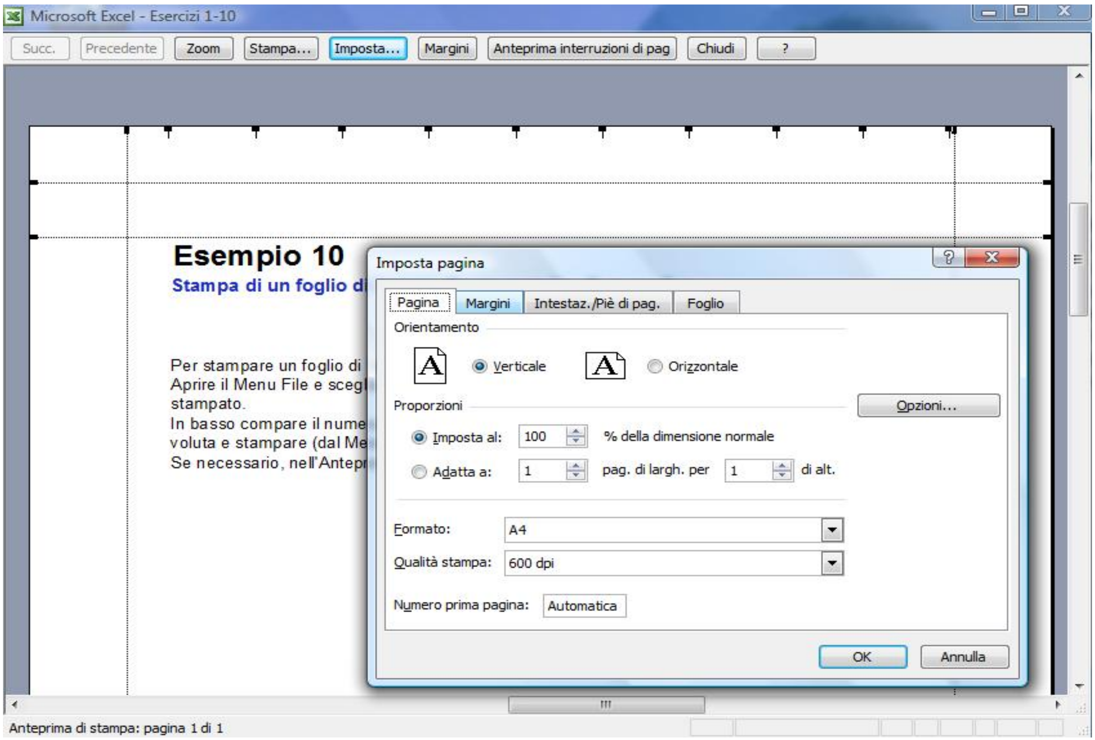
Torna su

Esempio 10
Stampa di un foglio di lavoro, di una cartella, di un grafico

Indice

Ritorna Esercizio 1

Prima di **stampare un foglio di lavoro** è consigliabile fare un'anteprima:
Aprire il Menu File e scegliere Anteprima di stampa: si visualizza il foglio nella forma in cui sarà stampato.
L'anteprima si può anche ottenere premendo il pulsante Anteprima di stampa  nella barra degli strumenti Standard
In basso compare il numero di pagine che saranno stampate; controllare che sia tutto nella forma voluta e stampare (dal Menu File o direttamente dall'anteprima, premendo il pulsante Stampa)
Se necessario, nell'Anteprima si possono modificare le impostazioni della pagina,margini, ecc. agendo sulle schede di Imposta pagina



Per **stampare una cartella di lavoro** completa, fare prima l'anteprima di controllo di tutti i fogli, poi dal Menu File, scegliere Stampa e selezionare Stampa tutta la cartella

Per **stampare** solo **una parte di un foglio di lavoro**, selezionare con il mouse la parte da stampare, aprire il Menu File e scegliere Area di stampa>Imposta area di stampa, poi fare l'anteprima e stampare

Per **stampare un grafico**, selezionarlo con il mouse, fare l'anteprima di controllo e stampare

Torna su

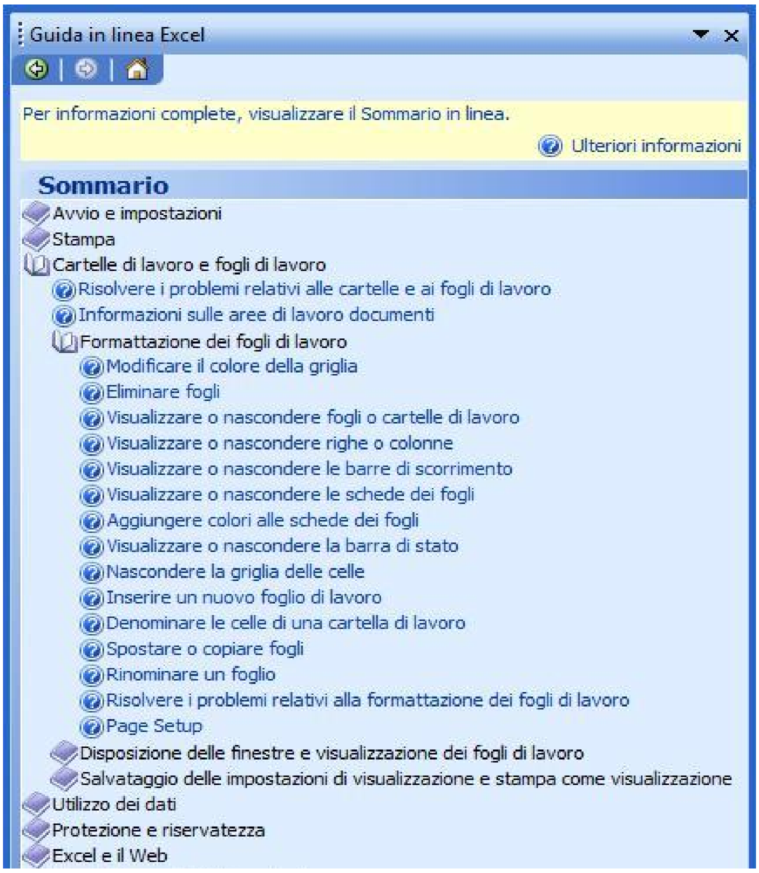
Esempio 11
Help on line

Indice

Ritorna Esercizio 1

Si può accedere alle funzioni di help cliccando sul pulsante **?** nella barra dei menu oppure premendo il tasto F1; l'uso della guida è molto intuitivo.
Nella finestra della **Guida in linea Excel** si può effettuare una ricerca per individuare le parti della guida che trattano l'argomento a cui si è interessati.
Ad esempio digitando nella casella **Cerca**: la parola Somma, e cliccando sul tasto 
si apre la finestra con i risultati della ricerca
Cliccando sul paragrafo che interessa si apre un'ulteriore pagina della guida contenente le informazioni richieste

Si può accedere all'indice della guida cliccando su **Sommario**
Nel sommario gli argomenti sono suddivisi in varie ampie aree, rappresentate da un'icona a forma di libro, che si apre e si chiude cliccando sull'icona
Cliccando sui titoli si aprono le pagine della guida



Torna su

Soluzione Esercizio 2
Le funzioni matematiche più semplici

Indice

Ritorna Esercizio 2

Esercizio 2.1

1 Nelle celle D29:D42 scrivere in colonna i seguenti testi
a ; b ; Somma ; Differenza ; Prodotto ; Quoziente ; Parte intera del quoziente ; Resto ;
Radice quadrata di a ; Radice cubica di a; a^2 ; Log10 a ; ln a; Massimo (a,b)
Allargare le colonne in modo da adattarle al testo

2 Inserire i due numeri 120 e 32 rispettivamente nelle celle E29 ed E30

3 Nelle celle E31:E34 inserire le formule per calcolare i risultati delle operazioni aritmetiche indicate dal testo nella cella alla sinistra.
Nota: Per calcolare la somma si può usare l'operatore + oppure la funzione SOMMA.

4 Nelle celle E35:E42 usare le funzioni predefinite INT RESTO RADQ LOG10 LN MAX

INT	calcola la parte intera di un numero
RESTO	calcola il resto della divisione fra due numeri
RADQ	calcola la radice quadrata di un numero
POTENZA	calcola la potenza di un numero
	Esempio: per calcolare la radice cubica di un numero usare l'esponente 1/3
LOG10	calcola il logaritmo in base 10 di un numero
LN	calcola il logaritmo naturale di un numero
MAX	calcola il massimo di un insieme di numeri

Vedere la guida in linea per eventuali informazioni sulle funzioni

5 Provare a cambiare i due numeri nelle celle e osservare che si aggiornano automaticamente i risultati: uno dei vantaggi più significativi del foglio elettronico è il **ricalcolo automatico** del risultato di tutte le formule, quando viene modificato il contenuto di una cella.

a	120
b	32
Somma	152
Differenza	88
Prodotto	3840
Quoziente	3,75
Parte intera del quoziente	3
Resto	24
Radice quadrata di a	10,954
Radice cubica di a	4,932
a^2	14400
Log10 a	2,079181
ln a	4,787492
Massimo(a,b)	120

Torna su

Soluzione Esercizio 3
Ordinamento e calcolo di massimo e minimo

Indice

Ritorna Esercizio 3

Esercizio 3.1

1 Inserire dei numeri qualsiasi nelle celle comprese tra C14 e C24.
2 Copiarli e ordinarli in modo decrescente nelle celle da D14 a D24 e in modo crescente nelle celle da E14 a E24
3 Nelle celle F14 e G14 calcolare il massimo e il minimo dei dati

Dati	Dati ordinati in modo decrescente	Dati ordinati in modo crescente	Massimo	Minimo
23	76	5	76	5
34	45	6		
23	45	12		
21	34	21		
6	23	21		
5	23	23		
45	21	23		
76	21	34		
12	12	45		
21	6	45		
45	5	76		

SUGGERIMENTI

2 Copiare i dati nella colonna D e usare il pulsante Ordinamento decrescente 

Copiare i dati nella colonna E e usare il pulsante Ordinamento crescente 

3 Attenzione: una volta ordinati i valori, è facile copiare il massimo e il minimo in una nuova cella. Esistono però le funzioni MAX e MIN che individuano il massimo e il minimo. Provare a utilizzarle!

Ordinamento di un elenco con più colonne

Esercizio 3.2

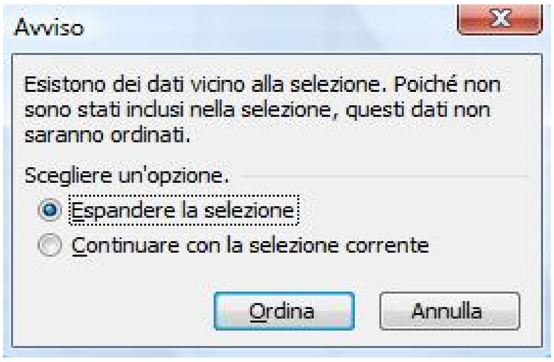
Ordinare in ordine alfabetico l'elenco della tabella seguente
Attenzione: le persone devono mantenere il proprio numero telefonico!

Fornitori	N° telefonico
Rossi	011 2345678
Bianchi	02 43657687
Verdi	0131 3465789
Neri	0171 3344675

Fornitori	N° telefonico
Bianchi	02 43657687
Neri	0171 3344675
Rossi	011 2345678
Verdi	0131 3465789

SUGGERIMENTI

1 Copiare la tabella nelle celle F41:G45 (per poter fare il confronto tra elenco disordinato e elenco ordinato)
2 Selezionare le celle con i nomi dei fornitori e premere il pulsante Ordinamento crescente
Fare attenzione alle opzioni proposte nella finestra Avviso!



Osservazione
Selezionando entrambe le colonne si ottiene lo stesso risultato in modo più semplice!

Torna su

Soluzione Esercizio 4

Somma di numeri. Funzionalità Copia/Incolla

Indice

Ritorna Esercizio 4

Esercizio 4.1

1 Inserire i primi cinque numeri pari nelle celle C16:C20
2 Nella cella C21 calcolare la somma usando il pulsante somma automatica.
3 Cambiare alcuni numeri nella colonna C e verificare che il totale si aggiorna automaticamente.
4 Copiare i dati della colonna C nella colonna H usando i pulsanti Copia e Incolla.
5 Nella cella D16 inserire una formula che moltiplichi per 2 il valore nella cella C16 e sommi 3.
6 Copiare la formula dalla cella D16 alla cella D20 trascinando con il mouse.
7 Copiare i dati della colonna D nella colonna E poi i dati della colonna E nella colonna F, trascinando con il mouse: verificare che usando i riferimenti relativi le formule e i risultati cambiano

	2	7	17	37	2
	4	11	25	53	4
	6	15	33	69	6
	8	19	41	85	8
	10	23	49	101	10
somma	30				

SUGGERIMENTI

2 Selezionare la cella C21, premere il pulsante Somma automatica
Excel propone in modo automatico l'intervallo di celle da sommare: se non sono quelle volute (in questo caso lo sono), puntare con il mouse sulla prima cella e trascinare fino all'ultima cella da sommare, poi premere Invio
Se le celle da sommare non sono contigue, selezionare le celle tenendo premuto il tasto Ctrl
4 Usare il Tasto Esc per togliere il tratteggio intorno alle celle copiate dopo aver eseguito la copia
5 La formula da utilizzare è =C16*2+3
6 Selezionare la cella D16; spostare il puntatore del mouse nell'angolo in basso a destra; quando il puntatore ha forma di croce nera premere il pulsante sinistro del mouse e trascinare verso il basso fino alla cella D20
7 Selezionare le celle da D16 a D20 e trascinare con il mouse nelle celle da E16 a E20;
Ripetere selezionando le celle E16:E20 trascinando nelle celle da F16 a F20
(puntatore del mouse a forma di croce nera)

Esercizio 4.2

Conoscendo il peso (in kg) e l'altezza (in cm) di alcuni soggetti (tabella 1), calcolare il loro "body mass index" espresso dala seguente formula

$$b.m.i. = 10000 \cdot \frac{peso}{altezza^2}$$

Costruire la formula usando i riferimenti relativi per le celle e usare il trascinamento

Tabella 1

peso	altezza	b.m.i.
66	170	22,84
53	156	21,78
94	175	30,69
97	165	35,63
61	170	21,11
70	168	24,80
52	154	21,93

Torna su

Soluzione Esercizio 5

Trascinare elenchi e successioni, uso di copia/incolla

Indice

Ritorna Esercizio 5

Esercizio 5.1

1 A partire dalla cella C17 costruire nella colonna C la successione dei primi dieci numeri pari.
Nella colonna D costruire la successione dei numeri 2, 6, 10,….
Utilizzare la funzione di trascinamento del mouse
2 Spostare tutte le celle appena create nelle colonne E e F trascinandole con il mouse.
Che cosa accade ai valori contenuti nelle celle? Perché?
3 Copiare (con Copia/incolla) i valori delle colonne C e D nelle colonne H,I: i valori cambiano?
4 Copiare il contenuto delle celle delle colonne C e D nelle righe 29 e 30

2	2	2	2
4	6	8	10
6	10	14	18
8	14	20	26
10	18	26	34
12	22	32	42
14	26	38	50
16	30	44	58
18	34	50	66
20	38	56	74

2	2
4	6
6	10
8	14
10	18
12	22
14	26
16	30
18	34
20	38

2	4	6	8	10	12	14	16	18	20
2	6	10	14	18	22	26	30	34	38

SUGGERIMENTI

1 Scrivere nelle celle C17, C18 i primi due numeri della successione, poi selezionare entrambe
le celle e trascinare premendo il pulsante sinistro del mouse (croce nera) fino alla cella C26
Scrivere nelle celle D17, D18 i numeri 2 e 6, poi ripetere come prima
2 Selezionare le celle C17:D26 e trascinarle con il mouse (croce nera) nelle colonne accanto
Nell'operazione di Copia/Incolla viene mantenuto il passo fra i numeri, sia in verticale che in
orizzontale
3 Selezionare le celle C17:D26 usare Copia/Incolla nelle colonne H,I (i numeri non cambiano, non ci
sono formule)
4 Selezionare le celle C17:D26, premere il tasto Copia; selezionare le celle B29, B30,
aprire il menu Modifica, sceglliere Incolla Speciale, selezionare Trasponi e dare OK
(alla fine premere Esc per eliminare il tratteggio attivo intorno alle celle da copiare)

Esercizio 5.2

Costruire una tabella che contenga:
1 nella prima colonna i primi dieci numeri interi
2 nella seconda colonna i primi dieci multipli del numero 20
3 nella terza colonna i quadrati dei primi dieci multipli del numero 15
4 nella quarta colonna la differenza fra i valori corrispondenti della terza e della seconda colonna
5 nella quinta colonna la radice cubica dei reciproci dei valori della quarta colonna
(Per calcolare la radice cubica si può usare la funzione POTENZA)

1	20	225	205	0,1696
2	40	900	860	0,1052
3	60	2025	1965	0,0798
4	80	3600	3520	0,0657
5	100	5625	5525	0,0566
6	120	8100	7980	0,0500
7	140	11025	10885	0,0451
8	160	14400	14240	0,0413
9	180	18225	18045	0,0381
10	200	22500	22300	0,0355

Torna su

Soluzione Esercizio 6
Uso dei riferimenti assoluti e relativi. Percentuali

Indice

Ritorna Esercizio 6

Esercizio 6.1

1 Scrivere 10 numeri qualsiasi nelle celle C16:C25
2 Nella cella C26 calcolarne la somma (pulsante SOMMA AUTOMATICA)
3 Nella cella D16 scrivere una formula che calcoli la percentuale di C16 rispetto al totale (C26)
e trascinare la formula in basso fino alla cella D27
4 Dare il formato percentuale alle celle D16:D26
5 Nella cella D26 verifcare che la somma delle percentuali è 100%

	12	5%
	15	6%
	35	15%
	14	6%
	6	3%
	4	2%
	22	9%
	45	19%
	57	24%
	30	13%
somma	240	100%

SUGGERIMENTI

3 Per calcolare la percentuale dividere il contenuto della cella C16 per la somma C26
Nella formula copiata nelle celle sottostanti (trascinando verso il basso con il mouse) il numeratore deve cambiare (riferimento relativo), mentre il denominatore deve restare invariato (riferimento assoluto). Scrivere quindi nella cella D16 la formula =C16/C26 e premere Invio
Selezionare la cella D16 e trascinare fino alla cella D26
Per rendere assoluto un riferimento di cella si deve inserire il simbolo $ davanti alla riga e alla colonna; si può anche procedere nel modo seguente: nella formula (cella D16) scrivere il riferimento alla cella C26 e premere il tasto F4
4 Per ottenere lo stile percentuale selezionare le celle D17:D27 e premere il pulsante


Stile percentuale nella barra Formattazione
oppure usare il Menu Formato>Stile>Stile percentuale

Esercizio 6.2

Scrivere i primi dieci termini della successione di Fibonacci.
I primi due termini sono uguali a 1, ogni termine successivo è uguale alla somma dei due termini precedenti:
1 1 2 3 5 ecc.

Successione di Fibonacci

1	1	2	3	5	8	13	21	34	55

SUGGERIMENTI

Inserire i primi due numeri nelle celle B48 e C48;
nella cella D48 calcolare la somma delle celle B48 e C48 (riferimenti relativi);
selezionare la cella D48 e trascinare fino alla cella K48

Torna su

Soluzione Esercizio 7

Costruzione della Tavola Pitagorica. Riferimenti misti

Indice

Ritorna Esercizio 7

Esercizio 7.1

Costruire la tavola pitagorica con le formule.

=C$12*$B13

	1	2	3	4	5	6	7	8	9	10
1	1	2	3	4	5	6	7	8	9	10
2	2	4	6	8	10	12	14	16	18	20
3	3	6	9	12	15	18	21	24	27	30
4	4	8	12	16	20	24	28	32	36	40
5	5	10	15	20	25	30	35	40	45	50
6	6	12	18	24	30	36	42	48	54	60
7	7	14	21	28	35	42	49	56	63	70
8	8	16	24	32	40	48	56	64	72	80
9	9	18	27	36	45	54	63	72	81	90
10	10	20	30	40	50	60	70	80	90	100

SUGGERIMENTI

1 Scrivere subito la prima riga (riga 12) e la prima colonna (colonna B) della tavola:
inserire i primi due numeri 1 e 2 nelle celle C12 e D12 e trascinare con il mouse lungo la riga;
analogamente inserire 1 e 2 nelle celle B13 e B14 e trascinare lungo la colonna

2 Scrivere la formula nella cella C13, tenendo conto di quali elementi rendere assoluti.
Usare i riferimenti misti!
La formula da scrivere è =C$12*$B13.
Copiare la formula su tutta la tabella

3 **Primo modo:**
Per copiare la formula su tutta la tabella in una volta sola:
Selezionare la cella C13, premere il pulsante Copia
Cliccare sulla cella L22 premendo il tasto Maiuscolo (per selezionare l'intera tabella), poi premere il pulsante Incolla
Secondo modo:
Selezionare la cella C13 e trascinare con il mouse lungo la colonna fino alla cella C22,
selezionare la colonna C13:C22 e trascinare con il mouse verso destra fino alla colonna L

Torna su

Soluzione Esercizio 8

Tabelle. Formule con riferimenti assoluti e relativi

Indice

Ritorna Esercizio 8

Esercizio 8.1

1 Creare una tabella per riassumere le spese di viaggio di un rappresentante nel mese di gennaio

10 gennaio	Spese Viaggio a Bergamo (110 Km) più € 40 (Altre Spese)
12 gennaio	Spese Viaggio a Brescia (210 Km) più € 30 (Altre Spese)
18 gennaio	Spese Viaggio a Como (78 Km) più € 35 (Altre Spese)
20 gennaio	Spese Viaggio a Bergamo (110 Km) più € 45 (Altre Spese)
22 gennaio	Spese Viaggio a Varese (128 Km) più € 45 (Altre Spese)

Il rimborso chilometrico è di 0,5 € al km.
Usare le seguenti formule:
Rimborso Viaggio = Km * Rimborso Chilometrico
Rimborso Totale = Rimborso Viaggio + Altre Spese
Rimborso del mese = Somma dei Rimborsi Totali

Nelle formule usare in modo appropriato i riferimenti relativi e assoluti
Usare il formato testo nelle celle destinate alla data

2 Inserire un titolo (Nota Spese) prima della tabella.
3 Scegliere uno sfondo colorato per le celle Rimborso del Mese.
4 Inserire un nuovo viaggio con i seguenti dati:
data **24 gennaio**
destinazione **Como**
Km **78**
Altre Spese € **35**
5 Usare nelle celle interessate il formato euro.

Inserire il titolo

		Nota Spese			
Data	**Destinazione**	**Km**	**Rimborso Viaggio**	**Altre Spese**	**Rimborso Totale**
10 gennaio	Bergamo	110	€ 55,00	€ 40,00	€ 95,00
12 gennaio	Brescia	210	€ 105,00	€ 30,00	€ 135,00
18 gennaio	Como	78	€ 39,00	€ 35,00	€ 74,00
20 gennaio	Bergamo	110	€ 55,00	€ 45,00	€ 100,00
22 gennaio	Varese	128	€ 64,00	€ 45,00	€ 109,00
24 gennaio	Como	78	€ 39,00	€ 35,00	€ 74,00
Rimborso/Km	€ 0,50			Rimborso del mese	€ 587,00

Inserire qui il rimborso chilometrico

Formula per il calcolo del rimborso mensile

Torna su

Soluzione Esercizio 9

Uso delle formule

Soluzione dell'equazione di secondo grado ax^2+bx+c=0

Indice

Ritorna Esercizio 9

Esercizio 9.1

1 Nella cella D17 scrivere il testo COEFFICIENTI e nella cella D22 il testo SOLUZIONI
Nelle celle D18, D19, D20 scrivere rispettivamente a , b , c
Nelle celle D23 e D24 scrivere rispettivamente x1 e x2
Nelle celle E18, E19, E20 scrivere i valori dei coefficienti dell'equazione
a = 1, b = –3, c = 2
2 Nelle celle E23 e E24 scrivere la formula risolutiva
3 Variare il valore dei coefficienti e verificare che si ottiene la soluzione della nuova equazione

COEFFICIENTI	
a	1
b	-3
c	2

SOLUZIONI	
x1	1
x2	2

SUGGERIMENTI

La formula risolutiva è

$$x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Scrivere i coefficienti della formula usando i riferimenti alle celle che li contengono.
La radice quadrata si calcola con la funzione RADQ; inserire fra parentesi come argomento l'espressione di cui si vuole calcolare la radice quadrata.

Torna su

Soluzione Esercizio 10
Progressioni aritmetiche e geometriche

Indice

Ritorna Esercizio 10

Esercizio 10.1

1 Costruire una tabella che contenga una progressione aritmetica con primo elemento a_0=3 e ragione q=2. L' elemento generico è $a_n=a_{n-1}+q$
Utilizzare due celle per a_0 e q e poi ricavare la progressione mediante formule.
Verificare che cambiando i valori di a_0 e q cambiano tutti i valori della progressione.

2 Aggiungere una colonna che contenga una progressione geometrica
Il generico elemento è $a_n=a_{n-1}*q$
Attenzione ai riferimenti relativi e assoluti!

		Progressione aritmetica	**Progressione geometrica**
Primo elemento	3	3	3
Ragione	2	5	6
		7	12
		9	24
		11	48
		13	96
		15	192
		17	384
		19	768
		21	1536
		23	3072
		25	6144
		27	12288
		29	24576
		31	49152
		33	98304
		35	196608
		37	393216
		39	786432

Torna su

2. GRAFICI

Soluzione Esercizio 11

Realizzazione di grafici: istogrammi, diagrammi circolari, grafici a dispersione, grafici a linee

Indice

Ritorna Esercizio 11

I grafici di base in Excel sono dei seguenti tipi:
istogrammi e diagrammi a barre
diagrammi circolari (a torta)
grafici a dispersione (grafici cartesiani)
grafici a linee
Istogrammi e **diagrammi a barre** si usano per rappresentare una frequenza sull'asse Y e dati qualitativi sull'asse X; per costruire questi grafici si seleziona un gruppo di celle (frequenze) e successivamente si selezionano le celle da indicare come etichette sull'asse X
Nei **diagrammi circolari** (**diagrammi a torta**) le fette sono proporzionali alle frequenze; non è necessario calcolare le frequenze percentuali: selezionando le frequenze, il programma calcola automaticamente la suddivisione percentuale necessaria per costruire le fette della torta
I **grafici a linee** sono indicati per descrivere l'andamento di un fenomeno nel tempo
In un grafico a linee i punti sull'asse orizzontale devono essere equamente distanziati.
I **grafici a dispersione** sono i noti diagrammi cartesiani, costruiti su due assi ortogonali su cui si fissa l'origine e l'unità di misura.
Per disegnare i diagrammi a dispersione occorre predisporre una tabella delle ascisse e delle corrispondenti ordinate della funzione, in un dato intervallo e con un opportuno incremento

Esercizio 11.1

La tabella 1 rappresenta il numero di studenti iscritti ai 5 anni di corso di un istituto superiore
Realizzare un **istogramma** della distribuzione di frequenza assoluta
Realizzare un **diagramma circolare** delle frequenze percentuali

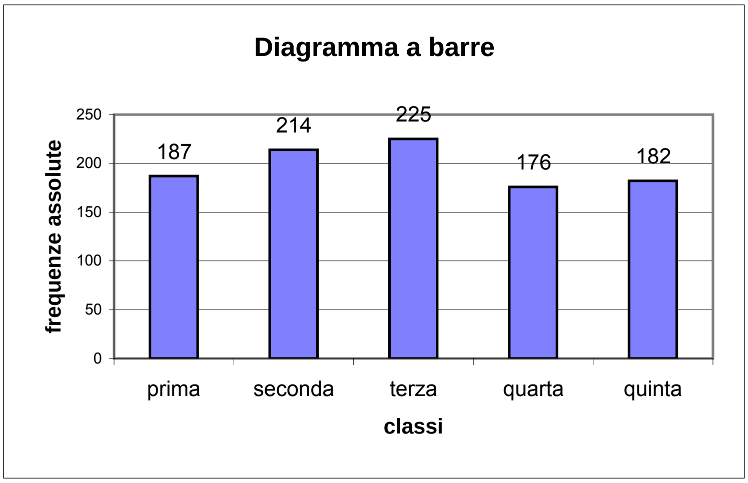
Tabella 1

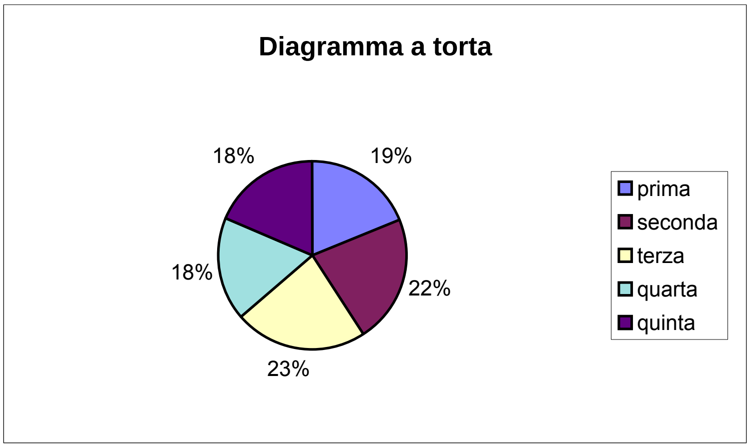
classi	frequenza assoluta
prima	187
seconda	214
terza	225
quarta	176
quinta	182
totale	984

SUGGERIMENTI

Selezionare la colonna delle frequenze assolute (attenzione a non selezionare il totale), cliccare sul pulsante **Creazione guidata Grafico**

Selezionare Tipo di grafico>Istogramma; seguire le istruzioni della finestra Creazione guidata grafico cliccando successivamente su Avanti
Nella scheda Serie (Passaggio 2), nella casella Etichette asse categorie X selezionare come etichette le celle contenenti i nomi delle classi (prima, seconda, ecc.)
Nel Passaggio 3 operare sulle varie schede inserendo titolo, nomi degli assi, ecc..
Nella scheda Etichette dati scegliere Valore per mostrare le frequenze assolute sul grafico
Nell'ultimo passaggio si può scegliere se il grafico deve essere collocato su un nuovo foglio di lavoro o su uno già esistente: la proposta di default è il foglio attivo
Al termine della realizzazione del grafico posizionarlo sul foglio trascinando con il mouse, ridimensionarlo cliccando sul bordo del grafico e agendo sui quadratini neri presenti sul bordo
Per realizzare il diagramma circolare seguire lo stesso procedimento del punto 1, selezionando la colonna delle frequenze assolute, poi scegliere Tipo di grafico>Torta
Nella scheda Serie (Passaggio 2), nella casella Etichette categorie selezionare come etichette le celle contenenti i nomi delle classi (prima, seconda, ecc.)
Nella scheda Etichette dati (passaggio 3) scegliere Percentuale per mostrare le frequenze percentuali sul grafico





Esercizio 11.2

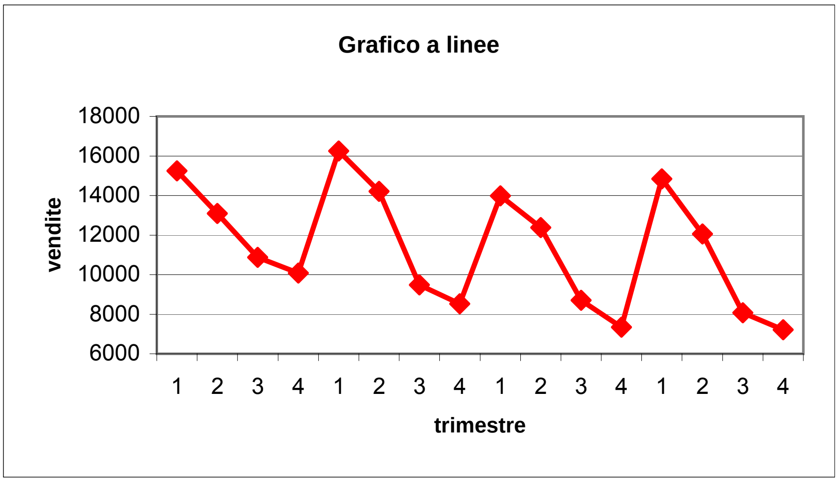
La tabella 2 rappresenta le vendite trimestrali di un prodotto (numero di confezioni vendute) negli anni 2001-2004.
Illustrare l'andamento delle vendite con un **grafico a linee**.

Tabella 2

Trimestre	Vendite
1	15239
2	13089
3	10879
4	10084
1	16257
2	14217
3	9478
4	8519
1	13981
2	12373
3	8711
4	7350
1	14838
2	12062
3	8070
4	7213

SUGGERIMENTI

Selezionare la colonna delle frequenze assolute (celle delle Vendite), cliccare sul pulsante Creazione guidata grafico
Selezionare Tipo di grafico>Linee; seguire le istruzioni della finestra Creazione guidata grafico cliccando successivamente su Avanti
Nella scheda Serie (Passaggio 2), nella casella Etichette asse categorie X selezionare come etichette le celle dei trimestri



Esercizio 11.3

Tracciare il grafico (**diagramma a dispersione**) della funzione y=x^2 nell'intervallo (-3,3)

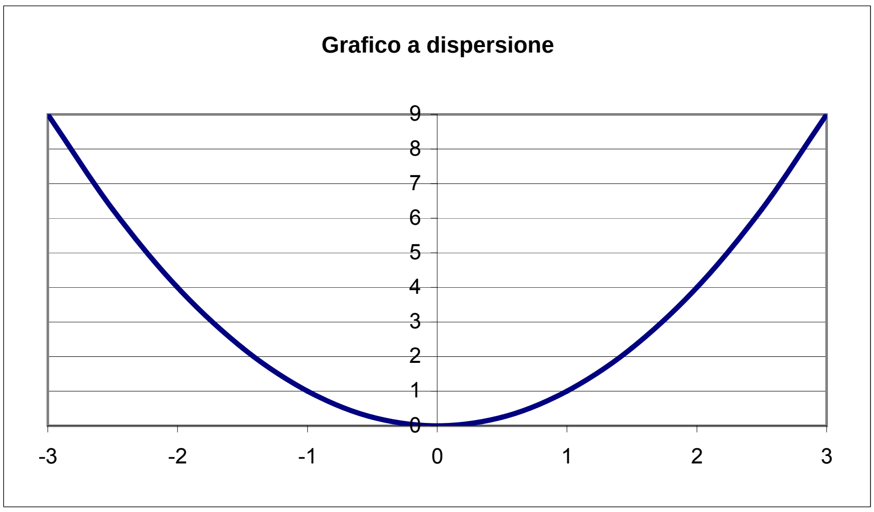
Tabella 3

x	f(x)
-3	9
-2,5	6,25
-2	4
-1,5	2,25
-1	1
-0,5	0,25
0	0
0,5	0,25
1	1
1,5	2,25
2	4
2,5	6,25
3	9

SUGGERIMENTI

Costruire la tabella 3: la prima colonna contiene i valori di x (ascisse) equidistanti, nell'intervallo (-3,3) con incremento h=0,5; la seconda colonna contiene le ordinate, ossia i corrispondenti valori della funzione y=x^2 nei punti x.

Selezionare insieme la colonna delle ascisse e quella delle ordinate, cliccare sul pulsante Creazione guidata grafico, e selezionare Tipo di grafico>Dispers. (XY); seguire le istruzioni della finestra Creazione guidata grafico, cliccando successivamente su Avanti



Torna su

Soluzione Esercizio 12
Grafico a barre orizzontali

Indice

Ritorna Esercizio 12

Esercizio 12.1

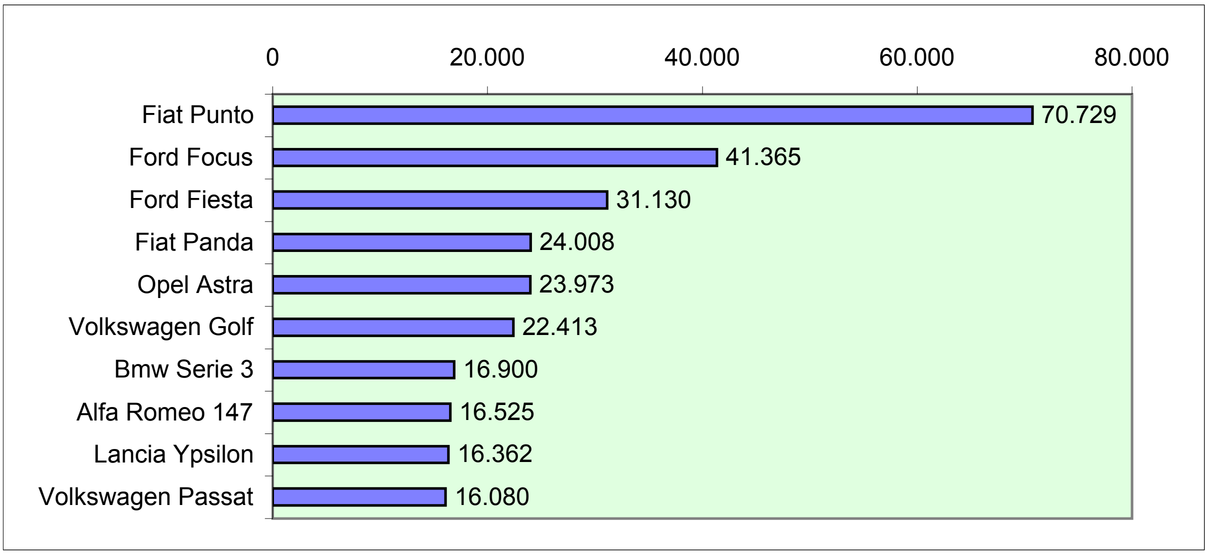
La tabella 1 contiene i dieci modelli di auto diesel più vendute in Italia nei primi sei mesi del 2006
Rappresentare i dati con un grafico a barre orizzontali, ordinandoli in modo decrescente
Indicare sull'asse verticale gli oggetti rappresentati

Tabella 1

	Modello	Numero auto vendute
1	Fiat Punto	70.729
2	Ford Focus	41.365
3	Ford Fiesta	31.130
4	Fiat Panda	24.008
5	Opel Astra	23.973
6	Volkswagen Golf	22.413
7	Bmw Serie 3	16.900
8	Alfa Romeo 147	16.525
9	Lancia Ypsilon	16.362
10	Volkswagen Passat	16.080

SUGGERIMENTI

1. Per realizzare il grafico a barre orizzontali selezionare i dati delle celle E12:E21 e cliccare sul pulsante Creazione guidata grafico; selezionare Tipo di grafico>Barre
2. Nella scheda Serie (Passaggio 2) nella casella Etichette asse categorie (X) selezionare le celle dei modelli delle auto (celle D12:D21)
   Per far comparire i numeri delle auto vendute accanto a ciascuna barra, al Passaggio 3 selezionare la scheda Etichette dati e scegliere Valore. Terminare il grafico.
3. Ridimensionare il grafico in verticale in modo che compaiano tutte le etichette sull'asse verticale
4. Per far comparire le barre in ordine inverso posizionare il puntatore del mouse su Asse delle categorie (asse verticale), cliccare con il tasto destro, scegliere Formato asse e nella Finestra Formato asse selezionare la scheda Scala, scegliere Categorie in ordine inverso
5. Togliere eventualmente la griglia verticale cliccando sulla griglia stessa nel grafico (si attivano i quadratini neri) e premendo il tasto Canc
6. Per togliere la Legenda cliccare sulla legenda e premere Canc



Esercizio 12.2

Spese per l'attività di ricerca svolta nelle Università italiane, suddivisa per regione, Anno 2005
La tabella 2 riporta la spesa per Ricerca e Sviluppo (R&S) suddivisa per regione; calcolare la spesa percentuale regione per regione e disegnare un diagramma a barre orizzontali (o anche verticali) indicando nel grafico nomi delle regioni e valori percentuali
Ordinare le barre in modo decrescente nel diagramma a barre orizzontali, (in modo crescente in quello a barre verticali)

Tabella 2

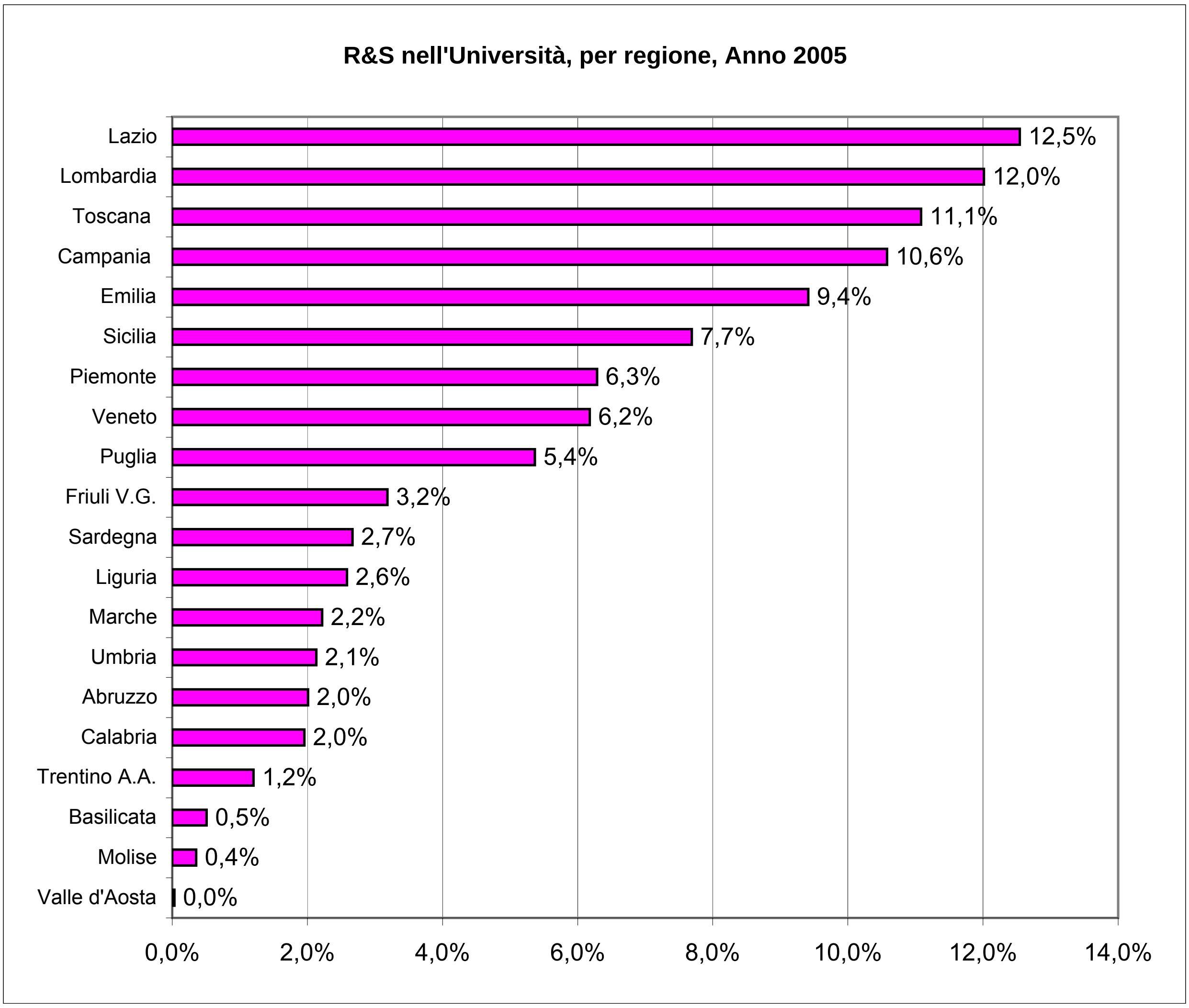
Regione	Migliaia di euro	Percentuale
Valle d'Aosta	1.395	0,0%
Molise	16.626	0,4%
Basilicata	23.966	0,5%
Trentino A.A.	56.562	1,2%
Calabria	92.090	2,0%
Abruzzo	94.653	2,0%
Umbria	100.485	2,1%
Marche	104.488	2,2%
Liguria	121.930	2,6%
Sardegna	125.748	2,7%
Friuli V.G.	149.970	3,2%
Puglia	252.892	5,4%
Veneto	291.112	6,2%
Piemonte	296.156	6,3%
Sicilia	362.320	7,7%
Emilia	443.494	9,4%
Campania	498.387	10,6%
Toscana	522.197	11,1%
Lombardia	566.080	12,0%
Lazio	591.119	12,5%
ITALIA (Totale)	4.711.670	100%

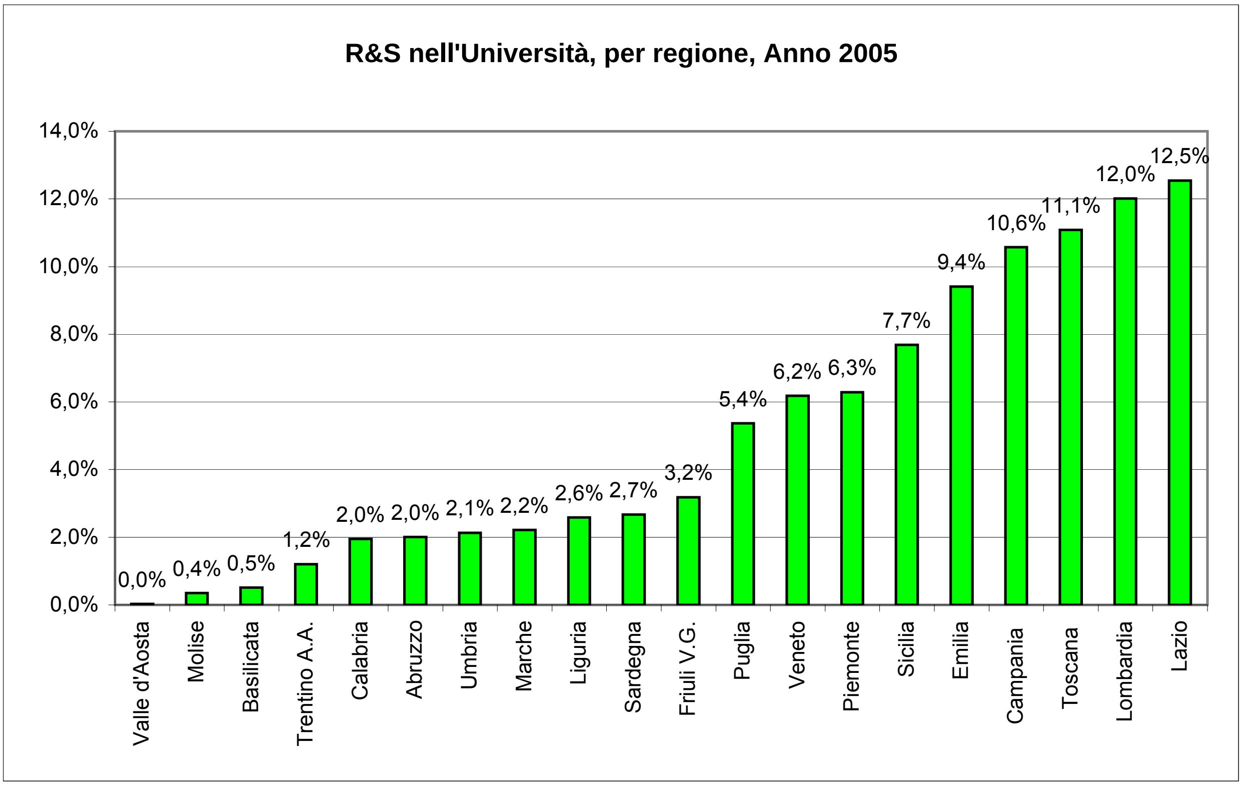
(Fonte Istat)

SUGGERIMENTI

Per ottenere le barre in modo decrescente (o crescente verso destra), prima di fare il grafico
Ordinare i dati in modo decrescente, selezionando la colonna delle percentuali e ricordando di espandere la selezione, in modo che ogni regione sia associata al suo dato

I grafici evidenziano che quattro regioni (Lazio, Lombardia, Toscana e Campania) esauriscono circa la metà (46%) della spesa.
Il valore basso di alcune regioni (ad esempio la regione Piemonte) mostra che in tale regione l'attività di ricerca viene svolta soprattutto fuori dell'Università.


R&S nell'Università, per regione, Anno 2005
Lazio 12,5%
Lombardia 12,0%
Toscana 11,1%
Campania 10,6%
Emilia 9,4%
Sicilia 7,7%
Piemonte 6,3%
Veneto 6,2%
Puglia 5,4%
Friuli V.G. 3,2%
Sardegna 2,7%
Liguria 2,6%
Marche 2,2%
Umbria 2,1%
Abruzzo 2,0%
Calabria 2,0%
Trentino A.A. 1,2%
Basilicata 0,5%
Molise 0,4%
Valle d'Aosta 0,0%
0,0% 2,0% 4,0% 6,0% 8,0% 10,0% 12,0% 14,0%


R&S nell'Università, per regione, Anno 2005
14,0% 12,0% 10,0% 8,0% 6,0% 4,0% 2,0% 0,0%
Valle d'Aosta 0,0%
Molise 0,4%
Basilicata 0,5%
Trentino A.A. 1,2%
Calabria 2,0%
Abruzzo 2,0%
Umbria 2,1%
Marche 2,2%
Liguria 2,6%
Sardegna 2,7%
Friuli V.G. 3,2%
Puglia 5,4%
Veneto 6,2%
Piemonte 6,3%
Sicilia 7,7%
Emilia 9,4%
Campania 10,6%
Toscana 11,1%
Lombardia 12,0%
Lazio 12,5%

Esercizio 12.2

Per valutare come si colloca l'Italia a livello internazionale per quanto riguarda la spesa per la ricerca osserviamo due indicatori di fonte OSCE
Nella tabella 3 sono riportati i dati relativi alla spesa per Ricerca e Sviluppo dell'Università in rapporto al PIL in alcuni paesi; nella tabella 4 per gli stessi paesi sono riportati i dati relativi alla spesa per R&S dell'Università rispetto alla spesa totale della nazione in R&S.
(Anno 2005, Fonte OCSE. Tutti i dati sono espressi in percentuale)
Realizzare due grafici a barre orizzontali per illustrare i dati (ordinare le barre in modo decrescente)

SUGGERIMENTI

Per colorare alcune (non tutte le barre) di un colore diverso cliccare una volta sulle barre per selezionarle tutte, poi cliccare una volta sulla barra di cui si vuole cambiare il colore, premere il tasto destro e scegliere Formato dato nella scheda Formato scegliere il colore nuovo dell'Area
Ripetere per ogni barra di cui si vuole cambiare il colore.
Per colorare tutte le barre di un colore diverso cliccare una volta sulle barre per selezionarle tutte, premere il tasto destro e scegliere Formato serie dati; nella scheda Opzioni della scheda Formato selezionare Varia colore per dato

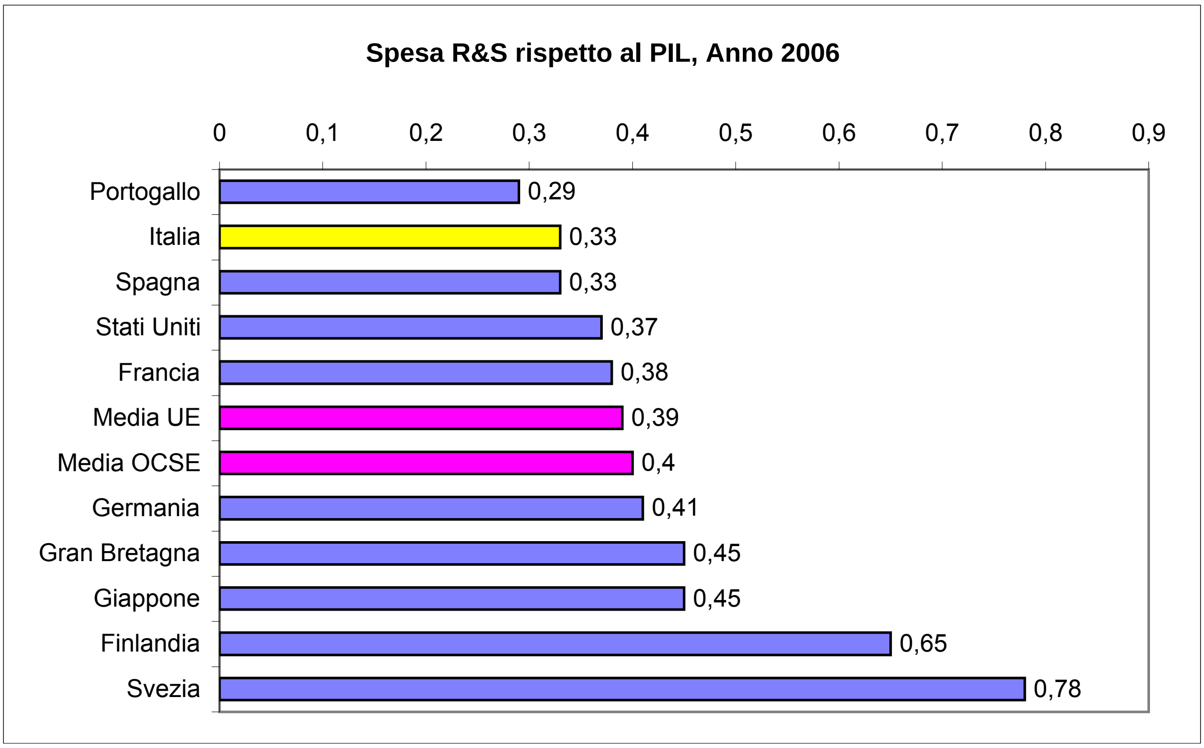
Tabella 3	
Portogallo	0,29
Italia	0,33
Spagna	0,33
Stati Uniti	0,37
Francia	0,38
Media UE	0,39
Media OCSE	0,4
Germania	0,41
Gran Bretagna	0,45
Giappone	0,45
Finlandia	0,65
Svezia	0,78

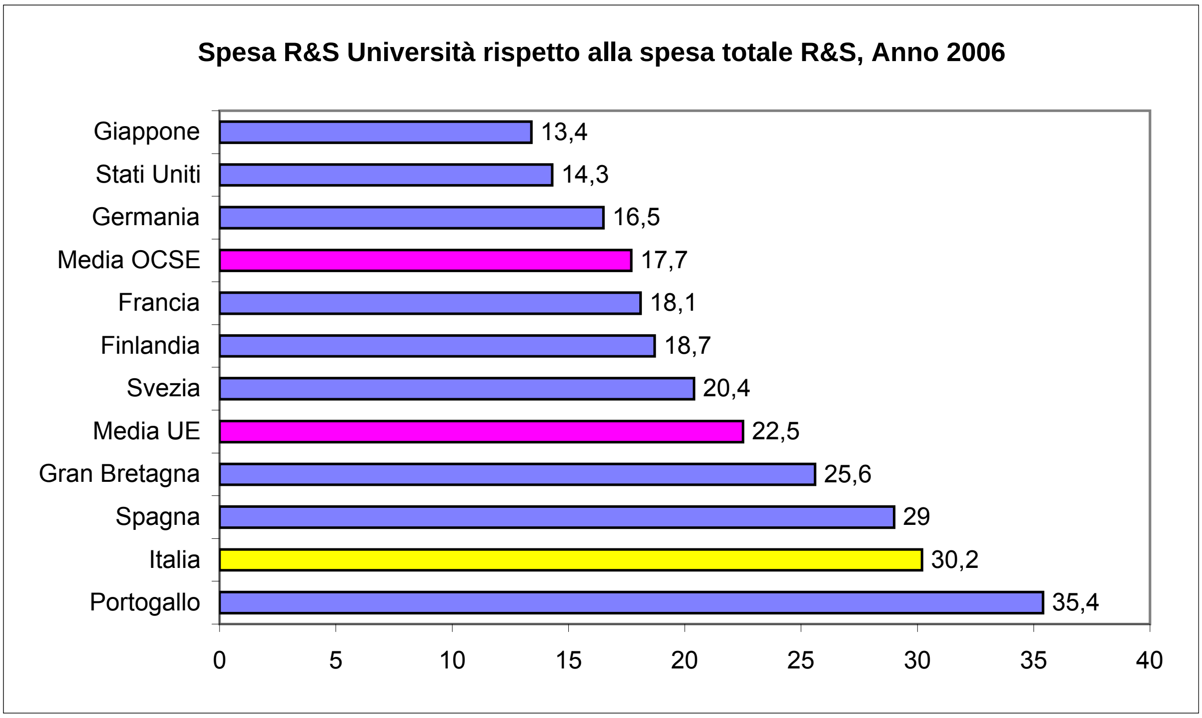
(Fonte OCSE)

Tabella 4	
Portogallo	35,4
Italia	30,2
Spagna	29
Gran Bretagna	25,6
Media UE	22,5
Svezia	20,4
Finlandia	18,7
Francia	18,1
Media OCSE	17,7
Germania	16,5
Stati Uniti	14,3
Giappone	13,4

(Fonte OCSE)

Il primo grafico evidenzia che l'Italia è in posizione di svantaggio sia rispetto alla media dei paesi UE che OCSE (0,39% e 0,4% rispettivamente): siamo al penultimo posto.
Il secondo grafico mostra che l'Italia è ancora al penultimo posto per la ricerca svolta da altri soggetti diversi dalle Università.


Spesa R&S rispetto al PIL, Anno 2006
0
0,1
0,2
0,3
0,4
0,5
0,6
0,7
0,8
0,9
Portogallo 0,29
Italia 0,33
Spagna 0,33
Stati Uniti 0,37
Francia 0,38
Media UE 0,39
Media OCSE 0,4
Germania 0,41
Gran Bretagna 0,45
Giappone 0,45
Finlandia 0,65
Svezia 0,78


Spesa R&S Università rispetto alla spesa totale R&S, Anno 2006
Giappone 13,4
Stati Uniti 14,3
Germania 16,5
Media OCSE 17,7
Francia 18,1
Finlandia 18,7
Svezia 20,4
Media UE 22,5
Gran Bretagna 25,6
Spagna 29
Italia 30,2
Portogallo 35,4
0
5
10
15
20
25
30
35
40

Torna su

Soluzione Esercizio 13
Diagramma circolare in due e tre dimensioni

Indice

Ritorna Esercizio 13

Esercizio 13.1

1 Copiare (Copia/incolla) la tabella 1 del foglio precedente nelle celle C14:E24
2 Costruire una nuova tabella (celle G14:H17) creando tre gruppi: Italia (Fiat, Lancia Alfa Romeo) Germania (Bmw, Opel, VW) e USA (Ford)
3 Rappresentare i dati di questa tabella con diagrammi circolari in due e tre dimensioni
Assegnare il titolo "Numero di auto vendute in Italia per paese di provenienza"
Aggiungere nel grafico una scritta in basso con la frase "Dati primo semestre 2006"

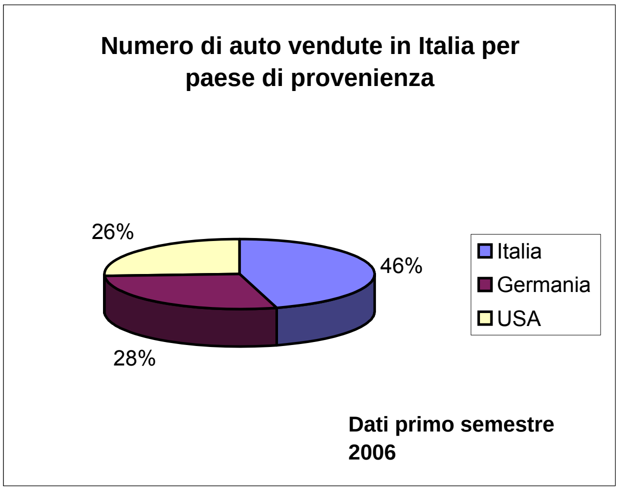
Tabella 1

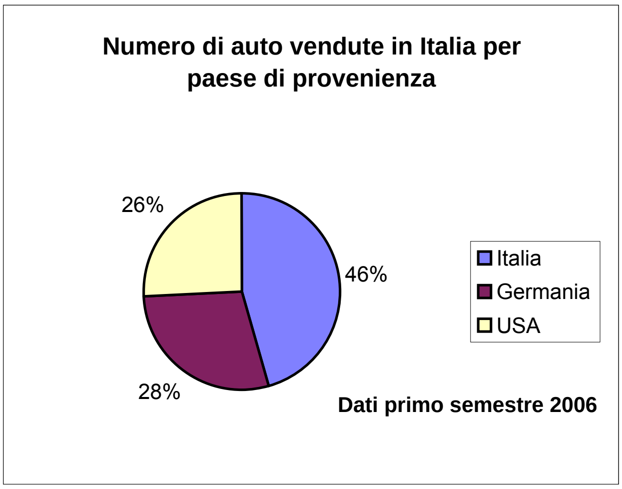
	Modello	Numero auto vendute
1	Fiat Punto	70.729
2	Ford Focus	41.365
3	Ford Fiesta	31.130
4	Fiat Panda	24.008
5	Opel Astra	23.973
6	Volkswagen Golf	22.413
7	Bmw Serie 3	16.900
8	Alfa Romeo 147	16.525
9	Lancia Ypsilon	16.362
10	Volkswagen Passat	16.080

Nazione	Numero auto vendute
Italia	127.624
Germania	79.366
USA	72.495

SUGGERIMENTI

2 Per calcolare le somme per nazione usare il pulsante Inserisci funzione, scegliere SOMMA
Nel riquadro degli argomenti (Num1) scrivere gli indirizzi delle celle che contengono i numeri da sommare separati da punto e virgola
Gli indirizzi delle celle possono essere inseriti selezionando direttamente le celle che contengono i numeri da sommare, premendo il tasto Ctrl (le celle non sono contigue)
3 Per realizzare il diagramma circolare selezionare le celle G14:H16 e cliccare su Creazione guidata grafico. Selezionare Tipo di grafico>Torta oppure Torta3D
Nel Passaggio 3, scheda Titoli, scrivere il titolo "Numero di auto vendute in Italia per paese di provenienza"; nella scheda Etichette dati scegliere percentuali
4 Per inserire la scritta in basso con la frase "Dati Primo semestre 2006" visualizzare la barra degli strumenti Disegno (Menu Visualizza>Barre degli strumenti>Disegno), cliccare sul pulsante Casella di testo: si apre un riquadro nell'area del grafico in cui scrivere la frase (usare il grassetto)




Numero di auto vendute in Italia per paese di provenienza
26%
46%
28%
Italia
Germania
USA
Dati primo semestre 2006

Torna su

Soluzione Esercizio 14
Istogrammi a barre multiple e in pila

Indice

Ritorna Esercizio 14

Esercizio 14.1
Nella tabella 1 si riportano i dati riguardanti l'istruzione universitaria in Italia (Fonte Istat, anno 2005/2006)

Tabella 1

	corsi di laurea	studenti in corso	studenti fuori corso	laureati
1	Facoltà scientifiche	101522	44255	13982
2	Facoltà di medicina	65211	7778	20361
3	Facoltà tecniche	128352	59339	23510
4	Facoltà economiche	119923	45616	19783
5	Facoltà giuridiche	218488	82153	30967
6	Facoltà letterarie	200692	80227	27768
7	Scienze motorie	13956	5339	1936
	Totali	848144	324707	138307

1 Calcolare i totali nella tabella 1
2 Realizzare un istogramma a barre multiple per i dati della tabella 1
3 Costruire la tabella 2 che riporta le percentuali di laureati in ciascuna facoltà
4 Costruire la tabella 3 che riporta le percentuali di laureati rispetto al numero di iscritti per ciascuna facoltà
5 Realizzare un diagramma circolare che illustri le percentuali della tabella 2
6 Realizzare un istogramma in pila per i dati della tabella 3
7 Perfezionare i grafici con titoli, legende, ecc.

Tabella 2

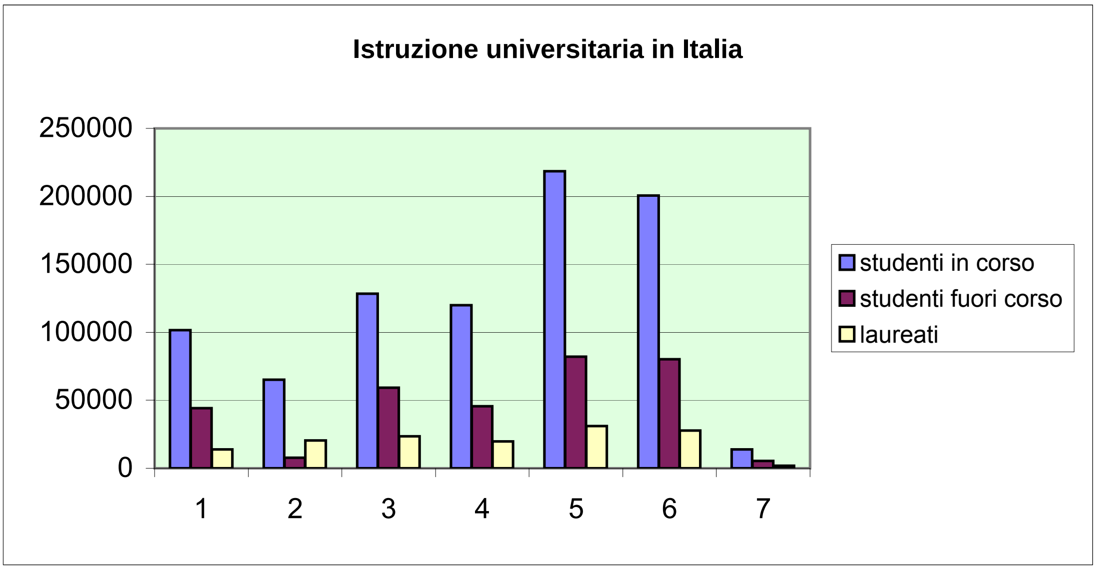
	corsi di laurea	laureati	percentuali laureati per Facoltà
1	Facoltà scientifiche	13982	10,1%
2	Facoltà di medicina	20361	14,7%
3	Facoltà tecniche	23510	17,0%
4	Facoltà economiche	19783	14,3%
5	Facoltà giuridiche	30967	22,4%
6	Facoltà letterarie	27768	20,1%
7	Scienze motorie	1936	1,4%
	Totali	138307	100,0%

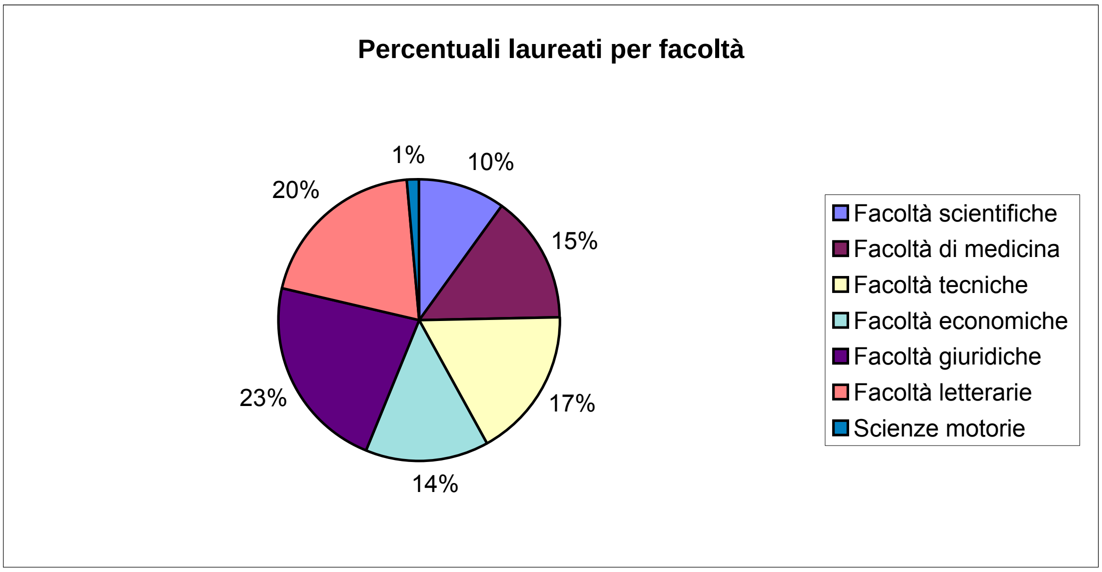
Tabella 3

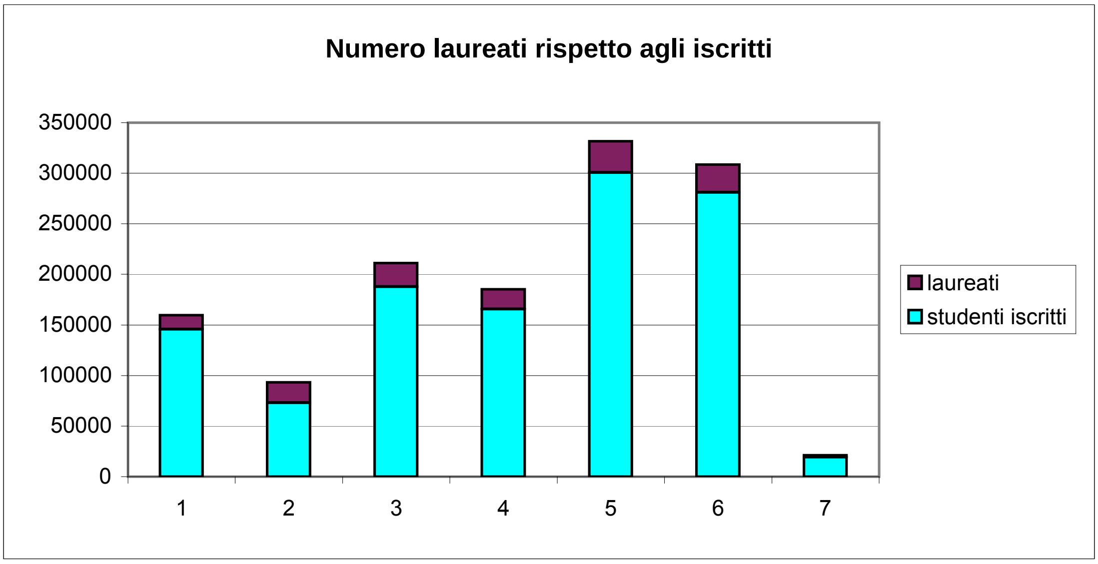
	corsi di laurea	studenti iscritti	laureati	percentuale laureati rispetto agli iscritti
1	Facoltà scientifiche	145777	13982	9,6%
2	Facoltà di medicina	72989	20361	27,9%
3	Facoltà tecniche	187691	23510	12,5%
4	Facoltà economiche	165539	19783	12,0%
5	Facoltà giuridiche	300641	30967	10,3%
6	Facoltà letterarie	280919	27768	9,9%
7	Scienze motorie	19295	1936	10,0%
	Totali	1172851	138307	11,8%

SUGGERIMENTI

2 Selezionare le celle D11:F18; realizzare il grafico cliccando sul pulsante Creazione guidata grafico; selezionare Tipo di grafico>Istogramma
Completare il grafico con il titolo.
Se non si selezionano le celle dell'intestazione (D11:F11) la legenda non è corretta
3 Attenzione ai riferimenti relativi e assoluti nella costruzione della tabella 2
4 Il risultato nella cella F49 rappresenta la percentuale di laureati (in tutte le discipline) rispetto al totale degli iscritti di tutte le facoltà.
5 Per ottenere la legenda corretta realizzare il grafico come segue:
Selezionare le celle C31:D37 della tabella 2 e procedere alla realizzazione del grafico con Creazione guidata grafico
Attenzione a non selezionare le celle dei totali
6 Selezionare i dati delle celle D42:E48 (tabella 3) e creare il grafico scegliendo Istogramma in pila






Numero laureati rispetto agli iscritti
350000
300000
250000
200000
150000
100000
50000
0
1
2
3
4
5
6
7
laureati
studenti iscritti

Torna su

Soluzione Esercizio 15

Istogrammi, diagrammi a barre, diagrammi circolari, grafici a linee

Indice

Ritorna Esercizio 15

Esercizio 15.1

Il reddito non è uguale per tutti!

Nella tabella 1 si riportano i dati relativi alla distribuzione del reddito delle famiglie italiane (Fonte Bankitalia 2008)

Nella prima colonna della tabella compare la percentuale di famiglie avente il reddito riportato nelle colonne accanto

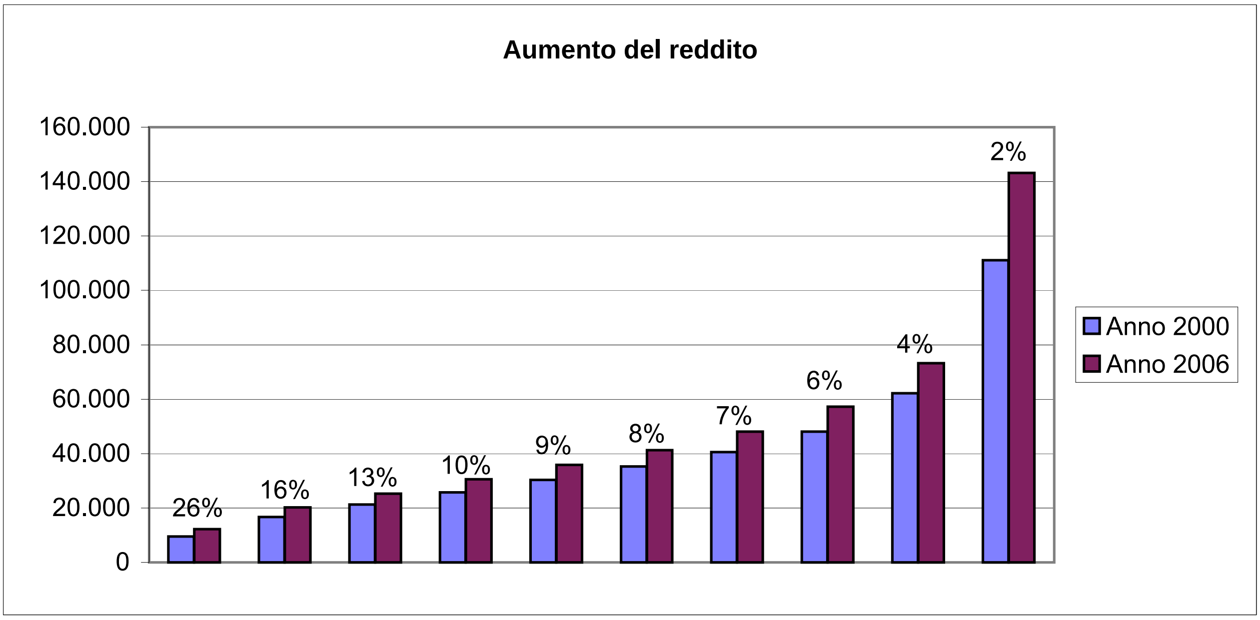
La seconda colonna è relativa all'anno 2000, la terza all'anno 2006

Tabella 1

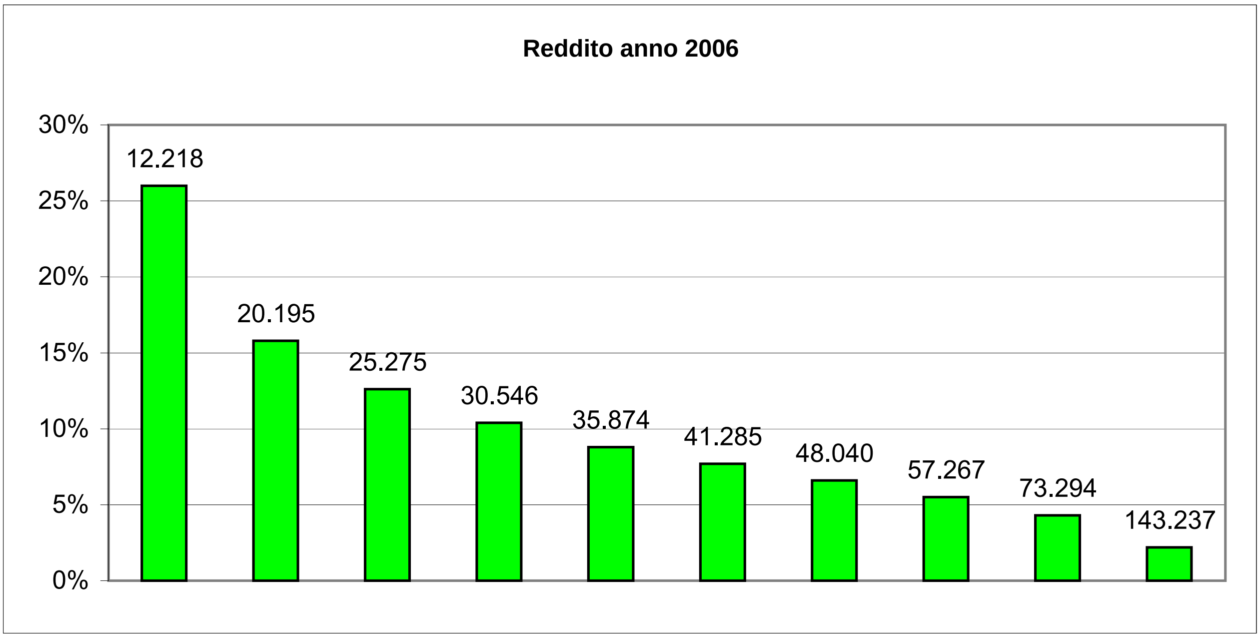
Percentuale di famiglie	Reddito anno 2000	Reddito anno 2006
26%	9.478	12.218
16%	16.735	20.195
13%	21.224	25.275
10%	25.688	30.546
9%	30.278	35.874
8%	35.222	41.285
7%	40.562	48.040
6%	48.100	57.267
4%	62.162	73.294
2%	111.072	143.237

1 Realizzare un istogramma a barre multiple che evidenzi il fatto che l'aumento di reddito non è uguale per tutti (Riportare in ordinata i valori del reddito)

2 Realizzare un istogramma relativo all'anno 2006, che evidenzi il legame decrescente fra livello di reddito e percentuale di famiglie che lo possiedono (Riportare in ordinata le percentuali di famiglie)



Il grafico mostra che l'aumento di reddito è molto più elevato nelle classi più abbienti.



Il grafico mostra che circa un quarto delle famiglie italiane appartiene alla prima classe di reddito e solo il 2% delle famiglie ha un reddito superiore a 140.000 euro!

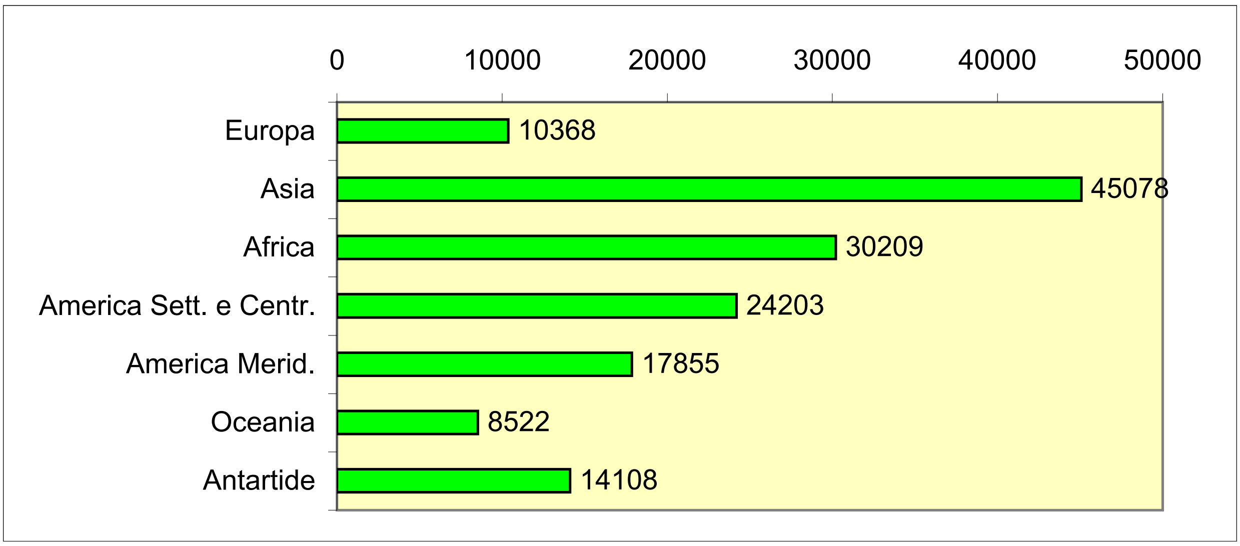
Esercizio 15.2

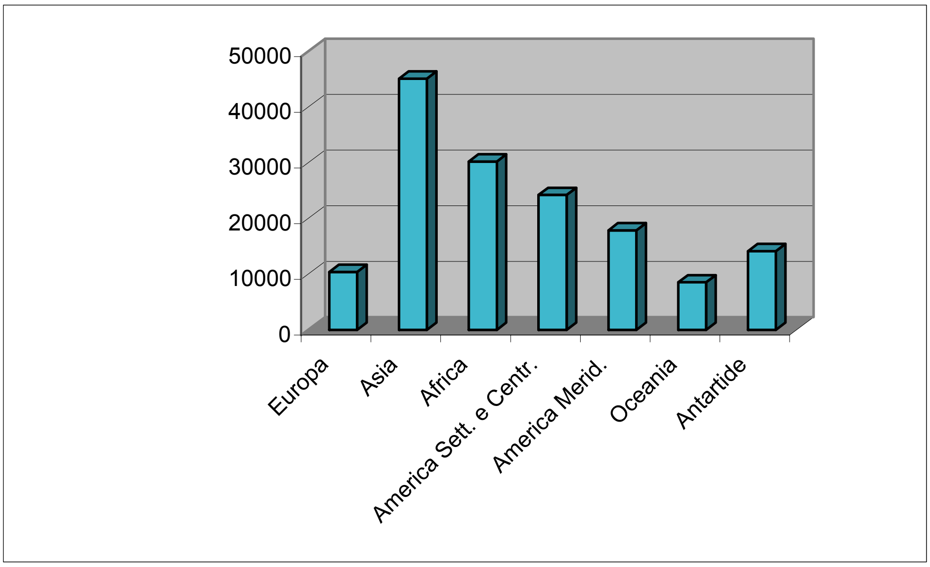
Nella tabella 2 si riportano le aree dei continenti del mondo, in migliaia di km quadrati.

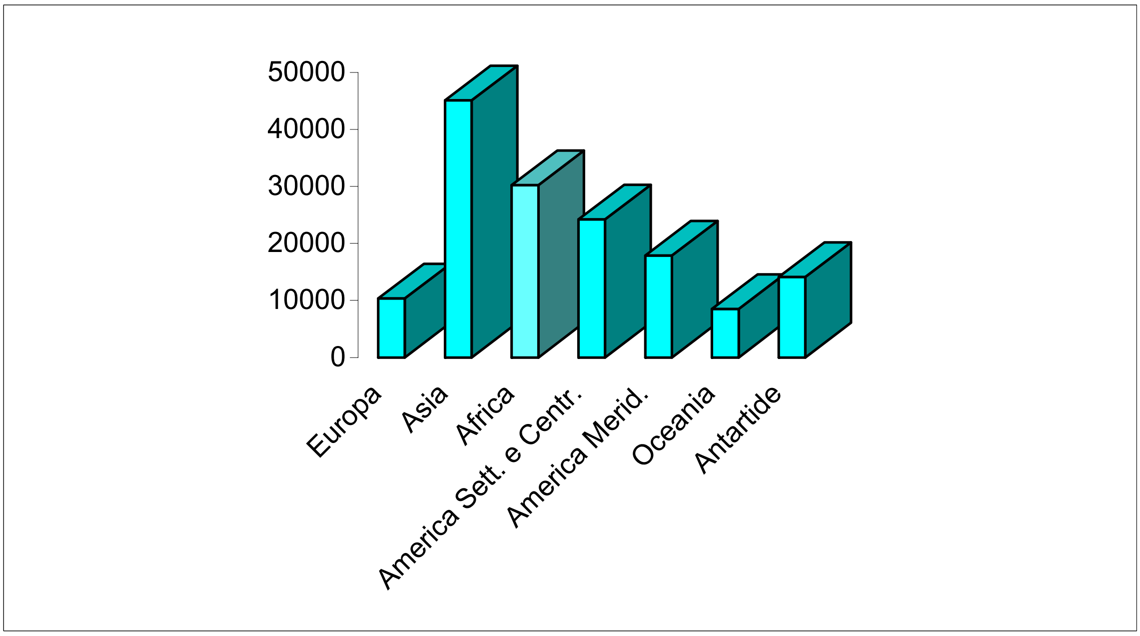
Tabella 2

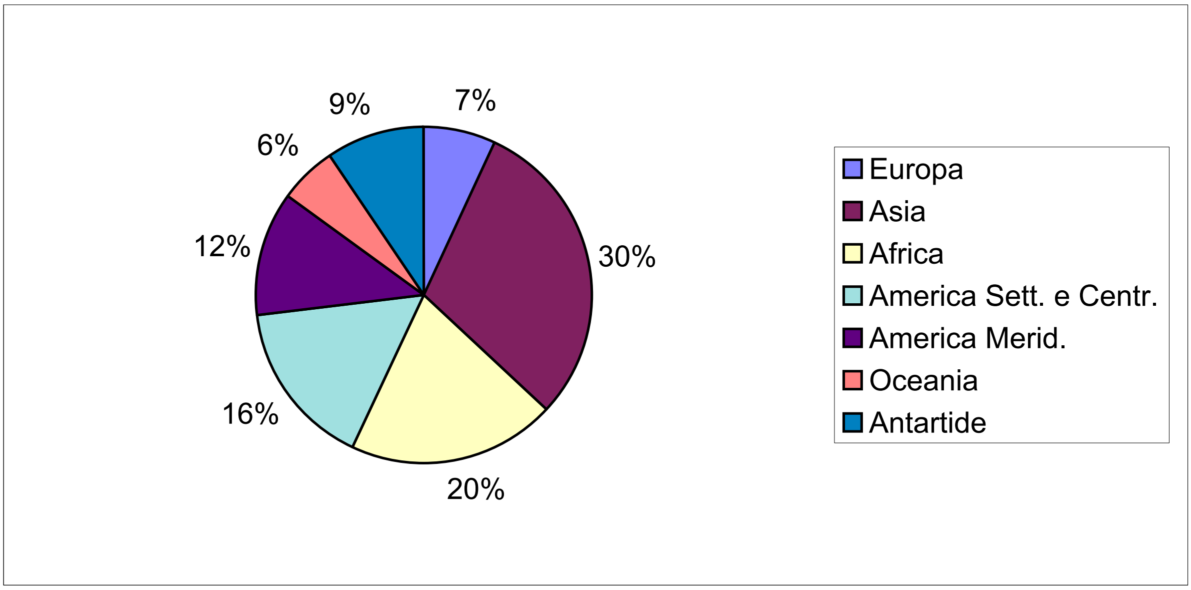
Continente	Area
Europa	10368
Asia	45078
Africa	30209
America Sett. e Centr.	24203
America Merid.	17855
Oceania	8522
Antartide	14108

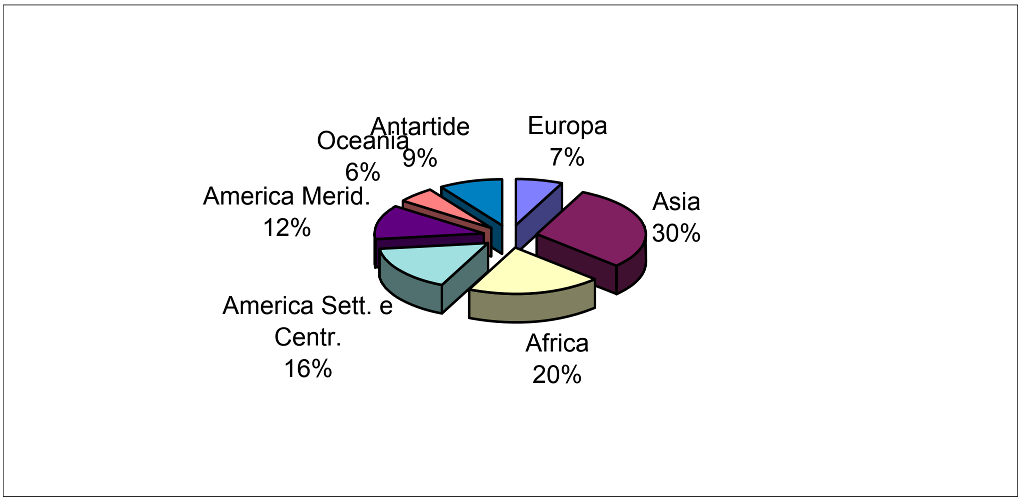
Rappresentare con istogrammi, diagrammi a barre orizzontali e con diagrammi circolari




50000
40000
30000
20000
10000
0
Europa
Asia
Africa
America Sett. e Centr.
America Merid.
Oceania
Antartide


50000
40000
30000
20000
10000
0
Europa
Asia
Africa
America Sett. e Centr.
America Merid.
Oceania
Antartide


7%
30%
20%
16%
12%
6%
9%
Europa
Asia
Africa
America Sett. e Centr.
America Merid.
Oceania
Antartide



Esercizio 3

La tabella 3 riporta i dati delle vendite di uno dei prodotti di un'azienda negli anni 2002-2007 e il fatturato totale dell'azienda stessa per ogni anno

Tabella 3

Anno	Vendite	Fatturato annuo	%
2002	€ 35.000	€ 255.000	14%
2003	€ 25.100	€ 305.000	8%
2004	€ 30.500	€ 285.000	11%
2005	€ 28.500	€ 405.000	7%
2006	€ 35.500	€ 505.000	7%
2007	€ 45.500	€ 455.000	10%

Per ogni anno calcolare le percentuali di vendita rispetto al fatturato totale annuo

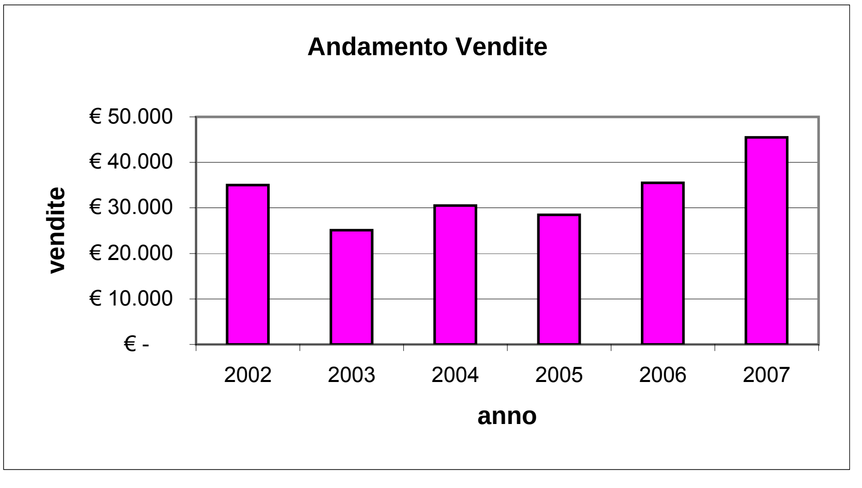
Realizzare i seguenti grafici

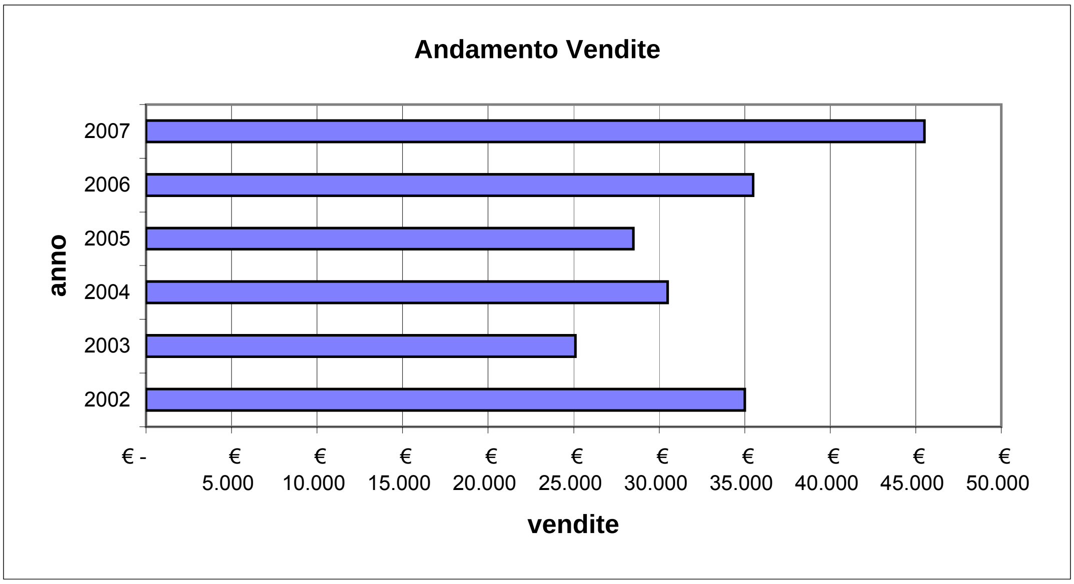
1 Un grafico a barre verticali che rappresenti gli anni sull'asse orizzontale e le vendite sull'asse verticale

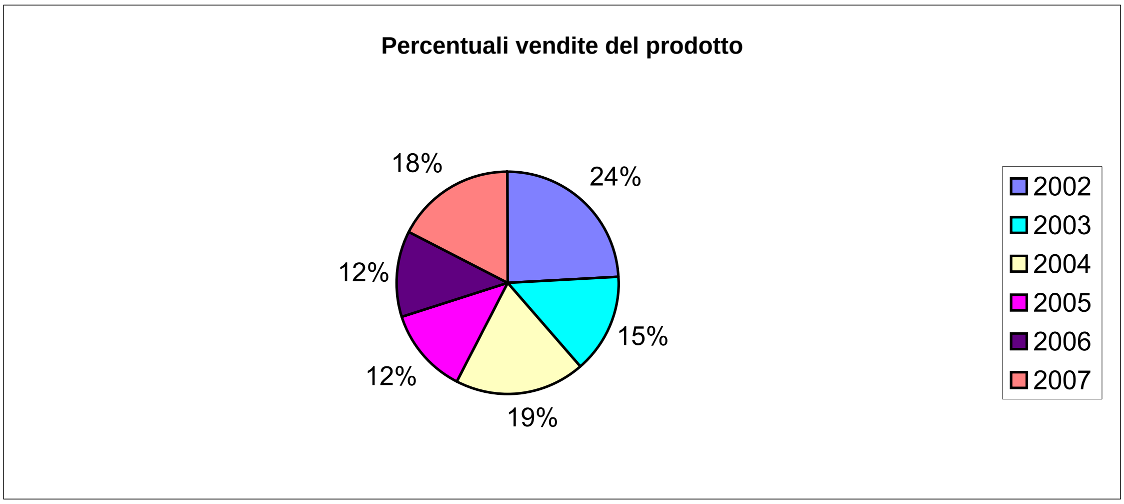
2 Un grafico a barre orizzontali che rappresenti gli anni sull'asse verticale e le vendite sull'asse orizzontale

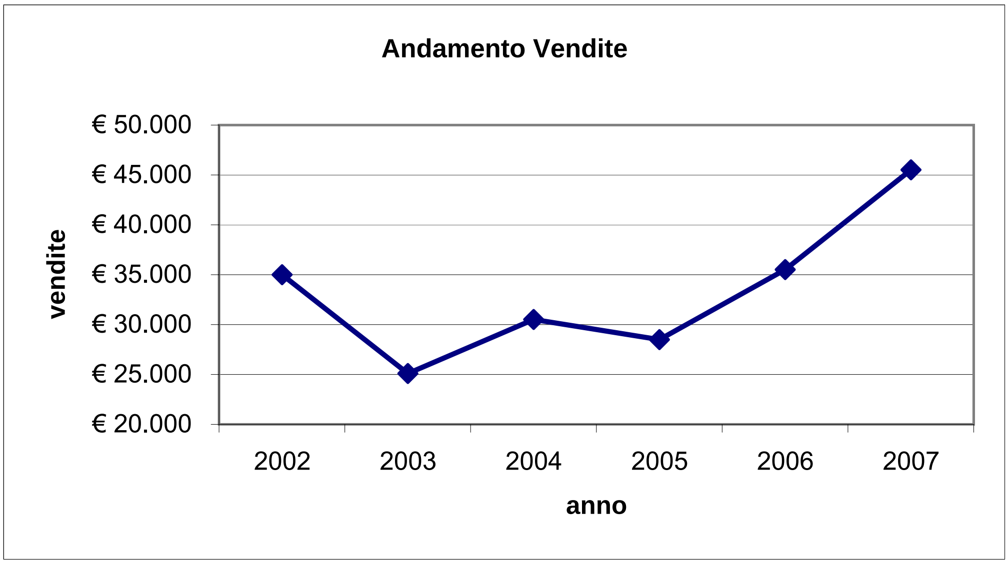
3 Un grafico a torta che rappresenti le percentuali di vendita nei sei anni rispetto al fatturato totale di ciascun anno; riportare come etichette di ciascuna fetta i valori delle percentuali dell'anno, nella legenda far comparire gli anni

4 Un grafico a linee che rappresenti gli anni sull'asse orizzontale e le vendite sull'asse verticale




Andamento Vendite
2007
2006
2005
2004
2003
2002
anno
€ -
€ 5.000
€ 10.000
€ 15.000
€ 20.000
€ 25.000
€ 30.000
€ 35.000
€ 40.000
€ 45.000
€ 50.000
vendite


Percentuali vendite del prodotto
18%
24%
12%
15%
12%
19%
2002
2003
2004
2005
2006
2007


Andamento Vendite
€ 50.000
€ 45.000
€ 40.000
€ 35.000
€ 30.000
€ 25.000
€ 20.000
vendite
2002
2003
2004
2005
2006
2007
anno

Torna su

Soluzione Esercizio 16
Grafici a dispersione

Indice

Ritorna Esercizio 16

Esercizio 16.1

La tabella seguente contiene il risultato di un test sul consumo di benzina di un'auto, misurato a differenti velocità, in litri per 100 Km.

1 Calcolare il consumo espresso in Km percorsi per litro

2 Rappresentare i dati dei consumi in litri per Km e in Km al litro in due diversi grafici a dispersione

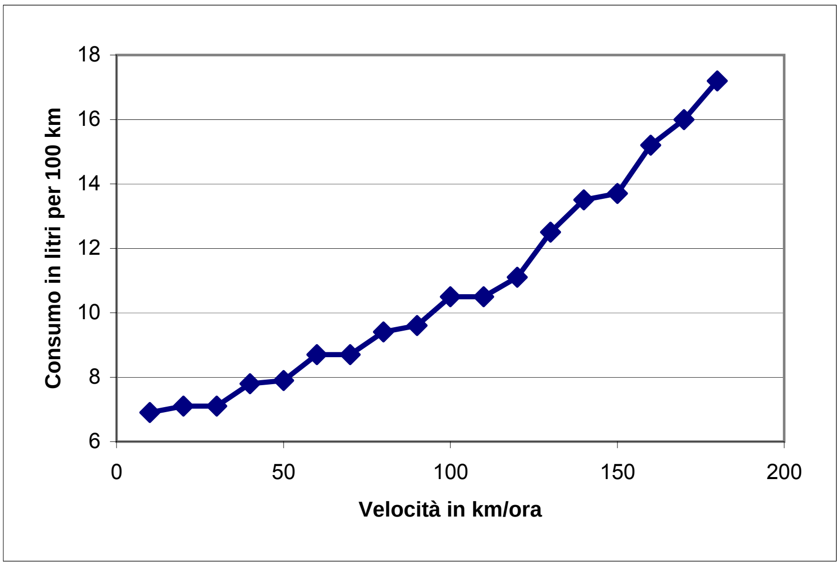
Velocità in Km/ora	Consumo in litri per 100Km	Consumo in Km per litro
10	6,9	14,5
20	7,1	14,1
30	7,1	14,1
40	7,8	12,8
50	7,9	12,7
60	8,7	11,5
70	8,7	11,5
80	9,4	10,6
90	9,6	10,4
100	10,5	9,5
110	10,5	9,5
120	11,1	9,0
130	12,5	8,0
140	13,5	7,4
150	13,7	7,3
160	15,2	6,6
170	16	6,3
180	17,2	5,8

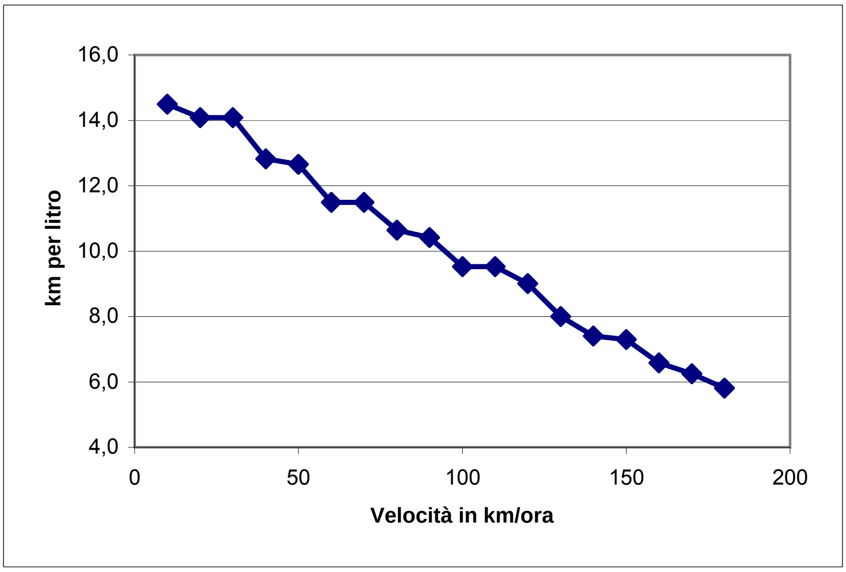
SUGGERIMENTI

1 Calcolare 100/(consumo in litri per 100 Km)
Usare il pulsante Diminuisci decimali per la rappresentazione con un decimale

2 Selezionare i dati delle celle B17:C34
Scegliere Tipo di grafico>Dispersione (XY), il grafico a spezzata (in basso a sinistra)
Passaggio 3 del grafico: aggiungere i nomi agli assi nella scheda Titoli

3 Selezionare i dati delle celle B17:B34 e D17:D34 (usare il pulsante Ctrl)
Scegliere Tipo di grafico>Dispersione (XY)
Aggiungere i nomi degli assi
Modificare eventualmente lo spessore delle linee


Consumo in litri per 100 km
18
16
14
12
10
8
6
0
50
100
150
200
Velocità in km/ora


km per litro
16,0
14,0
12,0
10,0
8,0
6,0
4,0
0
50
100
150
200
Velocità in km/ora

Torna su

Soluzione Esercizio 17
Grafico della funzione y=sen(kx)

Indice

Ritorna Esercizio 17

Esercizio 17.1

Costruire due colonne, una dove sono riportati i valori delle ascisse a intervalli crescenti equidistanziati di un incremento h (scrivere l'incremento nella cella C16) a partire da x=0 fino a x=π, e una con i valori corrispondenti della funzione sen(kx) dove k è un parametro (a scelta) da inserire nella cella C15.
Disegnare il grafico della funzione così tabulata.

k	2
Incremento	0,1

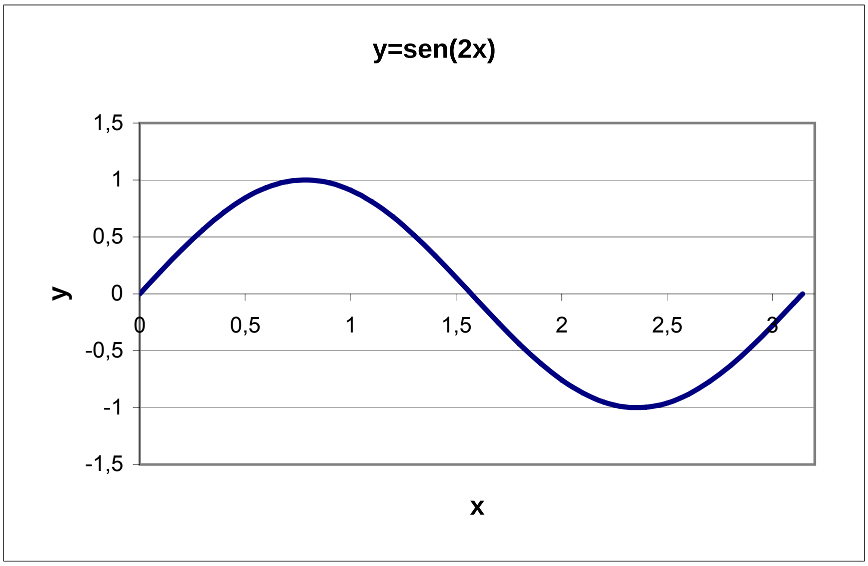
π	3,14159

x	sen(kx)
0	0
0,1	0,1987
0,2	0,3894
0,3	0,5646
0,4	0,7174
0,5	0,8415
0,6	0,9320
0,7	0,9854
0,8	0,9996
0,9	0,9738
1	0,9093
1,1	0,8085
1,2	0,6755
1,3	0,5155
1,4	0,3350
1,5	0,1411
1,6	-0,0584
1,7	-0,2555
1,8	-0,4425
1,9	-0,6119
2	-0,7568
2,1	-0,8716
2,2	-0,9516
2,3	-0,9937
2,4	-0,9962
2,5	-0,9589
2,6	-0,8835
2,7	-0,7728
2,8	-0,6313
2,9	-0,4646
3	-0,2794
3,1	-0,0831
3,14	-5,307E-06

SUGGERIMENTI

1 Scegliere il valore di k (ad esempio k=2) e scriverlo nella cella C15
2 Scegliere un passo h; la funzione viene disegnata per punti perciò il passo deve essere sufficientemente piccolo (altrimenti la curva appare come una spezzata: provare a prendere come passo 0,5 e verificare che la curva non è liscia). Scegliere h=0,1 e scriverlo nella cella C16

3 Nella cella E16 scrivere il primo valore di x (x=0); nella cella E17 scrivere il secondo valore di x con una formula (usare l'incremento; attenzione ai riferimenti assoluti e relativi)
Selezionare la cella E17 e trascinare con il mouse fino al valore x=3,1
Nella cella E48 aggiungere il valore approssimato di pi greco (copiare la cella C18).
4 Nella cella F16 scrivere la formula per il calcolo della funzione sen(kx) nei punti x della colonna E (attenzione ai riferimenti assoluti e relativi)
5 Selezionare la cella F16 e trascinare con il mouse fino al fondo della tabella
6 Selezionare le celle delle colonne E e F (tabella ascisse e ordinate)
Realizzare il grafico con Creazione guidata grafico, Tipo di grafico>Dispersione (XY)
Aggiungere il titolo, modificare lo spessore della linea cliccando due volte sulla linea,
Formato serie dati> scheda Motivo>Linea>Spessore



Soluzione Esercizio 18
Rappresentazione di più funzioni sullo stesso grafico

Indice

Ritorna Esercizio 18

Esercizio 18.1

Sia data la retta di equazione y = mx+n e la parabola di equazione y = ax^2+bx+c

1 Inserire i valori dei coefficienti nelle celle G20:G21 e G24:G26
m= –3 n=2 a= –2 b=2 c=12
2 Costruire una tabella dove per i valori di x compresi nell'intervallo (–5,5) e con passo h=0,5
si calcolano i corrispondenti valori di y per la retta e la parabola
3 Disegnare il grafico delle due funzioni sullo stesso grafico
4 Cambiare i valori dei coefficienti e osservare i cambiamenti

	retta	parabola
X	Y	Y
-5	17	-48
-4,5	15,5	-37,5
-4	14	-28
-3,5	12,5	-19,5
-3	11	-12
-2,5	9,5	-5,5
-2	8	0
-1,5	6,5	4,5
-1	5	8
-0,5	3,5	10,5
0	2	12
0,5	0,5	12,5
1	-1	12
1,5	-2,5	10,5
2	-4	8
2,5	-5,5	4,5
3	-7	0
3,5	-8,5	-5,5
4	-10	-12
4,5	-11,5	-19,5
5	-13	-28

Passo h:	0,5
Coefficienti retta	
m	-3
n	2

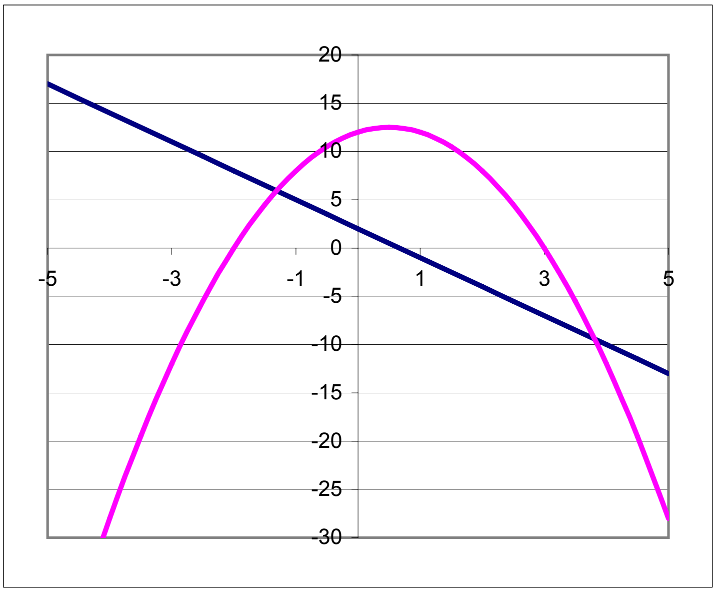
Coefficienti parabola	
a	-2
b	2
c	12

SUGGERIMENTI

1 Scrivere il valore del passo h e i coefficienti della retta e della parabola nelle celle indicate
2 Scrivere il primo valore di x nella cella B19; il secondo valore nella cella B20 con una formula
usando il passo (riferimenti assoluti e relativi!). Trascinare con il mouse fino al valore x=5
Calcolare l'ordinata sulla retta (cella C19) e sulla parabola (cella D19) con una formula
(riferimenti assoluti e relativi!.) Trascinare con il mouse fino al fondo della tabella.
3 Per realizzare i grafici delle due funzioni selezionare le tre colonne B,C,D (B19:D39) appena
costruite e selezionare Tipo di grafico>Dispersione (XY)
Togliere la legenda (in questo caso è chiaro chi è la retta e chi la parabola!)
Per fare in modo che il grafico non presenti regioni vuote a sinistra/destra e in alto/in basso,
dopo aver realizzato il grafico posizionare il puntatore del mouse prima su Asse dei valori (X)
poi su Asse dei valori (Y), cliccare due volte e nella scheda Scala scegliere Valore minimo e

Valore massimo, ossia gli estremi degli intervalli sugli assi x e y in cui si vuole disegnare il grafico
Si ricordi che è anche possibile cambiare colore e spessore delle linee, ecc.: cliccare due volte sulle linee e scegliere le modifiche nella finestra Formato serie dati, scheda Motivo

4 Cambiare i valori dei coefficienti; scegliere ad esempio:
m = 3 n = –2 a= 2 b=–2 c=–12
e osservare che tabelle e grafico vengono automaticamente aggiornati
Con il procedimento suggerito al punto 3 si realizzano i due grafici contemporaneamente.
Se si vogliono realizzare i grafici separatamente, si può procedere come segue:
Selezionare le celle B19:C39, scegliere il Tipo di grafico>Dispersione, al Passaggio 2 della composizione del grafico, nella scheda Serie premere il pulsante Aggiungi, nella casella Valori X selezionare le celle B19:B39, nella casella Valori Y selezionare le celle D19:D39



Torna su

Soluzione Esercizio 19
Funzione esponenziale e logaritmo

Indice

Ritorna Esercizio 19

Esercizio 19.1

1 Disegnare il grafico della funzione esponenziale   y=a^x , con a>0
2 Disegnare il grafico della funzione logaritmo $Y=\log_b(x)$, con b>0
3 Disegnare sullo stesso grafico le funzioni exp(x) e ln(x) (base e) e osservare la simmetria rispetto alla retta y=x (disegnare anche la retta)

base a	2
base b	2

SUGGERIMENTI

1 Scrivere il valore scelto per a nella cella J12
Costruire una tabella contenente i valori di ascisse e ordinate della funzione esponenziale (colonne B e C); far variare x nell'intervallo (-3,5) e usare un passo uguale a 0,5
Scrivere le formule nelle celle B36 e C35 facendo attenzione ai riferimenti assoluti e relativi; trascinare con il mouse per completare la tabella. Disegnare il grafico
Far variare a: scegliere valori di a >1, poi 0<a<1, infine a=1 e osservare l'aggiornamento automatico della tabella e del grafico
2 Con procedimento analogo disegnare il grafico della funzione logaritmo. Usare la funzione LOG che consente di scegliere la base b. Far variare x nell'intervallo (0,5 , 8)
Far variare b: scegliere valori di b >1, poi 0<b<1, e osservare l'aggiornamento automatico della tabella e del grafico
3 Completare la tabella 3 con i valori di x, y=exp(x), y=ln(x) e y=x
Disegnare i grafici nella stessa figura: osservare la simmetria rispetto alla bisettrice: da che cosa dipende tale simmetria?
La realizzazione di un grafico di buona qualità richiede la conoscenza delle proprietà delle due funzioni (dominio,…). La tabella 4 serve per completare il grafico nell'intervallo (0,1) in modo che la simmetria sia ben evidenziata. (il grafico in basso a destra è realizzato usando solo la tabella 3)

Tabella 1

x	y=a^x
-3	0,125
-2,5	0,176777
-2	0,25
-1,5	0,353553
-1	0,5
-0,5	0,707107
0	1
0,5	1,414214
1	2
1,5	2,828427
2	4
2,5	5,656854
3	8
3,5	11,31371
4	16
4,5	22,62742
5	32

Tabella 2

x	$y=\log_b(x)$
0,5	-1
1	0
1,5	0,584963
2	1
2,5	1,321928
3	1,584963
3,5	1,807355
4	2
4,5	2,169925
5	2,321928
5,5	2,459432
6	2,584963
6,5	2,70044
7	2,807355
7,5	2,906891
8	3

Tabella 3

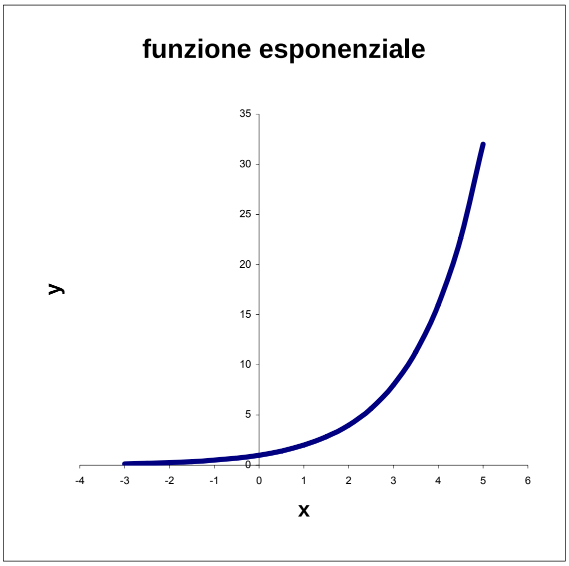
x	y=exp(x)	y=ln(x)	y=x
-3	0,049787		-3
-2,5	0,082085		-2,5
-2	0,135335		-2
-1,5	0,22313		-1,5
-1	0,367879		-1
-0,5	0,606531		-0,5
0	1		0
0,5	1,648721	-0,693147	0,5
1	2,718282	0	1
1,5	4,481689	0,405465	1,5
2	7,389056	0,693147	2
2,5	12,18249	0,916291	2,5
3	20,08554	1,098612	3
3,5	33,11545	1,252763	3,5
4	54,59815	1,386294	4
4,5		1,504077	4,5
5		1,609438	5
5,5		1,704748	5,5
6		1,791759	6
6,5		1,871802	6,5

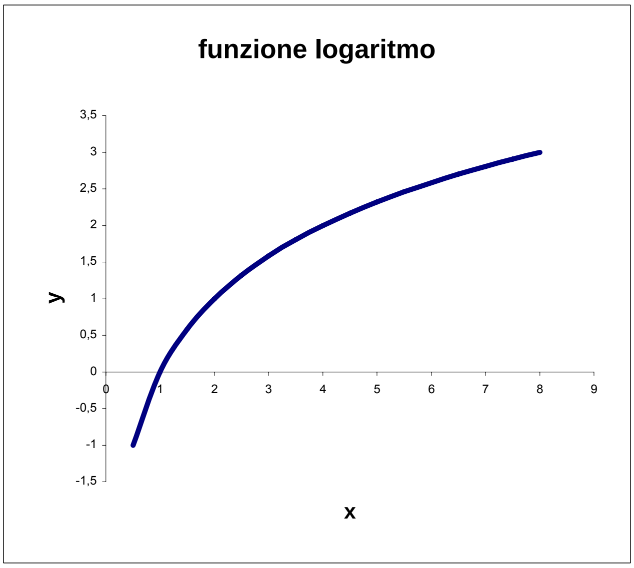
7		1,94591	7
7,5		2,014903	7,5
8		2,079442	8

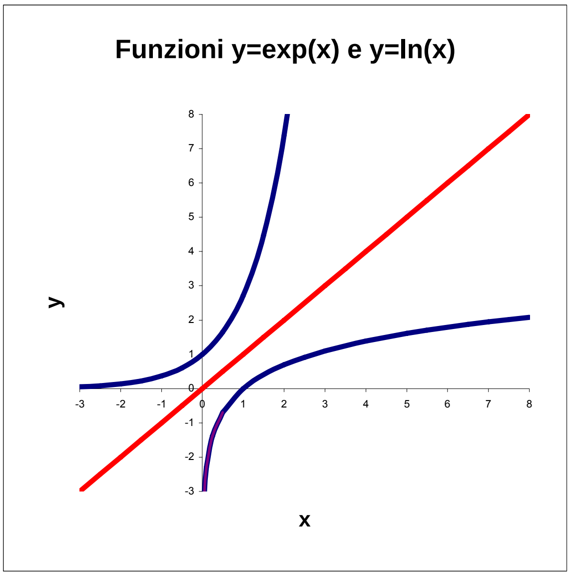
Tabella 4

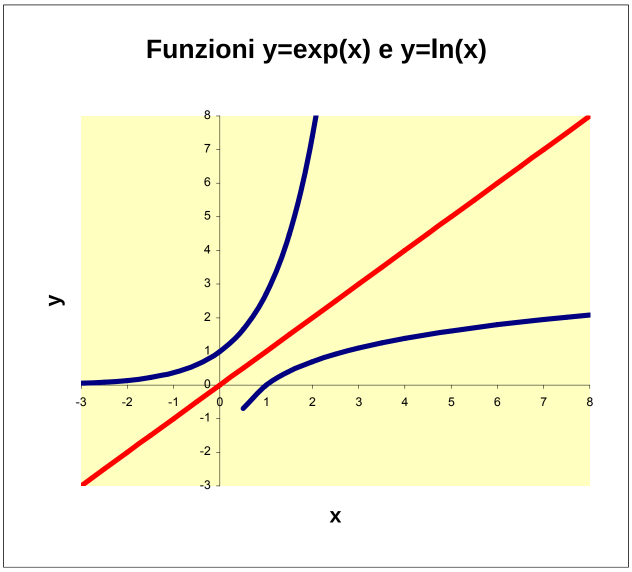
x	y=ln(x)
0,05	-2,995732
0,1	-2,302585
0,25	-1,386294
0,5	-0,693147

Grafici









Torna su

Soluzione Esercizio 20

Grafico di una funzione con punti di discontinuità (asintoti verticali)

Indice

Ritorna Esercizio 20

Esercizio 20.1

Disegnare il grafico della funzione $f(x)=\frac{\sqrt{x^2-1}}{x^2-4}$

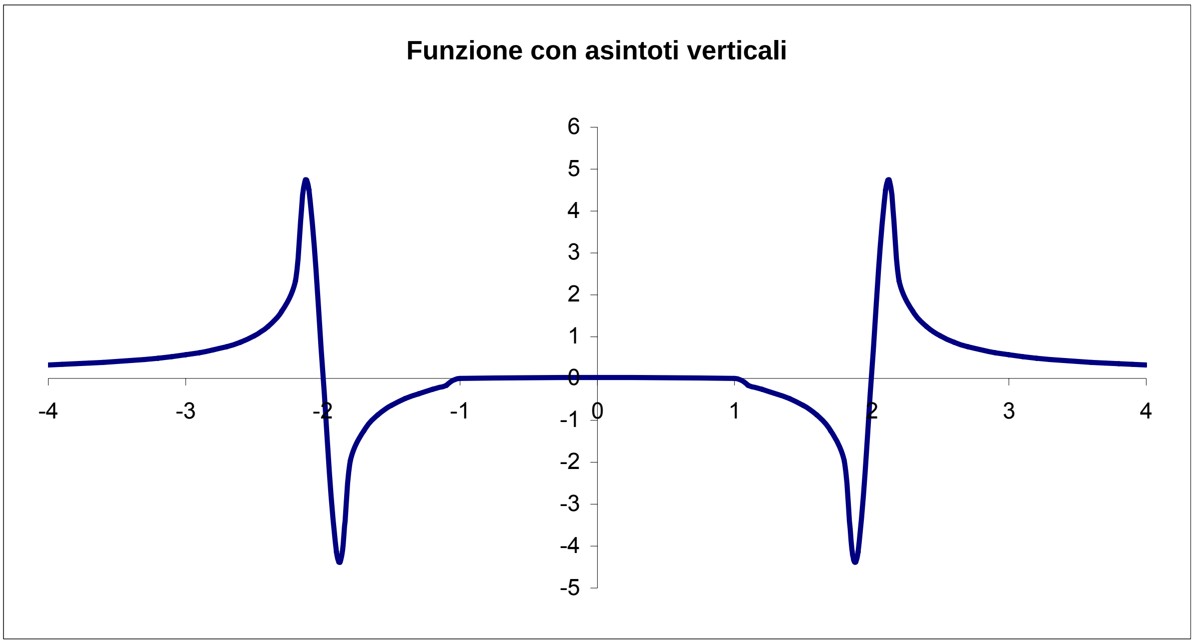
La funzione è definita per x<= –1 e x>=1 con x diverso da +2 e –2
Per disegnare il grafico costruire la tabella dei valori di x e y, facendo attenzione a non includere nella tabella i valori di x che non appartengono al dominio. Tracciare il grafico nell'intervallo (–4,4)

x	f(x)
-4	0,3227
-3,9	0,3363
-3,8	0,3512
-3,7	0,3676
-3,6	0,3860
-3,5	0,4066
-3,4	0,4298
-3,3	0,4564
-3,2	0,4871
-3,1	0,5230
-3	0,5657
-2,9	0,6173
-2,8	0,6811
-2,7	0,7623
-2,6	0,8696
-2,5	1,0184
-2,4	1,2396
-2,3	1,6056
-2,2	2,3328
-2,1	4,5039
-1,9	-4,1424
-1,8	-1,9693
-1,7	-1,2385
-1,6	-0,8674
-1,5	-0,6389
-1,4	-0,4803
-1,3	-0,3596
-1,2	-0,2591
-1,1	-0,1643
-1	0,0000
1	0,0000
1,1	-0,1643
1,2	-0,2591
1,3	-0,3596
1,4	-0,4803
1,5	-0,6389

1,6	-0,8674
1,7	-1,2385
1,8	-1,9693
1,9	-4,1424
2,1	4,5039
2,2	2,3328
2,3	1,6056
2,4	1,2396
2,5	1,0184
2,6	0,8696
2,7	0,7623
2,8	0,6811
2,9	0,6173
3	0,5657
3,1	0,5230
3,2	0,4871
3,3	0,4564
3,4	0,4298
3,5	0,4066
3,6	0,3860
3,7	0,3676
3,8	0,3512
3,9	0,3363
4	0,3227

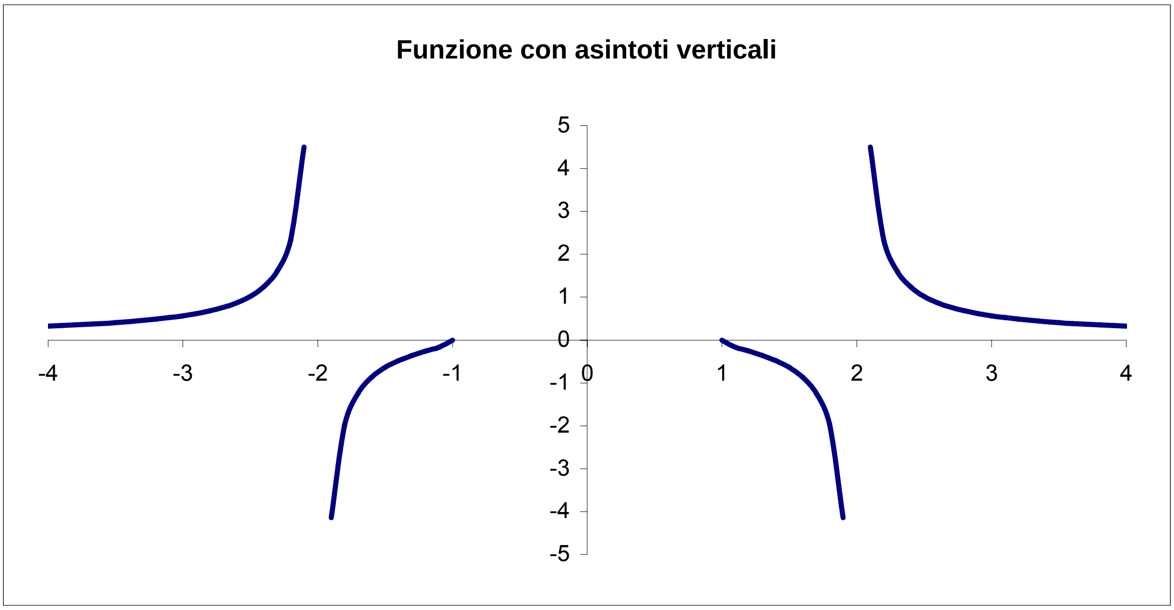
SUGGERIMENTI

Questo grafico non è corretto: i due segmenti verticali non devono essere disegnati
Non sono gli asintoti!! Cercare di capire perché vengono disegnati.
Neppure il segmento congiungente i punti (–1,0) e (1,0) deve essere disegnato!
Perché viene disegnato? Vedere dopo questo grafico come fare per ottenere il grafico corretto



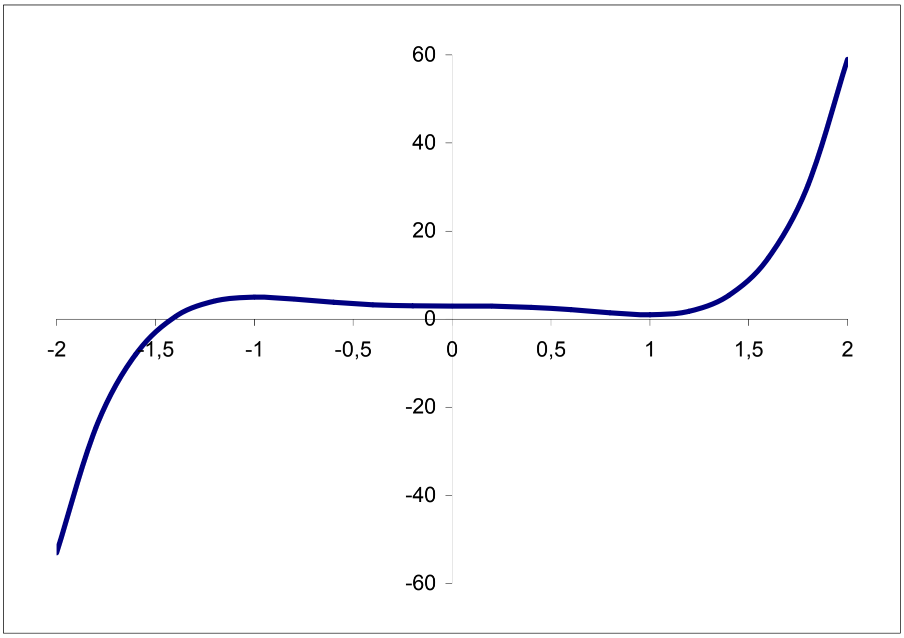
Il grafico seguente è corretto; per ottenerlo bisogna disegnarlo a tratti con il seguente procedimento
Selezionare le celle B20:C39 e iniziare a tracciare il grafico a dispersione;
Al Passaggio 2 cliccare sulla scheda Serie; cliccare su aggiungi; posizionarsi con il cursore nella casella Valori X, selezionare le celle B40:B49; nella casella Valori Y selezionare le celle C40:C49.

Cliccare ancora su aggiungi; Valori X: selezionare le celle B50:B59; Valori Y: selezionare le celle C50:C59.
Cliccare ancora su aggiungi; Valori X: selezionare le celle da B60:B79; Valori Y: selezionare le celle C60:C79. Terminare il grafico, poi perfezionarlo nell'aspetto estetico.



Esercizio 20.2

Disegnare il grafico della funzione $f(x) = 3x^5 - 5x^3 + 3$ nell'intervallo (-2,2)



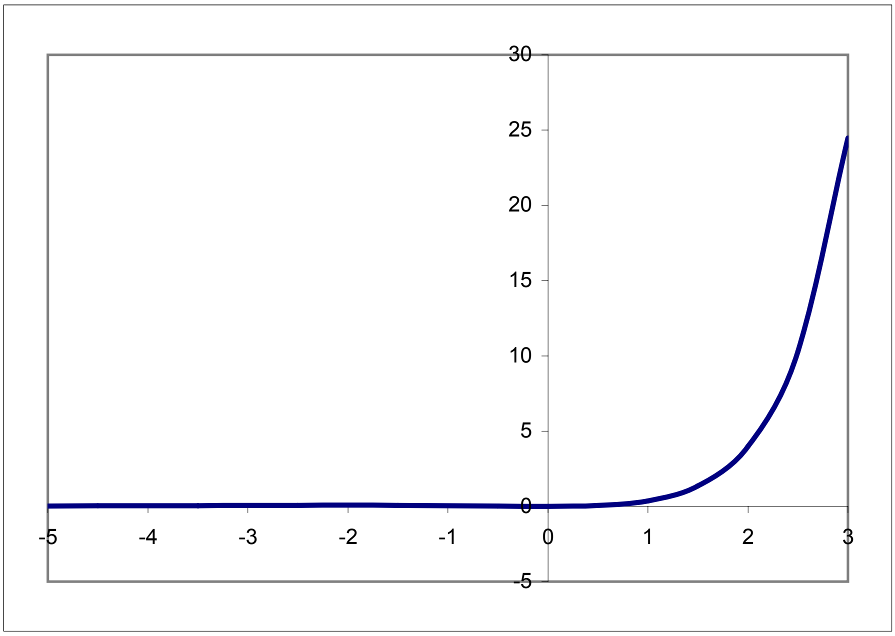
x	f(x)
-2	-53,00
-1,8	-24,53
-1,6	-7,98
-1,4	0,59
-1,2	4,18
-1	5,00
-0,8	4,58
-0,6	3,85
-0,4	3,29
-0,2	3,04
0	3,00
0,2	2,96
0,4	2,71
0,6	2,15
0,8	1,42
1	1,00
1,2	1,82
1,4	5,41
1,6	13,98
1,8	30,53
2	59,00

Esercizio 20.3

Disegnare il grafico della funzione $\boxed{f(x) = x^2 e^{x-2}}$ nell'intervallo (-5,3)

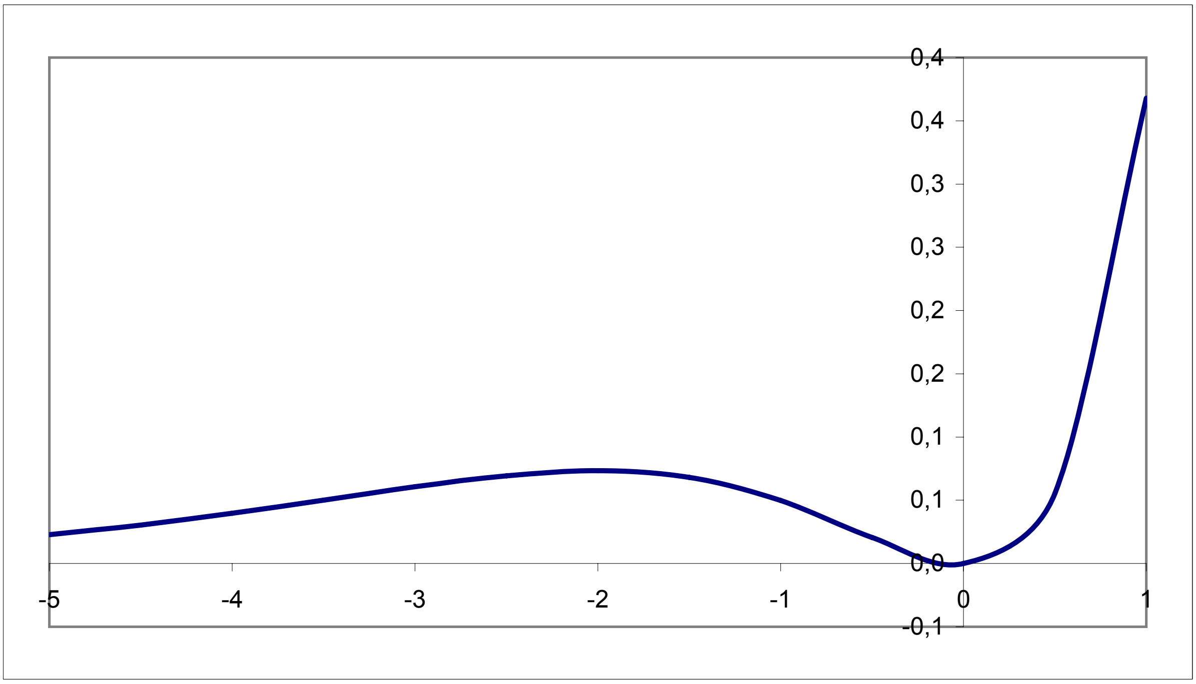
Osservare che il grafico ottenuto non descrive bene il comportamento di f(x); disegnare un altro grafico in un intervallo più adatto

Questo grafico non descrive bene il comportamento della funzione



x	f(x)
-5	0,0228
-4,5	0,0304
-4	0,0397
-3,5	0,0501
-3	0,0606
-2,5	0,0694
-2	0,0733
-1,5	0,0679
-1	0,0498
-0,5	0,0205
0	0,0000
0,5	0,0558
1	0,3679
1,5	1,3647
2	4,0000
2,5	10,3045
3	24,4645

Grafico corretto (si ottiene cambiando l'intervallo di valori di x)



Esercizio 20.4

Disegnare il grafico della funzione $f(x) = \dfrac{2x}{3 - 4x^2}$ nell'intervallo (-5,5)

Questo grafico non è corretto

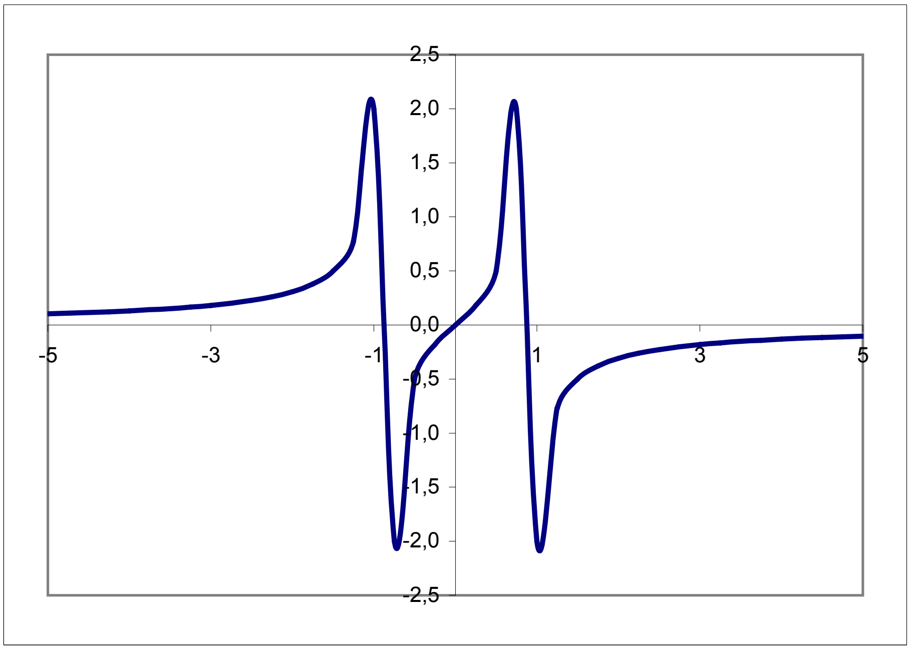
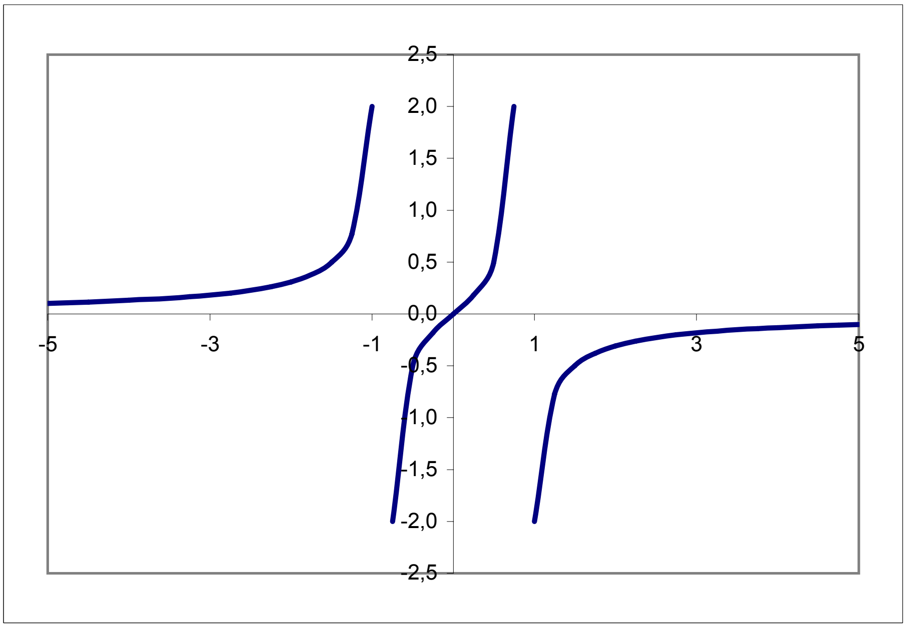


Grafico corretto
(bisogna eliminare le linee che appaiono come asintoti, ma non lo sono!)



x	f(x)
-5	0,1031
-4,75	0,1089
-4,5	0,1154
-4,25	0,1227
-4	0,1311
-3,75	0,1408
-3,5	0,1522
-3,25	0,1656
-3	0,1818
-2,75	0,2018
-2,5	0,2273
-2,25	0,2609
-2	0,3077
-1,75	0,3784
-1,5	0,5000
-1,25	0,7692
-1	2,0000
-0,75	-2,0000
-0,5	-0,5000
-0,25	-0,1818
0	0,0000
0,25	0,1818
0,5	0,5000
0,75	2,0000
1	-2,0000
1,25	-0,7692
1,5	-0,5000
1,75	-0,3784
2	-0,3077
2,25	-0,2609
2,5	-0,2273
2,75	-0,2018
3	-0,1818
3,25	-0,1656
3,5	-0,1522
3,75	-0,1408
4	-0,1311
4,25	-0,1227
4,5	-0,1154
4,75	-0,1089
5	-0,1031

Torna su

3. DISTRIBUZIONI DI FREQUENZA STATISTICHE

Soluzione Esercizio 21
Calcolo di media e varianza di un insieme di dati

Indice

Ritorna Esercizio 21

Esercizio 21.1

I dati contenuti nella Tabella 1 rappresentano il peso alla nascita di 20 bambini nati in una settimana in un ospedale
Nella cella D19 calcolare il valor medio dei dati; nelle celle D20 e D21 la varianza e lo scarto quadratico medio (deviazione standard)
Usare le funzioni del foglio elettronico: MEDIA, VAR, DEV.ST
Nella cella D23 calcolare il numero di dati della tabella con la funzione CONTA.NUMERI
Nelle celle D25:D27 calcolare media, varianza e scarto quadratico medio usando le formule che li definiscono (vedi Suggerimenti)

Valor medio	3241
Varianza	167546,32
Scarto quadratico medio	409,32
Numero dati	20
Valor medio (formula)	3241
Varianza (formula)	167546,32
Scarto quadr. medio (formula)	409,32

Tabella 1
3280
3260
3240
3480
4160
3320
3650
3200
3020
2580
2500
2840
3600
2840
3540
2760
3250
3320
3200
3780

SUGGERIMENTI

Formule per il calcolo di media e varianza di n dati x_i , $i=1,2,\ldots,n$

media
$$\boxed{\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i}$$

varianza
$$\boxed{s^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n\bar{x}^2\right)}$$

Torna su

Lo scarto quadratico medio è la radice quadrata della varianza

Soluzione Esercizio 22

Frequenze assolute, relative, percentuali. Frequenze cumulative
Diagrammi a barre. Grafici delle frequenze cumulative

Indice

Ritorna Esercizio 22

Esercizio 22.1

Variabile qualitativa

Nell'ambito di uno studio sui mezzi di trasporto usati dagli studenti universitari torinesi viene chiesto a un campione di 30 studenti qual è il mezzo di trasporto usato per recarsi in università.

I risultati sono riportati nella tabella 1

Tabella 1

auto	auto	mezzo pubblico	auto	bicicletta
a piedi	mezzo pubblico	mezzo pubblico	bicicletta	moto
auto	auto	bicicletta	moto	mezzo pubblico
moto	a piedi	mezzo pubblico	mezzo pubblico	mezzo pubblico
moto	auto	mezzo pubblico	auto	moto
moto	bicicletta	mezzo pubblico	moto	mezzo pubblico

Costruire la tabella delle frequenze assolute (tabella 2) con la funzione CONTA.SE (vedi Suggerimenti)

Disegnare un diagramma a barre della distribuzione di frequenza assoluta

Disegnare un diagramma circolare

Tabella 2

mezzo di trasporto	frequenza assoluta
auto	7
moto	7
mezzo pubblico	10
bicicletta	4
a piedi	2
totale	30

=CONTA.SE(B18:F23;"auto")

SUGGERIMENTI

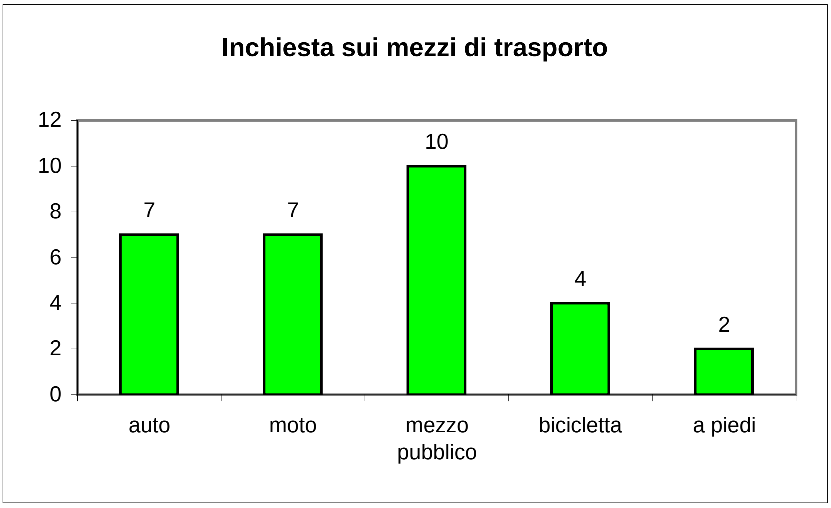
Per realizzare la Tabella 2 usare la funzione CONTA.SE, che conta il numero di celle di un dato intervallo che corrispondono a un dato criterio

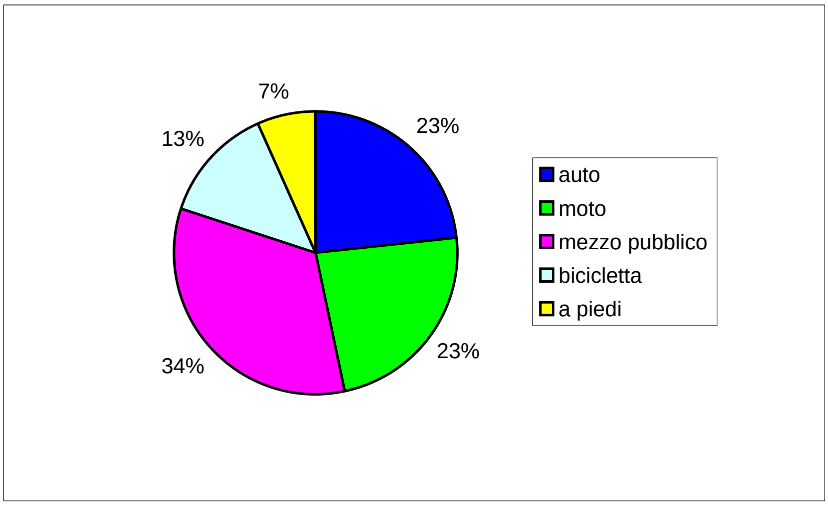
Sintassi

CONTA.SE(**intervallo**;**criteri**)

intervallo	intervallo delle celle contenenti i dati
criteri	criterio, scritto fra doppi apici " " (vedi commento)

Per ottenere la colonna delle frequenze assolute non si può usare il trascinamento, occorre scrivere singolarmente le formule in tutte le celle destinate alle frequenze assolute.





Esercizio 22.2
Variabile qualitativa
Una società di marketing ha condotto un'indagine sui consumatori che usano sempre la stessa marca di dentifricio. Dopo aver provato un nuovo tipo di dentifricio da poco immesso sul mercato, al campione di consumatori è stato chiesto di rispondere alla domanda "potresti decidere di acquistare il nuovo prodotto?"
La risposta è stata valutata su una scala da 1 a 5, con il seguente criterio
A - "non lo comprerei mai"
B - "lo comprerei raramente"
C - "mi è indifferente comprarlo o no"
D - "lo comprerei qualche volta"
E - "lo comprerei sempre"
I dati del sondaggio sono riportati nella Tabella 3

Tabella 3

A	E	D	A	D
E	D	B	B	D
C	C	A	D	B
E	D	B	D	C
A	C	A	C	B
B	C	E	D	E
D	D	C	D	D

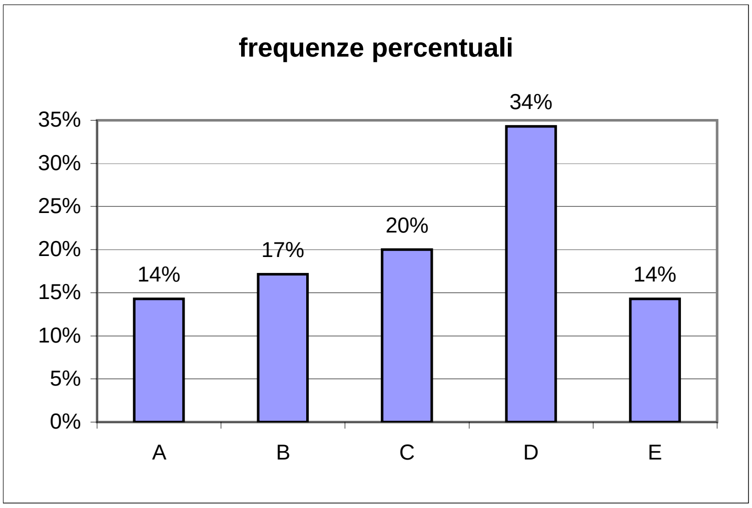
Costruire la tabella delle frequenze assolute (tabella 4) con la funzione CONTA.SE
Calcolare le frequenze relative e percentuali.
Disegnare un diagramma a barre della distribuzione di frequenza percentuale e un diagramma a torta, mostrando le percentuali

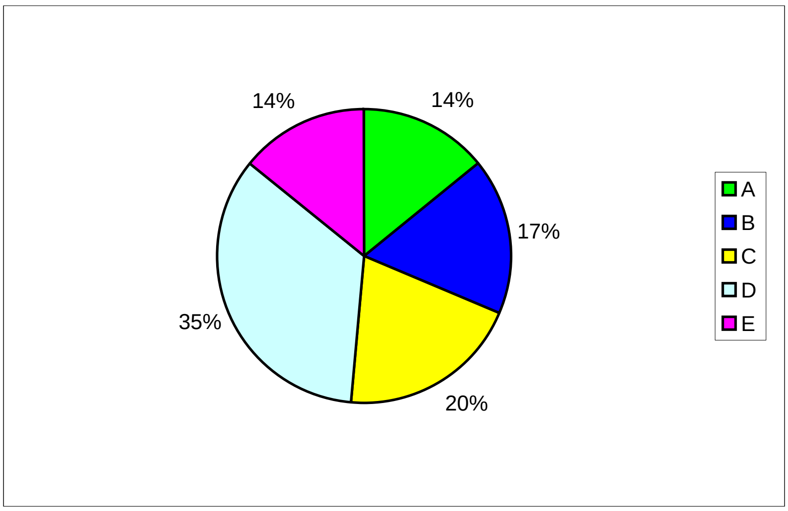
SUGGERIMENTI

Calcolare le frequenze relative dividendo ogni frequenza assoluta per il numero totale dei dati (Attenzione al riferimento assoluto)
Scegliere il formato percentuale per le celle delle frequenze percentuali; assegnare i valori delle frequenze relative alle celle delle frequenze percentuali con una formula, poi trascinare con il mouse nelle celle successive
Per disegnare il diagramma a torta selezionare la colonna delle frequenze assolute e mostrare come etichette le percentuali

Tabella 4

risposta	frequenza assoluta	frequenza relativa	frequenza percentuale
A	5	0,14	14%
B	6	0,17	17%
C	7	0,20	20%
D	12	0,34	34%
E	5	0,14	14%
totale	35		





Esercizio 22.3

Variabile numerica discreta

La tabella 5 contiene la distribuzione dei punteggi ottenuti con 500 lanci di due dadi.
Disegnare l'istogramma; inserire sull'asse orizzontale le etichette corrispondenti ai punteggi
Calcolare le frequenze relative e percentuali
Costruire la tabella delle frequenze cumulative
Disegnare il grafico delle frequenze cumulative

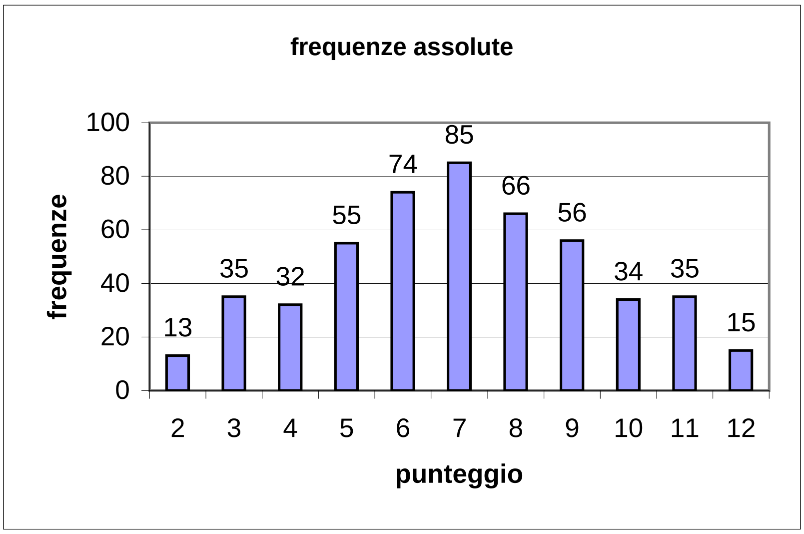
Tabella 5

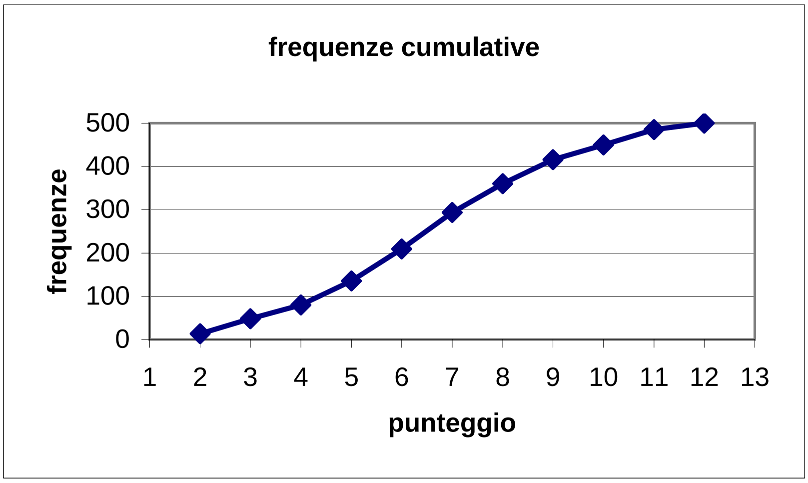
punteggio	frequenza assoluta	frequenza relativa	frequenza percentuale
2	13	0,026	2,6%
3	35	0,07	7,0%
4	32	0,064	6,4%
5	55	0,11	11,0%
6	74	0,148	14,8%
7	85	0,17	17,0%
8	66	0,132	13,2%
9	56	0,112	11,2%
10	34	0,068	6,8%
11	35	0,07	7,0%
12	15	0,03	3,0%
Totale	500	1	100,0%

SUGGERIMENTI

Per calcolare le frequenze cumulative usare le fomule, procedendo nel modo seguente:
la prima frequenza cumulativa è uguale alla prima frequenza assoluta;
calcolare la seconda frequenza cumulativa con una formula opportuna: la seconda frequenza cumulativa è la somma della frequenza cumulativa precedente e della seconda frequenza assoluta (vedere i commenti alle celle)
trascinare con il mouse la formula della seconda frequenza cumulativa nelle celle successive
Nella colonna C sono indicate le classi per le frequenze cumulative: si intende che ogni classe è del tipo "minore o uguale a…."
Per disegnare il grafico della distribuzione qualitativa usare il tipo di grafico Dispersione (XY) (coordinate unite da linee)

punteggio	frequenza cumulativa
2	13
3	48
4	80
5	135
6	209
7	294
8	360
9	416
10	450
11	485
12	500


frequenze assolute
frequenze
100
80
60
40
20
0
13
35
32
55
74
85
66
56
34
35
15
2
3
4
5
6
7
8
9
10
11
12
punteggio


frequenze cumulative
frequenze
500
400
300
200
100
0
1
2
3
4
5
6
7
8
9
10
11
12
13
punteggio

Torna su

Soluzione Esercizio 23

Costruzione di una tabella di distribuzione di frequenza. Istogramma

Indice

Ritorna Esercizio 23

Esempio 23.1

Variabile numerica continua

Nella Tabella 1 sono riportate le misure dell'emissione giornaliera di gas inquinanti in un impianto industriale

1 Calcolare il numero dei dati, il minimo e il massimo dei dati e il campo di variazione
2 Calcolare il numero di classi per costruire la distribuzione di frequenza assoluta.
3 Raggruppare i dati in una distribuzione di frequenza con il numero di classi scelto.
Scegliere le classi (chiuse a destra e di uguale ampiezza) e usare la funzione **FREQUENZA**
4 Calcolare le frequenze relative e percentuali
5 Disegnare l'istogramma delle frequenze assolute.

Tabella 1

15,8	24,6	24,8	13,5
22,7	19,4	26,1	24,6
26,8	12,3	20,9	20
19,1	15,9	21,4	24,1
18,5	11,2	18	9
14,4	14,7	24,3	17,6
8,3	20,5	11,8	16,7
25,9	26,6	17,9	16,9
26,4	20,1	18,7	23,5
9,8	17	12,8	18,4
22,7	22,3	15,5	25,7
15,2	27,5	19,2	20,1
23	23,9	7,7	13,2
29,6	17,5	22,5	23,7
21,9	11	19,3	10,7
10,5	20,4	9,4	19
17,3	16,2	13,9	14,5
6,2	20,8	28,6	18,1
18	13,3	19,4	31,8
22,9	18,1	21,6	28,5

SUGGERIMENTI

1 Usare le funzioni CONTA.NUMERI, MIN, MAX
Il campo di variazione è la differenza tra il minimo e il massimo dei dati
2 Per il calcolo del numero di classi usare la seguente regola empirica:
n = numero dati
k = numero classi

$$\boxed{k = 1 + 3{,}322\ Log_{10}(n)}$$

Il numero di classi deve essere un intero: usare la funzione INT per arrotondare (per difetto) il valore trovato con la regola empirica.
L'ampiezza delle classi può essere trovata dividendo il campo di variazione per il numero di classi e arrotondando per eccesso con la funzione ARROTONDA.ECCESSO

Sintassi

ARROTONDA.ECCESSO(**num**; **peso**)

num — numero da arrotondare
peso — indica a quale multiplo intero del parametro peso si vuole arrotondare il numero

Esempi

Per arrotondare il numero 7,83 all'intero successivo 8 usare peso = 1

7,83	8

=ARROTONDA.ECCESSO(C58;1)

Per arrotondare il numero 17,65 al più vicino multiplo di 10 usare peso = 10

17,65	20

=ARROTONDA.ECCESSO(C62;10)

Per arrotondare il numero 227,65 al più vicino multiplo di 100 usare peso = 100

227,65	300

=ARROTONDA.ECCESSO(D79;100)

3 Scegliere l'estremo destro della prima classe, costruire il successivo estremo destro con una formula in base all'ampiezza scelta; trascinare con il mouse per ottenere gli altri estremi
Controllare che le classi scelte comprendano tutti i dati del campione
Calcolare i valori centrali: calcolare il primo valore con una formula e trascinare nelle celle successive
Per calcolare le frequenze assolute si usa la funzione FREQUENZA

Sintassi
FREQUENZA(**matrice_dati**;**matrice_classi**)
matrice_dati   intervallo di celle contenente i dati
matrice_classi   intervallo di celle contenente gli estremi destri delle classi

Calcolare le frequenze assolute con la funzione FREQUENZA con il seguente procedimento: selezionare **tutte le celle** in cui dovranno comparire le frequenze assolute, premere il pulsante Incolla Funzione e scegliere la funzione FREQUENZA
Nella finestra della funzione FREQUENZA posizionare il cursore nella casella Matrice_dati e selezionare le celle della tabella dei dati; nella casella Matrice_classi selezionare le celle degli estremi destri delle classi.

La funzione FREQUENZA è una funzione di tipo matrice: richiede un uso diverso dalle altre funzioni. Per calcolare le frequenze premere Ctrl+Maiuscolo+Invio (e non solo Invio oppure OK)

Se si è operato correttamente, la formula deve comparire racchiusa tra due parentesi graffe nella barra della formula e le celle D119:D125 devono contenere tutte le frequenze

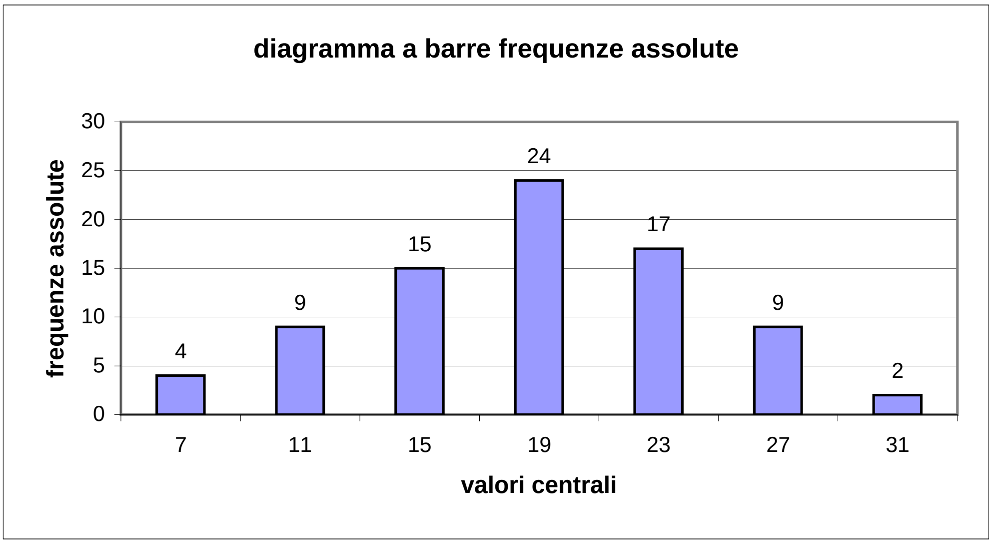
4 Calcolare le frequenze relative dividendo ogni frequenza assoluta per il numero totale dei dati (Attenzione al riferimento assoluto)
Scegliere il formato percentuale per le celle F119:F125; assegnare i valori delle frequenze relative alle celle delle frequenze percentuali nel modo seguente: nella cella F119 scrivere la formula =E119, poi trascinare con il mouse nelle celle successive

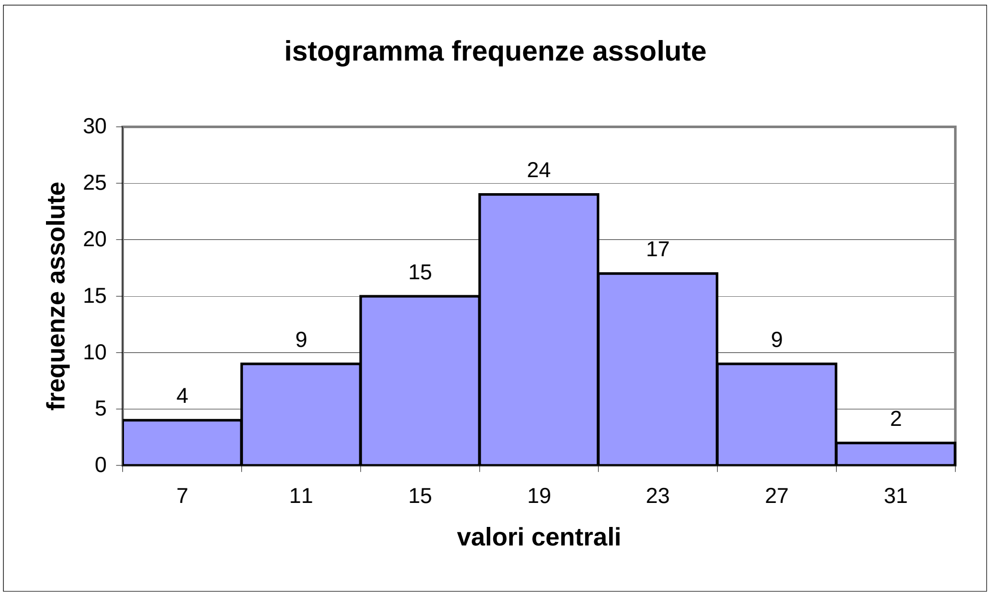
5 **Istogramma**: selezionare le frequenze assolute (non il totale!) e disegnare l'istogramma con Creazione guidata Grafico, Tipo di grafico: Istogramma
Per inserire i valori centrali sull'asse X: selezionare l'area del grafico (quadratini neri sugli spigoli) e cliccare sul pulsante Creazione guidata Grafico; al Passaggio 2 della creazione del grafico selezionare la scheda Serie e nella casella Etichette assi categorie (X) inserire le etichette selezionando le celle dei valori centrali; concludere la realizzazione del grafico che viene aggiornato con la nuova scelta delle etichette
Excel disegna di default un diagramma a barre (primo grafico);
per ottenere un istogramma occorre modificare la larghezza delle barre (secondo grafico)
Si ricordi che il diagramma a barre è utilizzato per variabili discrete, mentre l'istogramma è utilizzato per variabili continue (come in questo esempio)
Per cambiare la larghezza delle barre del diagramma, puntare con il mouse su una delle barre e premere il pulsante destro, selezionare Formato serie dati, Scheda Opzioni, Distanza tra le barre: aumentare o diminuire la distanza.
Con lo stesso procedimento si può cambiare il colore delle barre, il tipo di riempimento, … (scheda Motivo)

numero dati	80
minimo	6,2
massimo	31,8

campo variazione	25,6
numero classi	7
ampiezza classi	4

classi	estremi destri	frequenza assoluta	frequenza relativa	frequenza percentuale	valori centrali
1	9	4	0,05	5%	7
2	13	9	0,1125	11%	11
3	17	15	0,1875	19%	15
4	21	24	0,3	30%	19
5	25	17	0,2125	21%	23
6	29	9	0,1125	11%	27
7	33	2	0,025	3%	31
Totali		80	1	100%	





Osservazione.
La scelta delle classi (ampiezza e estremi destri) è un punto importante: la funzione FREQUENZA può lavorare in modo da evitare il rischio che qualche dato del campione sia fuori dalle classi scelte e non venga conteggiato
Procedendo come segue, tutti i dati del campione vengono conteggiati e non è più necessario controllare che le classi scelte comprendano tutti i dati del campione

Scegliamo ad esempio sette classi, di ampiezza 3,5 e estremi destri indicati nelle celle D177:D183: queste classi non contengono tutti i dati: ci sono dati minori dell'estremo sinistro 6,5 della prima classe e dati maggiori dell'estremo destro 30 dell'ultima classe
Dopo aver scelto le classi (estremi destri D177:D183), selezionare le celle che dovranno contenere tutte le frequenze assolute prendendo **una cella in più** (E177:E184): la cella E184 conterrà il numero di dati maggiori dell'estremo destro dell'ultima classe; la prima cella E177 contiene in effetti il numero di tutti i dati minori dell'estremo destro della prima classe

estremo destro	frequenza assoluta
9	4
12,5	8
16	12
19,5	21
23	16
26,5	12
30	6
	1

Esercizio 23.2

Variabile numerica continua

Sono assegnati i dati della tabella 2 (misure di peso in g)

1 Calcolare il numero dei dati, il minimo e il massimo dei dati e il campo di variazione
Calcolare media, varianza e scarto quadratico medio dei dati
2 Determinare la distribuzione di frequenza assoluta, raggruppando i dati in un numero adatto di classi chiuse a destra di uguale ampiezza
3 Determinare le distribuzioni di frequenza relativa e percentuale.
4 Disegnare il grafico della distribuzione di frequenza assoluta

SUGGERIMENTI

Valgono i suggerimenti dell'esempio precedente.
L'ampiezza si ottiene calcolando prima il quoziente fra campo di variazione e numero di classi e poi arrotondando al multiplo di 10 con la funzione ARROTONDA.ECCESSO (peso = 10)
Numerare le classi con le etichette 1, 2, …

Tabella 2

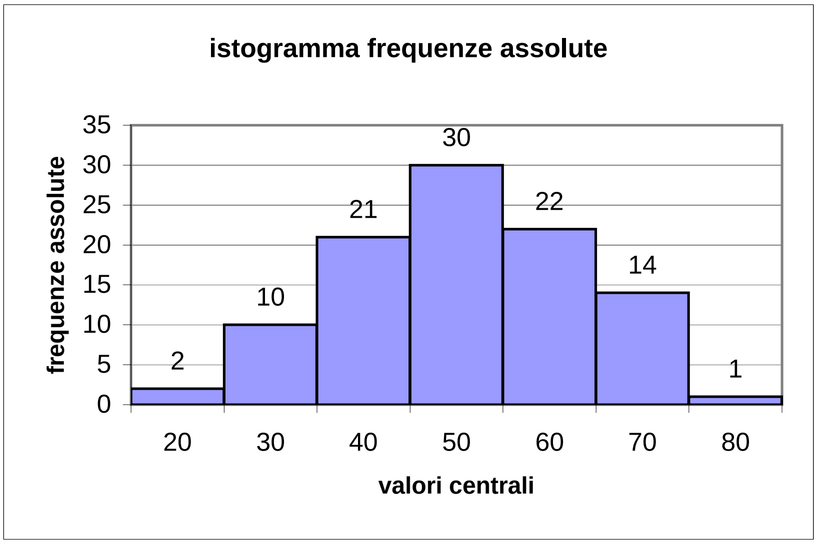
64	76	43	59	49
69	51	53	50	37
67	38	54	69	62
41	58	52	28	35
55	71	49	37	41
71	57	29	40	35
52	51	48	55	41
75	48	26	44	43
63	75	57	50	60
38	64	65	46	68
49	53	64	50	51
62	41	25	30	39
72	71	65	58	41
27	35	73	47	65
47	49	42	40	60
47	49	42	40	60
39	32	66	47	53
59	68	33	38	24
61	51	54	58	56
39	52	56	67	54

numero dati	100
minimo	24
massimo	76

campo variazione	52
numero classi	7
ampiezza classi	10

media	51,10
varianza	165,00
scarto quadratico medio	12,85

classi	estremi destri	frequenza assoluta	frequenza relativa	frequenza percentuale	valori centrali
1	25	2	0,02	2%	20
2	35	10	0,1	10%	30
3	45	21	0,21	21%	40
4	55	30	0,3	30%	50
5	65	22	0,22	22%	60
6	75	14	0,14	14%	70
7	85	1	0,01	1%	80
Totali		100	1	100%	



Torna su

Soluzione Esercizio 24

Costruzione di una tabella di distribuzione di frequenza. Istogramma

Indice

Ritorna Esercizio 24

Esercizio 24.1

Variabile numerica continua

Nella tabella 1 sono riportati i pesi alla nascita di 100 bambini.

1 Calcolare la media e lo scarto quadratico medio dei dati.

2 Calcolare il numero di dati, il minimo e il massimo dei dati e il campo di variazione

3 Raggruppare i dati in una distribuzione di frequenza con un adeguato numero di classi. Scegliere un'ampiezza multipla di 100 per semplicità

4 Disegnare un istogramma delle frequenze assolute.

5 Costruire la distribuzione di frequenza cumulativa assoluta e disegnare il grafico.

Tabella 1

2720	1640	3340	2600	3060
3600	2340	2440	3260	3340
1200	3480	1800	2660	1900
3280	2940	3740	2780	4120
3260	3440	1940	3040	2360
4560	940	2200	3500	2960
2300	2580	3460	4100	2800
1980	2940	2260	1900	2980
3200	3620	3000	3540	3060
2400	3780	3260	3600	3820
2120	3740	900	3980	3900
2380	2700	2360	3180	3620
3060	3500	4380	2960	2840
3500	1740	2640	2400	2660
3260	3580	2480	2520	3060
2860	3540	2880	3460	3880
1080	3260	2940	2760	2520
980	4080	2460	2480	2920
3100	2780	3760	2940	2360
1800	2520	3440	3180	4100

SUGGERIMENTI

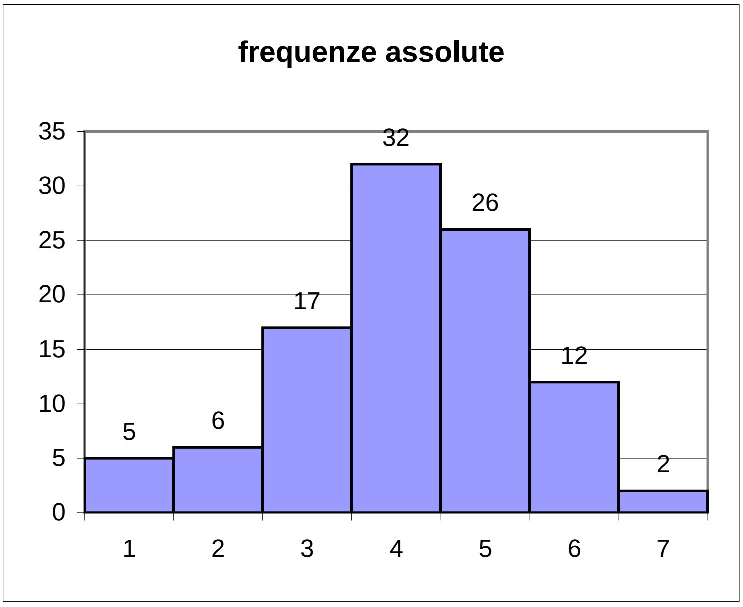
Valgono i suggerimenti dell'esercizio precedente.
L'ampiezza si può scegliere calcolando il quoziente fra campo di variazione e numero di classi e poi arrotondando al multiplo di 100 con la funzione ARROTONDA.ECCESSO (peso = 100)
Numerare le classi con le etichette 1, 2, …

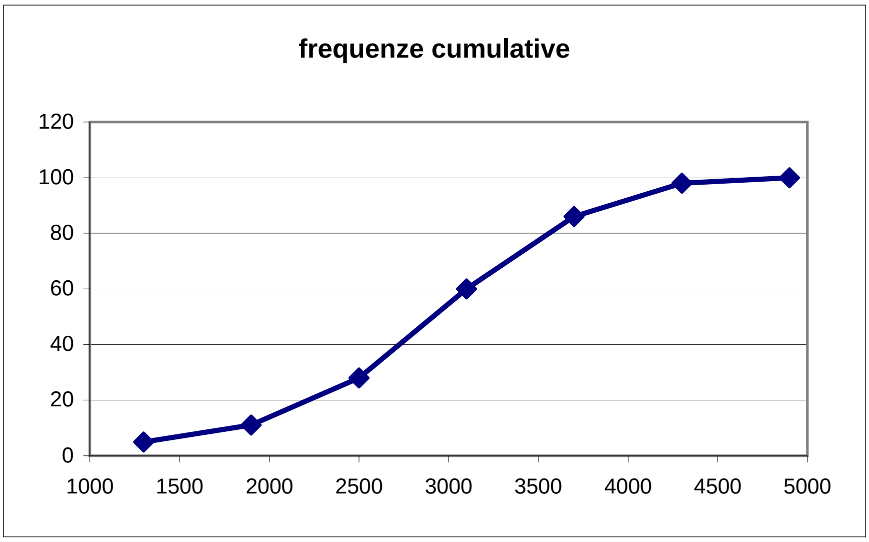
minimo	900
massimo	4560
range	3660

numero dati	100
numero classi	7
ampiezza classi	600
media	2905,60
varianza	581083,47
scarto quadr. medio	762,29

classi	estremo destro	frequenza assoluta	frequenza relativa	frequenza percentuale	valori centrali
1	1300	5	0,05	5%	1100
2	1900	6	0,06	6%	1700
3	2500	17	0,17	17%	2300
4	3100	32	0,32	32%	2900
5	3700	26	0,26	26%	3500
6	4300	12	0,12	12%	4100
7	4900	2	0,02	2%	4700
Totali		100	1	100%	

classi	frequenza cumulativa
1300	5
1900	11
2500	28
3100	60
3700	86
4300	98
4900	100





Torna su

Soluzione Esercizio 25
Costruzione di tabelle di distribuzione di frequenza e grafici

Indice

Ritorna Esercizio 25

Esercizio 25.1

Variabile numerica continua

Sono assegnati i dati della tabella 1

1 Calcolare il numero di dati, il minimo e il massimo dei dati e il campo di variazione
2 Calcolare media, varianza e scarto quadratico medio del campione di dati
3 Costruire una distribuzione di frequenza assoluta, raggruppando i dati in 6 classi
4 Costruire le distribuzione di frequenza relativa e percentuale.
5 Disegnare l'istogramma della distribuzione di frequenza assoluta
6 Costruire la distribuzione di frequenza cumulativa assoluta e disegnare il grafico.

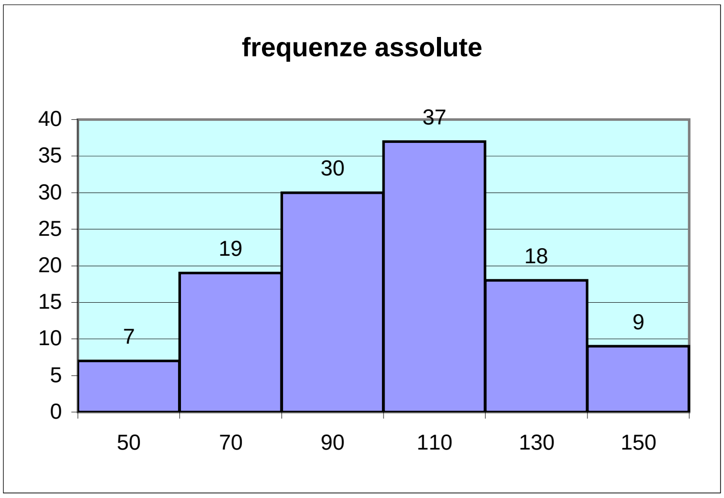
Tabella 1

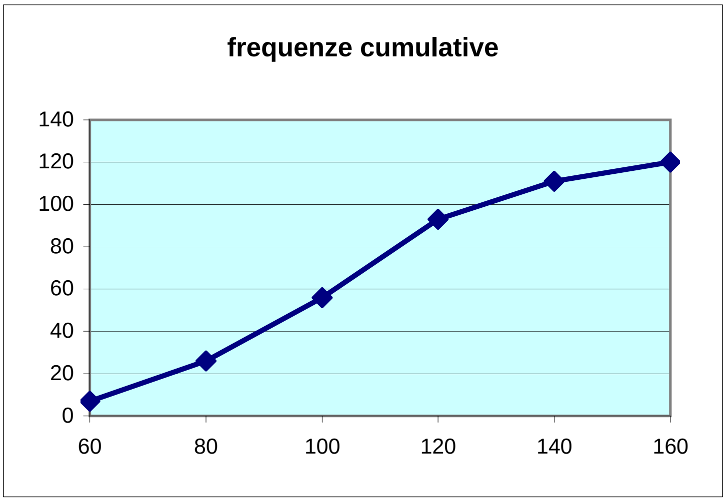
128	152	87	118	97
138	102	106	100	74
134	76	109	138	123
81	115	104	57	71
111	142	99	74	82
142	114	59	80	70
73	119	108	160	126
105	102	96	110	82
150	96	52	88	86
67	151	114	100	120
76	128	130	92	136
119	101	108	124	116
99	86	128	100	103
123	82	91	59	78
144	143	130	117	81
87	102	120	67	75
118	137	80	109	117
121	101	85	112	135
79	104	97	107	113
110	58	128	42	108
79	63	146	95	98
87	94	94	85	80
105	120	130	95	98
105	106	80	55	70

numero dati	120
minimo	42
massimo	160
campo di variazione	118
media	102,33
varianza	614,79
scarto quadr. medio	24,80
ampiezza classi	20

classi	estremo destro	frequenza assoluta	frequenza relativa	frequenza percentuale	valori centrali
1	60	7	0,058	5,8%	50
2	80	19	0,158	15,8%	70
3	100	30	0,250	25,0%	90
4	120	37	0,308	30,8%	110
5	140	18	0,150	15,0%	130
6	160	9	0,075	7,5%	150
Totali		120	1,000	100,0%	

classi	frequenza cumulativa
x<=60	7
x<=80	26
x<=100	56
x<=120	93
x<=140	111
x<=160	120





Esercizio 25.2

Variabile numerica continua

Viene condotta un'indagine sulla modalità con cui si distribuiscono i ritardi di alcuni treni a lunga percorrenza in arrivo nella stazione di Torino nell'arco di due settimane.
I dati della tabella 2 rappresentano i minuti di ritardo

1 Calcolare il numero dei dati, il minimo e il massimo dei dati
Calcolare media, varianza e scarto quadratico medio dei dati
2 Costruire una distribuzione di frequenza assoluta, raggruppando i dati in classi chiuse a destra di uguale ampiezza e disegnare il grafico della distribuzione di frequenza assoluta
3 Trovare le distribuzioni di frequenza relativa e percentuale.
4 Trovare la distribuzione di frequenza cumulativa assoluta e disegnare il grafico.

Tabella 2

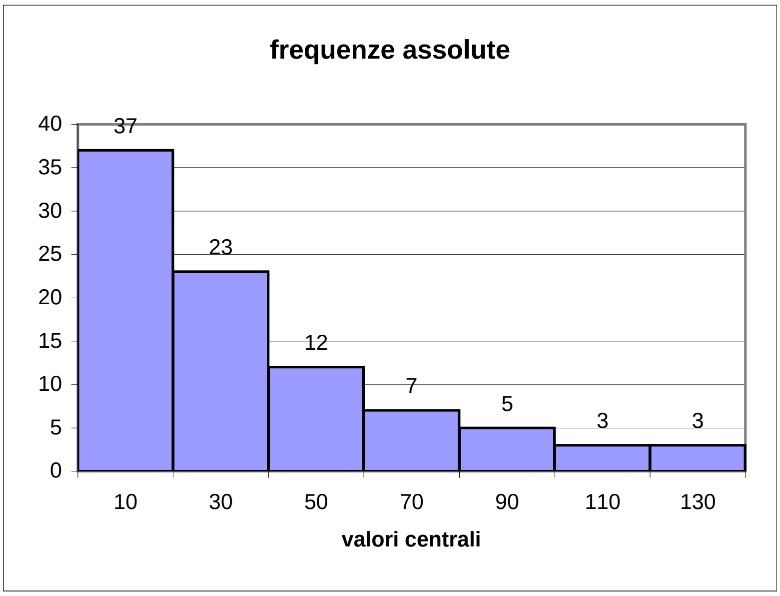
50	6	67	50	36	15
57	56	31	5	8	4
102	32	39	34	116	29
32	6	5	7	7	5
10	16	5	49	138	46
36	8	12	24	140	33
22	22	16	8	22	32
6	76	27	17	33	54
121	26	16	7	6	4
23	97	11	5	36	41
31	16	119	54	16	23
36	47	94	68	6	34
14	46	17	69	69	90
18	7	44	4	91	67
9	6	4	88	6	64

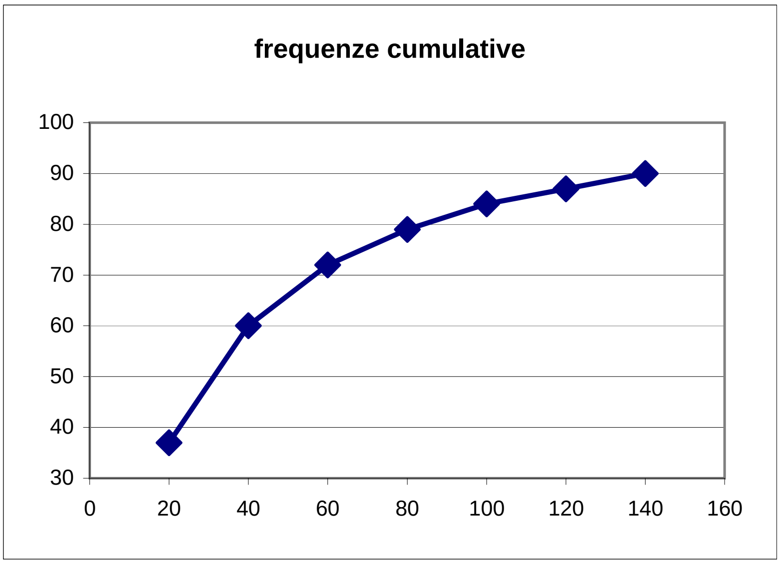
numero dati	90
minimo	4
massimo	140
media	36,68
varianza	1131,16
scarto quadratico medio	33,63

range	136
numero classi	7
ampiezza	20

classi	estremo destro	frequenza assoluta	frequenza relativa	frequenza percentuale	valori centrali
1	20	37	0,41	41%	10
2	40	23	0,26	26%	30
3	60	12	0,13	13%	50
4	80	7	0,08	8%	70
5	100	5	0,06	6%	90
6	120	3	0,03	3%	110
7	140	3	0,03	3%	130
Totali		90	1	100%	

classi	frequenza cumulativa
x<=20	37
x<=40	60
x<=60	72
x<=80	79
x<=100	84
x<=120	87
x<=140	90





Esercizio 22.3

Variabile numerica discreta

Nella tabella 5 sono riportati gli stipendi base dei dipendenti di un gruppo di filiali di una grande banca. Gli stipendi base sono classificati secondo sei livelli contrattuali, riportati nella tabella 4

1 Calcolare lo stipendio medio
2 Costruire una distribuzione di frequenza assoluta, raggruppando i dati in classi chiuse a destra di uguale ampiezza e disegnare il grafico della distribuzione di frequenza assoluta
3 Trovare le distribuzioni di frequenza relativa e percentuale.
4 Trovare la distribuzione di frequenza cumulativa assoluta e disegnare il grafico.

Tabella 4

classe 1	30000
classe 2	35000
classe 3	40000
classe 4	45000
classe 5	50000
classe 6	55000

Tabella 5

30000	35000	35000	35000	35000
30000	35000	35000	35000	35000
30000	35000	35000	35000	35000
30000	35000	35000	35000	35000
30000	35000	35000	35000	35000
30000	35000	35000	35000	35000
30000	35000	35000	35000	35000
30000	35000	35000	35000	35000
30000	35000	35000	35000	35000
30000	35000	35000	35000	35000
35000	35000	35000	40000	40000
40000	40000	40000	40000	40000
40000	40000	40000	40000	40000
40000	40000	40000	40000	40000
40000	40000	40000	40000	40000
40000	40000	40000	40000	40000
40000	40000	40000	40000	40000
40000	40000	40000	40000	40000
40000	40000	40000	40000	40000
40000	40000	40000	40000	40000
40000	40000	40000	40000	40000
40000	40000	40000	40000	40000
45000	45000	40000	45000	45000
45000	45000	45000	45000	45000
45000	45000	45000	45000	50000
50000	50000	50000	50000	55000
35000	35000	35000	35000	35000
35000	35000	35000	35000	35000
35000	35000	35000	35000	35000
35000	35000	35000	35000	35000
35000	35000	35000	35000	35000
35000	35000	35000	35000	35000
35000	35000	35000	35000	35000
35000	35000	35000	35000	35000
35000	35000	35000	35000	35000
35000	35000	35000	35000	35000
40000	40000	40000	40000	40000
40000	40000	40000	40000	40000
40000	40000	40000	40000	40000
40000	40000	40000	40000	40000
40000	40000	40000	40000	40000
40000	40000	40000	40000	40000
40000	40000	40000	40000	40000
40000	40000	40000	40000	40000
40000	40000	40000	40000	40000
40000	40000	40000	40000	40000
40000	40000	40000	40000	40000
40000	45000	45000	45000	45000
45000	45000	45000	45000	45000
45000	45000	45000	45000	45000
50000	50000	50000	50000	45000
55000	55000	55000	55000	55000

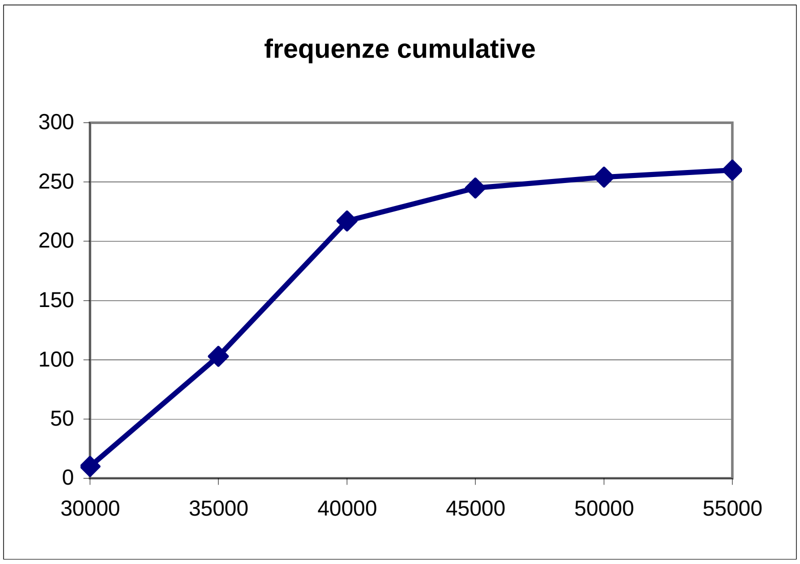
SUGGERIMENTI

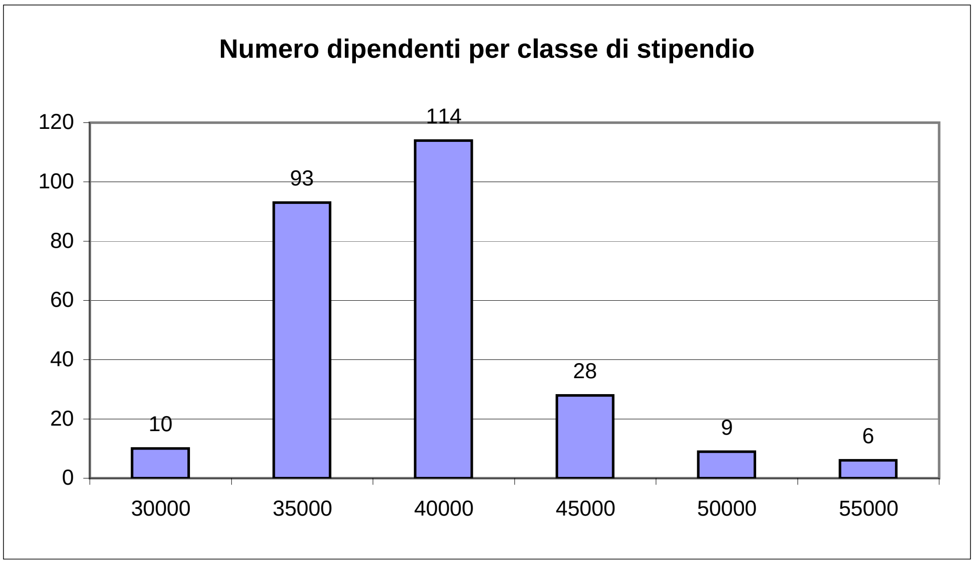
I dati della tabella 5 sono valori assunti da una variabile discreta; per costruire la distribuzione di frequenza si può usare la funzione FREQUENZA con le stesse modalità già illustrate negli esempi precedenti.
Scegliere sei classi, identificate dai sei valori degli stipendi contrattuali.
In questo caso, trattandosi di una variabile discreta, l'estremo "destro" della classe è l'unico valore assunto dalla variabile "stipendio base" in quella classe

Stipendio medio	39058

classi	estremo destro	frequenza assoluta	frequenza relativa	frequenza percentuale
1	30000	10	0,038	4%
2	35000	93	0,358	36%
3	40000	114	0,438	44%
4	45000	28	0,108	11%
5	50000	9	0,035	3%
6	55000	6	0,023	2%
Totali		260	1,000	100%

classi	frequenza cumulativa
30000	10
35000	103
40000	217
45000	245
50000	254
55000	260





Torna su

Soluzione Esercizio 26

Distribuzioni di frequenza. Istogrammi, confronto fra ampiezze diverse

Indice

Ritorna Esercizio 26

Esercizio 26.1

Variabile numerica continua

La tabella 1 contiene le misure dei diametri di 100 sferette in cm.

1 Costruire la distribuzione di frequenza assoluta scegliendo più ampiezze diverse per le classi (di conseguenza si utilizzano più o meno classi)

Trovare minimo e massimo dei dati e usare per le ampiezze i valori seguenti

0,4 0,3 0,2 0,1 0,08

L'estremo destro della prima classe si stabilisce in base al minimo dei dati

Il numero di classi viene stabilito di conseguenza, in base all'ampiezza scelta

2 Confrontare gli istogrammi nei vari casi e scegliere il numero di classi più adatto

In base al confronto degli istogrammi si può decidere qual è il numero migliore di classi?

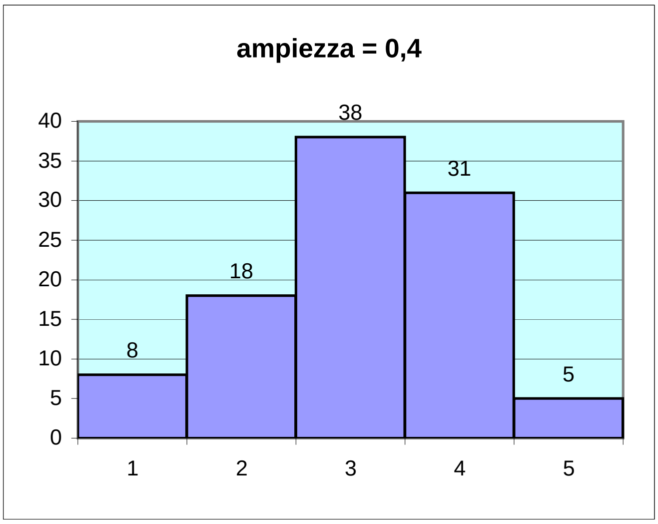
Tabella 1

1,81	2,13	1,28	2	2,28
1,26	1,4	2,12	1,86	2,36
2,06	2,32	1,53	2,49	2,42
2,13	2,73	2,63	1,16	1,56
1,49	1,69	1,64	2,19	2,09
2,53	2,38	2,24	2,4	2,11
2,53	2,56	2,1	2,33	1,55
1,98	1,29	1,59	2,26	1,67
2,15	1,36	1,03	2,02	2,48
2,08	2,26	1,97	2,3	1,94
1,92	1,82	1,55	2,25	2,17
2,32	2,31	2,27	1,89	2,04
1,74	2,36	2,23	1,83	1,72
2,98	2,32	2,76	1,87	1,75
1,94	2,58	2,26	1,34	2,2
2,05	2,3	1,71	1,9	1,58
2,48	2,53	2,17	2,05	2,35
2,03	1,46	1,55	2,14	2,25
1,96	1,99	1,99	2,65	1,63
1,63	1,93	1,98	1,84	1,88

minimo	1,03
massimo	2,98

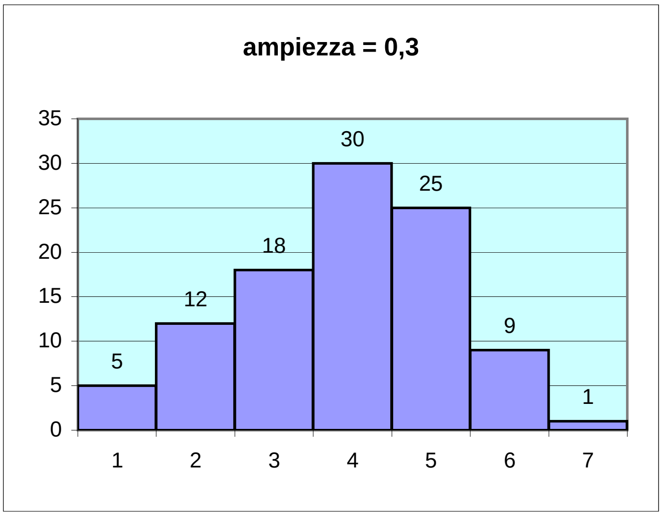
ampiezza	**0,4**

classi (estremo destro)	**frequenze assolute**
1,4	8
1,8	18
2,2	38
2,6	31
3	5
Totale	100



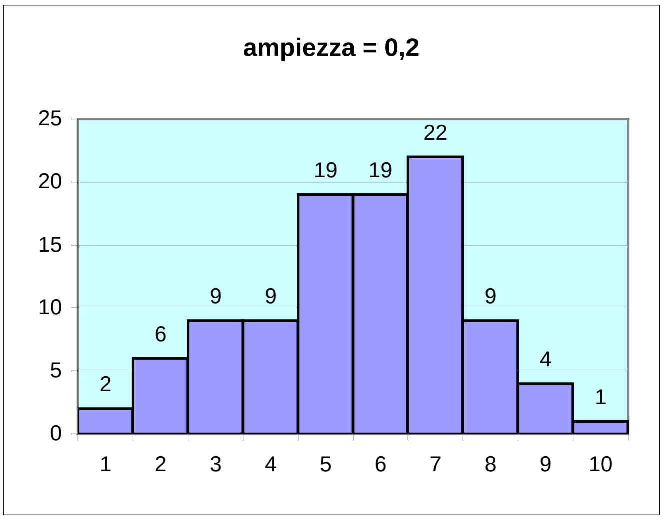
ampiezza	**0,3**

classi (estremo destro)	**frequenze assolute**
1,3	5
1,6	12
1,9	18
2,2	30
2,5	25
2,8	9
3,1	1
Totale	100



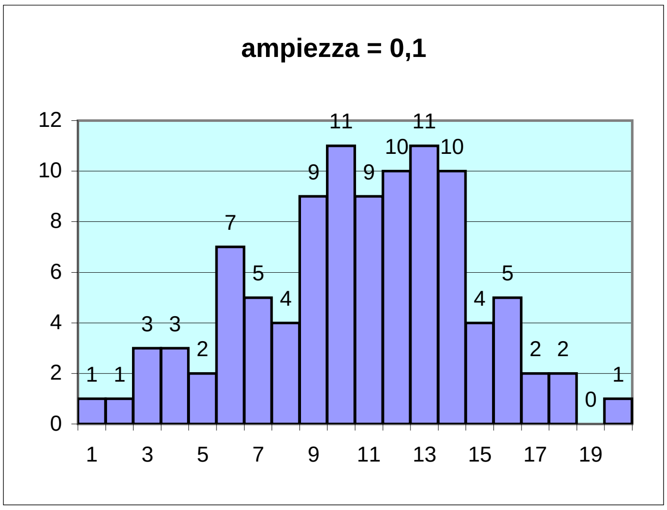
ampiezza	**0,2**

classi (estremo destro)	**frequenze assolute**
1,2	2
1,4	6
1,6	9
1,8	9
2	19
2,2	19
2,4	22
2,6	9
2,8	4
3	1
Totale	100



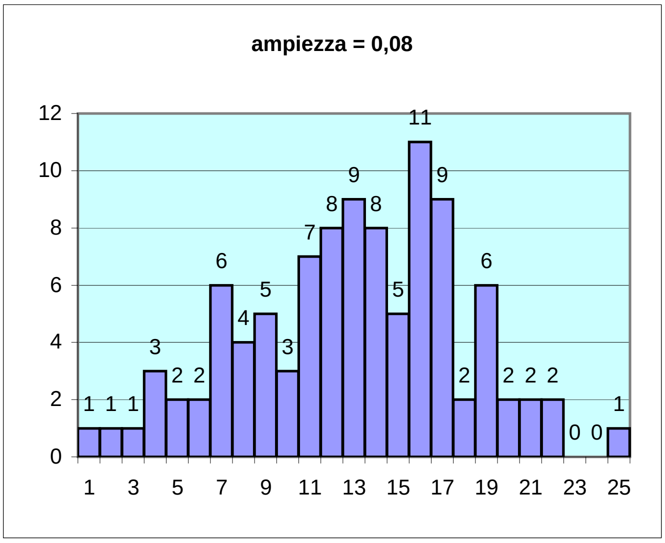
ampiezza	**0,1**

classi (estremo destro)	**frequenze assolute**
1,1	1
1,2	1
1,3	3
1,4	3
1,5	2
1,6	7
1,7	5
1,8	4
1,9	9
2	11
2,1	9
2,2	10
2,3	11
2,4	10
2,5	4
2,6	5
2,7	2
2,8	2
2,9	0
3	1
Totale	100



ampiezza	0,08

classi (estremo destro)	frequenze assolute
1,1	1
1,18	1
1,26	1
1,34	3
1,42	2
1,5	2
1,58	6
1,66	4
1,74	5
1,82	3
1,9	7
1,98	8
2,06	9
2,14	8
2,22	5
2,3	11
2,38	9
2,46	2
2,54	6
2,62	2
2,7	2
2,78	2
2,86	0
2,94	0
3,02	1
Totale	100



L'ampiezza più adatta è 0,3, corrispondente a 7 classi

Torna su

Soluzione Esercizio 27
Calcolo di media e varianza per dati raggruppati

Indice

Ritorna Esercizio 27

Esercizio 27.1

Nell'esercizio 23 è stata calcolata la distribuzione di frequenza sotto riportata per i dati della tabella 1

1 Calcolare media e varianza usando i dati raggruppati

Questo problema non può essere risolto con le funzioni di Excel, che non prevedono il calcolo di media e varianza per dati raggruppati in classi di frequenza: si devono usare le formule (riportate nei suggerimenti).

2 Calcolare media e varianza con le funzioni di Excel usando la tabella dei dati

Tabella 1

15,8	24,6	24,8	13,5
22,7	19,4	26,1	24,6
26,8	12,3	20,9	20
19,1	15,9	21,4	24,1
18,5	11,2	18	9
14,4	14,7	24,3	17,6
8,3	20,5	11,8	16,7
25,9	26,6	17,9	16,9
26,4	20,1	18,7	23,5
9,8	17	12,8	18,4
22,7	22,3	15,5	25,7
15,2	27,5	19,2	20,1
23	23,9	7,7	13,2
29,6	17,5	22,5	23,7
21,9	11	19,3	10,7
10,5	20,4	9,4	19
17,3	16,2	13,9	14,5
6,2	20,8	28,6	18,1
18	13,3	19,4	31,8
22,9	18,1	21,6	28,5

ampiezza classi	4

classi	estremo destro	frequenza assoluta f_i	valori centrali m_i	fi*mi	f_i*m_i^2
5<x<=9	9	4	7	28	196
9<x<=13	13	9	11	99	1089
13<x<=17	17	15	15	225	3375
17<x<=21	21	24	19	456	8664
21<x<=25	25	17	23	391	8993
25<x<=29	29	9	27	243	6561
29<x<=33	33	2	31	62	1922
Totali		80		1504	30800

media dai dati raggruppati	18,8
varianza dai dati raggruppati	31,9595
media calcolata dai dati	18,90
varianza calcolata dai dati	31,9956

SUGGERIMENTI

Formule per il calcolo di media e varianza usando i dati raggruppati:

Valor medio

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k} m_i f_i$$

Varianza

$$s^2 = \frac{1}{n-1}\left[\sum_{i=1}^{k} f_i m_i^2 - n\bar{x}^2\right]$$

m_i	valori centrali delle classi
f_i	frequenze assolute di ogni classe
n	numero dei dati
k	numero delle classi

Esercizio 27.2

E' data la distribuzione di frequenza assoluta della tabella 2

Tabella 2

classi	frequenza assoluta
40<x<=50	2
50<x<=60	14
60<x<=70	29
70<x<=80	35
80<x<=90	28
90<x<=100	10
100<x<=110	2

Calcolare la media e la varianza con le formule dei dati raggruppati
Calcolare lo scarto quadratico medio

ampiezza classi	10

classi	estremo destro	frequenza assoluta f_i	valori centrali m_i	fi*mi	f_i*m_i^2
40<x<=50	50	2	45	90	4050
50<x<=60	60	14	55	770	42350
60<x<=70	70	29	65	1885	122525
70<x<=80	80	35	75	2625	196875
80<x<=90	90	28	85	2380	202300
90<x<=100	100	10	95	950	90250
100<x<=110	110	2	105	210	22050
	Totali	120		8910	680400

media dai dati raggruppati	74,25
varianza dai dati raggruppati	158,256
scarto quadratico medio	12,58

Torna su

Soluzione Esercizio 28

Calcolo di media e varianza e loro utilizzo

Indice

Ritorna Esercizio 28

Esercizio 28.1

Per la partecipazione a una gara di matematica una scuola deve formare una squadra di 6 studenti; con una selezione preliminare, attraverso un test con un punteggio massimo di 100 punti, sulla base della media dei migliori punteggi risultano tre squadre a pari merito.
Con quale criterio può essere scelta la squadra da mandare alla gara?
Calcolare i valori di media, varianza e scarto quadratico medio con le funzioni di Excel (MEDIA, VAR, DEV.STD)

squadra	punteggi degli studenti					
A	73	76	77	85	88	90
B	74	74	78	84	88	91
C	72	77	79	82	84	95

media squadra A	81,5
media squadra B	81,5
media squadra C	81,5

varianza squadra A	49,9
varianza squadra B	52,7
varianza squadra C	61,1

scarto quadratico medio squadra A	7,06
scarto quadratico medio squadra B	7,26
scarto quadratico medio squadra C	7,82

Criterio di scelta: si sceglie la squadra che ha il minor scarto quadratico medio (squadra A)
(o equivalentemente la minor varianza)

Torna su

Soluzione Esercizio 29

Calcolo di percentili e quartili

Indice

Ritorna Esercizio 29

Per il calcolo di percentili e quartili si usano le funzioni PERCENTILE e QUARTILE

Sintassi

PERCENTILE(matrice;k)

Matrice	intervallo di celle contenente i dati
k	valore percentile (valore in decimale compreso fra 0 e 1)
	Esempio: k=0,25 per il 25° percentile

QUARTILE(matrice, quarto)

Matrice	intervallo di celle contenente i dati
quarto	numero del quartile (1 per il primo, ecc.)

La mediana (secondo quartile e 50-esimo percentile) si calcola anche con la funzione MEDIANA

Esercizio 29.1

Calcolare i percentili indicati per i dati della tabella 1

Tabella 1

32,2
32
30,4
31
31,2
31,3
30,3
29,6
30,5
30,7

25° percentile	30,425
75° percentile	31,275

Mediana	30,85	con la funzione MEDIANA
Mediana	30,85	con la funzione PERCENTILE come 50° percentile

Esercizio 29.2

Calcolare i percentili e i quartili indicati per i dati della tabella 2

Tabella 2

6,2	7,7	8,3	9	9,4
11,8	12,3	12,8	13,2	13,3
15,2	15,5	15,8	15,9	16,2
17,6	17,9	18	18	18,1
19,1	19,2	19,3	19,4	19,4
20,8	20,9	21,4	21,6	21,9
23	23,5	23,7	23,9	24,1
25,9	26,1	26,4	26,6	26,8
9,8	10,5	10,7	11	11,2
13,5	13,9	14,4	14,5	14,7
16,7	16,9	17	17,3	17,5
18,1	18,4	18,5	18,7	19
20	20,1	20,1	20,4	20,5
22,3	22,5	22,7	22,7	22,9
24,3	24,6	24,6	24,8	25,7
27,5	28,5	28,6	29,6	31,8

Primo quartile	15,08	
25-esimo percentile	15,08	
Terzo quartile	22,93	
Mediana	19,05	(come secondo quartile, con la funzione QUARTILE)
95-esimo percentile	27,55	

Torna su

Soluzione Esercizio 30
Strumenti Analisi Dati
Statistica descrittiva, Istogramma

Indice

Ritorna Esercizio 30

Gli **Strumenti Analisi Dati** sono uno dei Componenti Aggiuntivi inclusi in Excel
(Vedere Esercizio 1, Esempio 8, per l'eventuale installazione degli Strumenti Analisi Dati)
Gli strumenti di analisi consentono di ridurre i passaggi necessari per sviluppare analisi statistiche o ingegneristiche.
Una volta forniti i dati e i parametri per ciascuna analisi, lo strumento utilizzerà le funzioni statistiche o ingegneristiche appropriate, visualizzando i risultati in una tabella di output.
Alcuni strumenti generano anche dei grafici.
Esaminiamo in questi esercizi alcuni di questi strumenti, utili per le applicazioni statistiche
Statistica descrittiva è lo strumento di analisi che permette di calcolare i diversi indici di posizione e dispersione
Istogramma è lo strumento che permette di costruire la distribuzione di frequenza e l'istogramma per un insieme di dati

Esercizio 30.1
Nella tabella 1 sono riportate le misure dell'emissione giornaliera di gas inquinanti in un impianto industriale (Vedi Esercizio 23)
Usare gli Strumenti **Statistica Descrittiva** e **Istogramma** per calcolare le statistiche, la distribuzione di frequenza e l'istogramma per i dati della tabella

Tabella 1

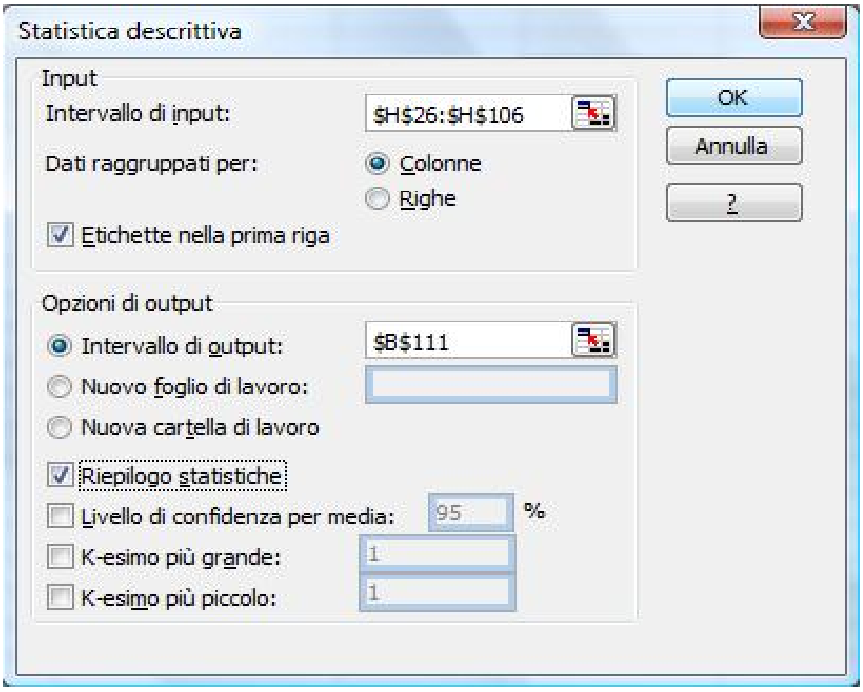
15,8	24,6	24,8	13,5
22,7	19,4	26,1	24,6
26,8	12,3	20,9	20
19,1	15,9	21,4	24,1
18,5	11,2	18	9
14,4	14,7	24,3	17,6
8,3	20,5	11,8	16,7
25,9	26,6	17,9	16,9
26,4	20,1	18,7	23,5
9,8	17	12,8	18,4
22,7	22,3	15,5	25,7
15,2	27,5	19,2	20,1
23	23,9	7,7	13,2
29,6	17,5	22,5	23,7
21,9	11	19,3	10,7
10,5	20,4	9,4	19
17,3	16,2	13,9	14,5
6,2	20,8	28,6	18,1
18	13,3	19,4	31,8
22,9	18,1	21,6	28,5

Dati
15,8
22,7
26,8
19,1
18,5
14,4
8,3
25,9
26,4
9,8
22,7
15,2
23
29,6
21,9
10,5
17,3
6,2
18
22,9
24,6
19,4
12,3
15,9
11,2
14,7
20,5
26,6

SUGGERIMENTI
Strumenti Analisi: Statistica Descrittiva
Lo strumento Statistica Descrittiva richiede che i dati siano disposti in un'unica colonna (o riga): disporre i dati in colonna copiandoli nella colonna H
(usare il Menu Modifica>Incolla speciale>Valori, se non si vuole modificare la formattazione delle celle della colonna H)

Dal Menu Strumenti scegliere Analisi dati > Statistica descrittiva
Riempire la finestra di dialogo secondo le indicazioni seguenti (figura 1)
Premere il tasto Tab per passare al campo successivo nella finestra di dialogo
Intervallo di input: selezionare con il mouse le celle contenenti i dati in colonna
Dati raggruppati per: scegliere Colonne (i dati sono stati copiati nella colonna H)
Selezionare **Etichette nella prima riga** se si vuole far comparire l'etichetta
nella tabella di output (in questo esempio il testo Dati, presente nella cella H26)
Nelle **opzioni di output** (Intervallo di output) indicare l'indirizzo della prima
cella (in alto a sinistra) in cui si vuole far comparire l'output, oppure scegliere
Nuovo foglio di lavoro per disporre l'output in un nuovo foglio.
Selezionare **Riepilogo statistiche** per ottenere fra i risultati di output i valori
delle statistiche (media, mediana, varianza, ecc.)
Livello di confidenza per media: selezionare questa opzione se si vuole
ottenere l'intervallo di confidenza per la media (vedere Stima dei parametri, es. 62)
Dopo aver riempito i campi necessari, cliccare su OK
Il risultato è una tabella contenente le statistiche dei dati
Se necessaio formattare la tabella di output per visualizzarla meglio:
allargare la colonna B in modo da adattarla al testo più lungo contenuto nella tabella
Con il pulsante Diminuisci decimali ridurre il numero dei decimali visualizzati.
I risultati ottenuti con lo strumento di analisi sono statici: se si cambiano i dati
della colonna H, i risultati delle colonne B:C non vengono automaticamente
aggiornati; per avere i risultati aggiornati bisogna utillizzare di nuovo lo strumento
Statistica descrittiva.

Figura 1



20,1
17
22,3
27,5
23,9
17,5
11
20,4
16,2
20,8
13,3
18,1
24,8
26,1
20,9
21,4
18
24,3
11,8
17,9
18,7
12,8
15,5
19,2
7,7
22,5
19,3
9,4
13,9
28,6
19,4
21,6
13,5
24,6
20
24,1
9
17,6
16,7
16,9
23,5
18,4
25,7
20,1
13,2
23,7
10,7
19
14,5
18,1
31,8
28,5

Dati	
Media	18,89625
Errore standard	0,632411608
Mediana	19,05
Moda	22,7
Deviazione standard	5,656461383
Varianza campionaria	31,99555538
Curtosi	-0,498253885
Asimmetria	-0,102510166
Intervallo	25,6
Minimo	6,2
Massimo	31,8
Somma	1511,7
Conteggio	80

Notare che fra i risultati compaiono gli indici di Curtosi e Asimmetria, non descritti in questi esercizi; inoltre compare l'errore standard, definito come segue

$$\boxed{\text{errore standard} = \frac{\text{deviazione standard}}{\sqrt{\text{numero dati}}}}$$

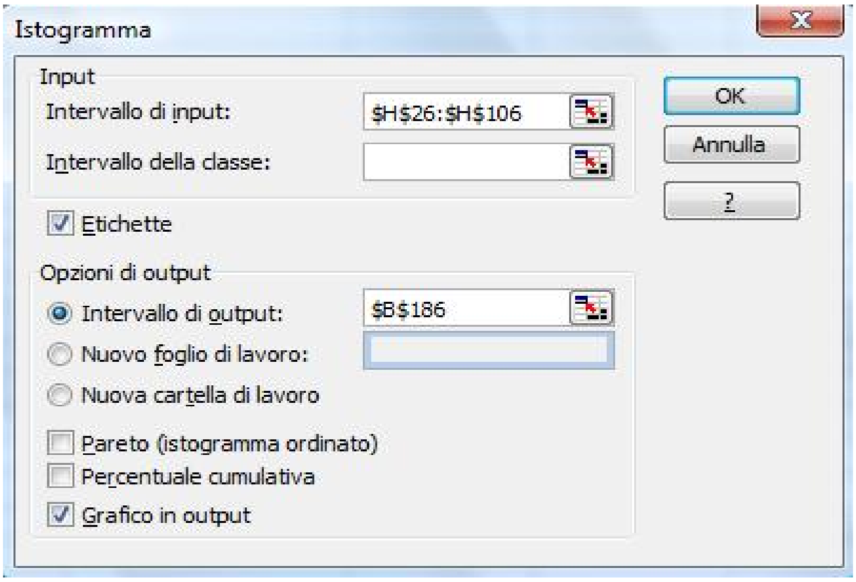
L'errore standard viene utillizzato per l'inferenza statistica (intervalli di confidenza e test di ipotesi).

SUGGERIMENTI

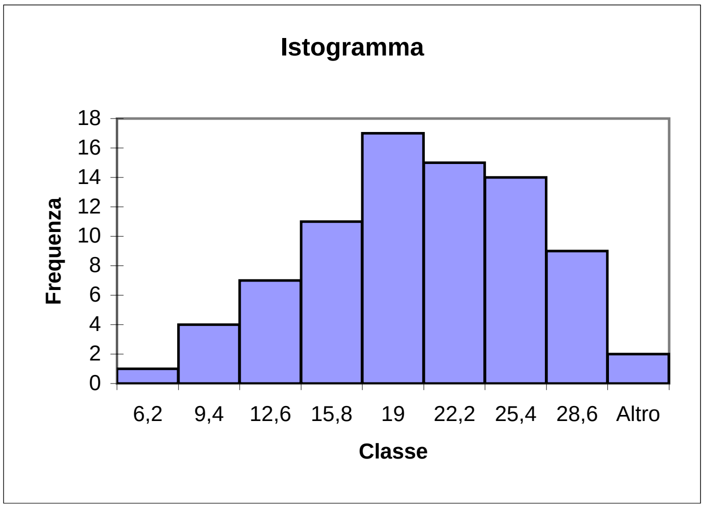
Strumenti Analisi: Istogramma

Anche lo strumento istogramma richiede che i dati siano disposti in colonna
Scegliere Menu Strumenti > Analisi dati > Istogramma
Riempire la finestra di dialogo secondo le indicazioni seguenti (figura 2)
Premere il tasto Tab per passare al campo successivo nella finestra di dialogo
Intervallo di input: selezionare con il mouse le celle contenenti i dati in colonna
Intervallo della classe: se non si indica la scelta, Excel utilizza un numero di classi aprrossimativamente uguale alla radice quadrata del numero di dati, con classi di uguale ampiezza; l'estremo sinistro della prima classe è uguale al più piccolo dei dati.
Nelle **opzioni di output** (Intervallo di output) indicare l'indirizzo della prima cella in alto a sinistra in cui si vuole far comparire l'output, oppure scegliere Nuovo foglio di lavoro per disporre l'output in un nuovo foglio.
Percentuale cumulativa: attivare questa scelta per ottenere le frequenze cumulative
Grafico in output: attivare questa scelta per ottenere l'istogramma delle frequenze assolute.
Dopo aver riempito i campi necessari, cliccare su OK
Il risultato è una tabella contenente la distribuzione di frequenza e l'istogramma
Nella prima colonna della tabella (Classe) compaiono gli estremi destri delle **classi scelte in modo automatico**; l'ultima classe denominata Altro contiene gli eventuali dati maggiori dell' estremo destro dell'ultima classe, che non sono quindi conteggiati nella classe precedente
Le etichette di classe sono di default gli estremi destri delle classi
L'istogramma è in effetti un diagramma a barre e occorre ridurre la distanza fra le barre per ottenere il vero e proprio istogramma; le etichette di classe sono di default gli estremi destri delle classi
Si possono aggiungere titoli, modificare colori, ecc.

Figura 2



Classe	*Frequenza*
6,2	1
9,4	4
12,6	7
15,8	11
19	17
22,2	15
25,4	14
28,6	9
Altro	2



SUGGERIMENTI

Scelta personalizzata delle classi

Se la scelta automatica delle classi non è soddisfacente, le **classi** possono anche essere **scelte dall'utente**; occorre predisporre una tabella con gli estremi destri delle classi e nella finestra di dialogo (figura 3) completare l'input con l'Intervallo di classe, contenente gli estremi destri scelti.

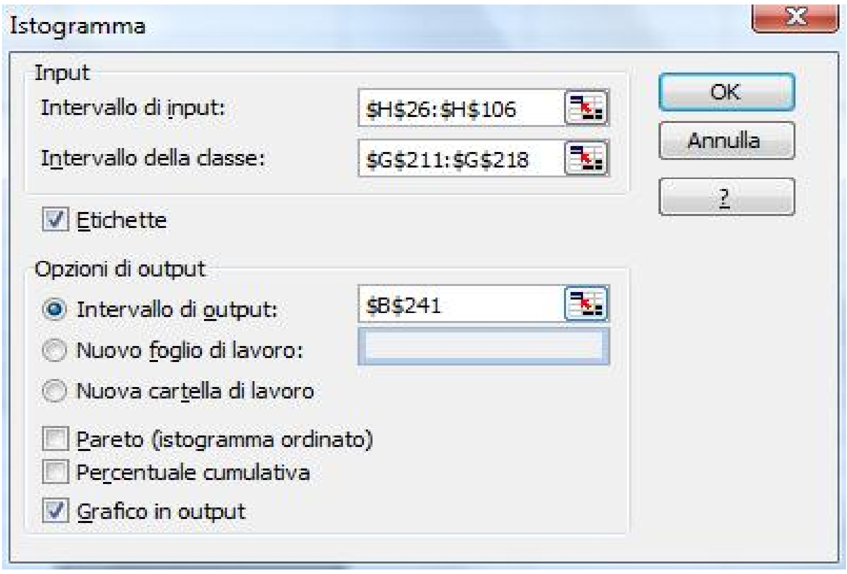
Per la scelta delle classi: numero, ampiezza, estremi destri, vedi esercizio 23

Nell'esercizio 23 sono state scelte sette classi di ampiezza = 4, i cui estremi destri sono riportati qui a destra

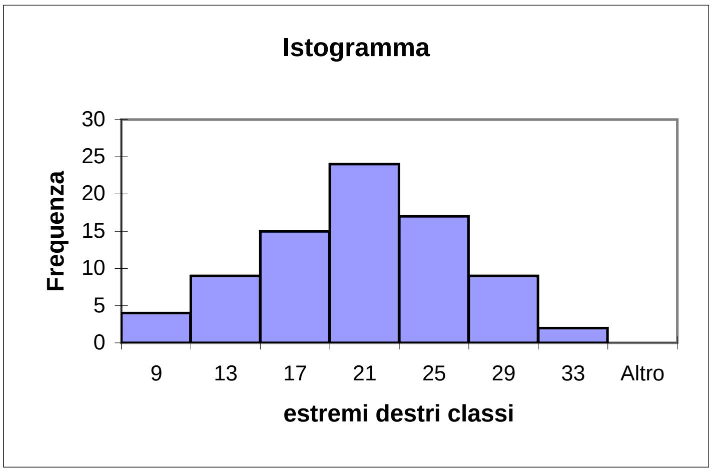
estremi destri classi
9
13
17
21
25
29
33

Attenzione! Excel aggiunge una classe in più denominata Altro; se le classi vengono scelte dall'utente in modo che l'estremo destro dell'ultima classe comprenda alla sua sinistra tutti i dati, alla classe denominata Altro corrisponderà la frequenza 0

Figura 3



estremi destri classi	*Frequenza*
9	4
13	9
17	15
21	24
25	17
29	9
33	2
Altro	0



Esercizio 30.2

La tabella 2 contiene le misure dei diametri di 100 sferette in cm. (Esercizio 26)
Effettuare l'analisi dei dati con gli strumenti Statistica Descrittiva e Istogramma
Ricordare che i dati devono essere disposti in colonna
Scegliere prima le classi in modo automatico e poi in modo personalizzato

Tabella 2

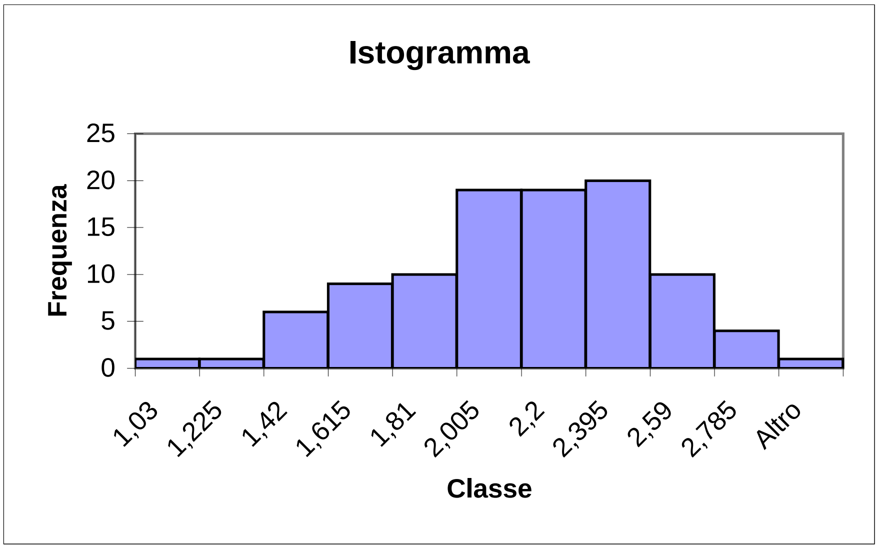
1,81	2,13	1,28	2	2,28
1,26	1,4	2,12	1,86	2,36
2,06	2,32	1,53	2,49	2,42
2,13	2,73	2,63	1,16	1,56
1,49	1,69	1,64	2,19	2,09
2,53	2,38	2,24	2,4	2,11
2,53	2,56	2,1	2,33	1,55
1,98	1,29	1,59	2,26	1,67
2,15	1,36	1,03	2,02	2,48
2,08	2,26	1,97	2,3	1,94
1,92	1,82	1,55	2,25	2,17
2,32	2,31	2,27	1,89	2,04
1,74	2,36	2,23	1,83	1,72
2,98	2,32	2,76	1,87	1,75
1,94	2,58	2,26	1,34	2,2

Dati
1,81
1,26
2,06
2,13
1,49
2,53
2,53
1,98
2,15
2,08
1,92
2,32
1,74
2,98
1,94

2,05	2,3	1,71	1,9	1,58
2,48	2,53	2,17	2,05	2,35
2,03	1,46	1,55	2,14	2,25
1,96	1,99	1,99	2,65	1,63
1,63	1,93	1,98	1,84	1,88

Dati	
Media	2,022
Errore standard	0,039
Mediana	2,05
Moda	2,53
Deviazione standard	0,388
Varianza campionaria	0,151
Curtosi	-0,242
Asimmetria	-0,250
Intervallo	1,95
Minimo	1,03
Massimo	2,98
Somma	202,19
Conteggio	100

Classe	*Frequenza*
1,03	1
1,225	1
1,42	6
1,615	9
1,81	10
2,005	19
2,2	19
2,395	20
2,59	10
2,785	4
Altro	1

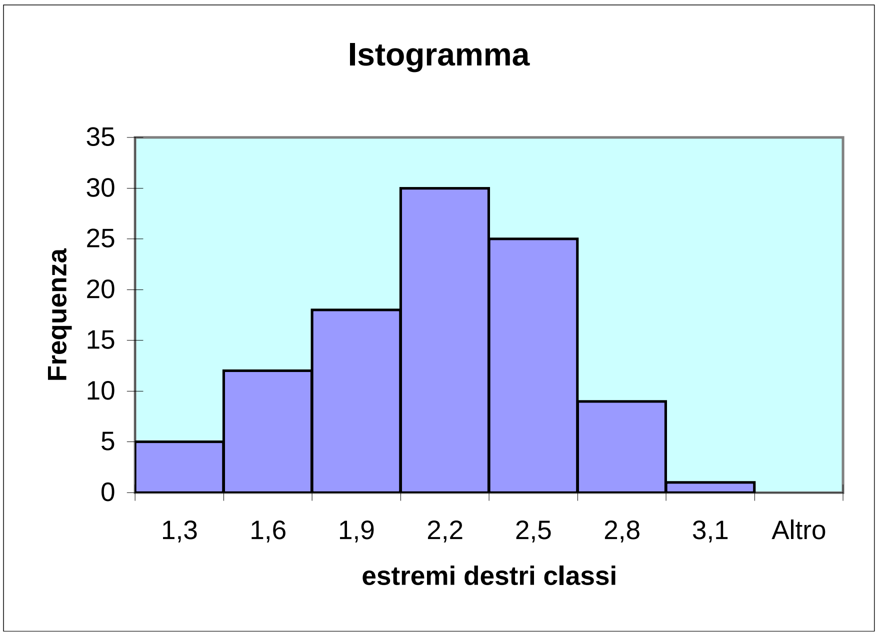


2,05
2,48
2,03
1,96
1,63
2,13
1,4
2,32
2,73
1,69
2,38
2,56
1,29
1,36
2,26
1,82
2,31
2,36
2,32
2,58
2,3
2,53
1,46
1,99
1,93
1,28
2,12
1,53
2,63
1,64
2,24
2,1
1,59
1,03
1,97
1,55
2,27
2,23
2,76
2,26
1,71
2,17
1,55
1,99
1,98
2
1,86
2,49
1,16
2,19
2,4
2,33
2,26
2,02
2,3
2,25

La scelta automatica delle classi non è la migliore (ci sono troppe classi, vedere anche esercizio 26).
Una scelta migliore può essere fatta dall'utente: ad esempio le classi di cui si elencano gli estremi destri qui sotto (vedere esercizio 26)

estremi destri classi
1,3
1,6
1,9
2,2
2,5
2,8
3,1

estremi destri classi	*Frequenza*
1,3	5
1,6	12
1,9	18
2,2	30
2,5	25
2,8	9
3,1	1
Altro	0



1,89
1,83
1,87
1,34
1,9
2,05
2,14
2,65
1,84
2,28
2,36
2,42
1,56
2,09
2,11
1,55
1,67
2,48
1,94
2,17
2,04
1,72
1,75
2,2
1,58
2,35
2,25
1,63
1,88

Esercizio 30.3
Nella tabella 3 sono riportati 90 dati (misure ddi lunghezza in mm)
Effettuare l'analisi dei dati con gli strumenti Statistica Descrittiva e Istogramma
Ricordare che i dati devono essere disposti in colonna
Scegliere prima le classi in modo automatico e poi in modo personalizzato

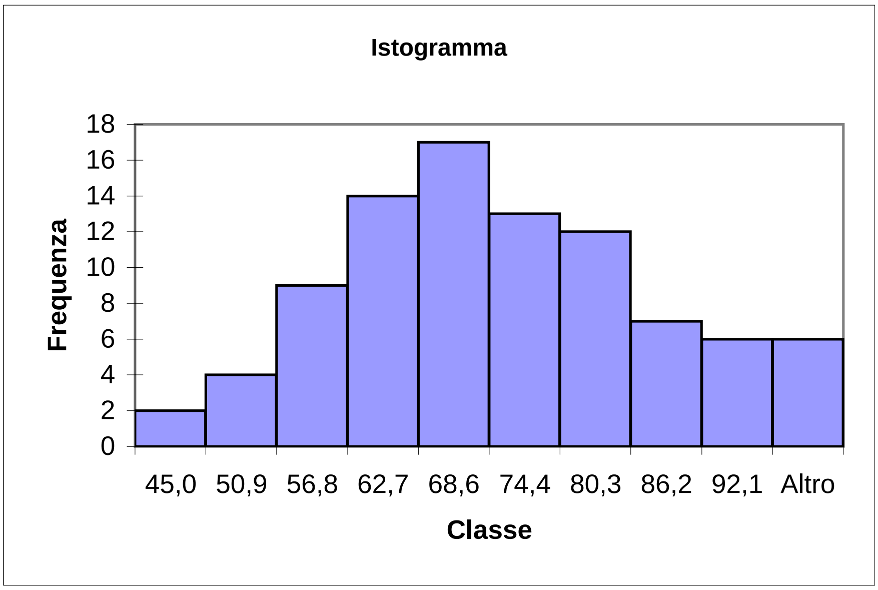
Tabella 3

76	56	77	76	63
75	67	53	67	45
73	58	63	58	53
52	61	72	46	61
60	77	92	49	58
72	83	64	64	81
47	80	95	72	88
52	61	55	72	75
72	77	68	53	62
87	86	95	84	72
64	80	67	88	57
69	71	86	63	52
93	68	90	63	68
68	57	94	78	77
59	82	74	98	69
73	62	62	67	63
83	87	56	74	93
48	59	78	45	66

Dati	
Media	69,511
Errore standard	1,391
Mediana	68
Moda	72
Deviazione standard	13,195
Varianza campionaria	174,095
Curtosi	-0,669
Asimmetria	0,201
Intervallo	53
Minimo	45
Massimo	98
Somma	6256
Conteggio	90

Classe	*Frequenza*
45,0	2
50,9	4
56,8	9
62,7	14
68,6	17
74,4	13
80,3	12
86,2	7
92,1	6
Altro	6

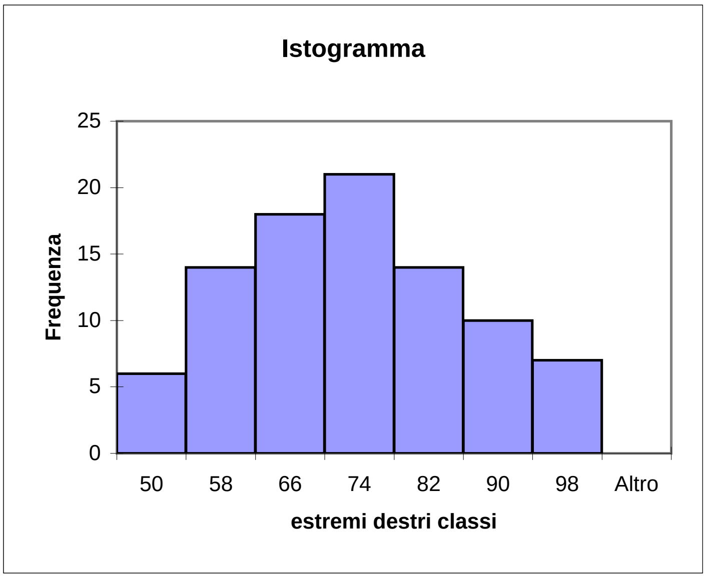
Dati
76
75
73
52
60
72
47
52
72
87
64
69
93
68
59
73
83
48
56
67
58
61
77
83
80
61
77
86
80
71
68
57
82
62
87
59
77
53
63
72
92
64
95
55
68
95
67
86
90
94
74
62
56
78
76



La scelta automatica delle classi non è la migliore (ci sono troppe classi)
Una scelta migliore può essere fatta dall'utente nel modo seguente

numero classi	7
ampiezza	8

estremi destri classi
50
58
66
74
82
90
98



estremi destri classi	*Frequenza*
50	6
58	14
66	18
74	21
82	14
90	10
98	7
Altro	0

67
58
46
49
64
72
72
53
84
88
63
63
78
98
67
74
45
63
45
53
61
58
81
88
75
62
72
57
52
68
77
69
63
93
66

Torna su

4. CORRELAZIONE E REGRESSIONE

Soluzione Esercizio 31
Calcolo di covarianza e coefficiente di correlazione lineare

Indice

Ritorna Esercizio 31

Esercizio 31.1

Nella tabella 1 sono riportati i punteggi conseguiti da dieci studenti negli esami di Analisi I e Analisi II (punteggio massimo = 100)
Calcolare il coefficiente di correlazione lineare usando la funzione di Excel e stabilire se i dati sono linearmente correlati. Usare la funzione CORRELAZIONE
I dati sono linearmente correlati? Perché?
Disegnare in un grafico i dati assegnati (punti)

SUGGERIMENTI

Per stabilire se due insiemi di dati sono **linearmente** correlati di usa la funzione CORRELAZIONE
Sintassi
CORRELAZIONE(**matrice1**;**matrice2**)
matrice1 intervallo di celle contenente i valori del primo insieme di dati
matrice2 intervallo di celle contenente i valori del secondo insieme di dati

Coefficiente di correlazione R	0,955

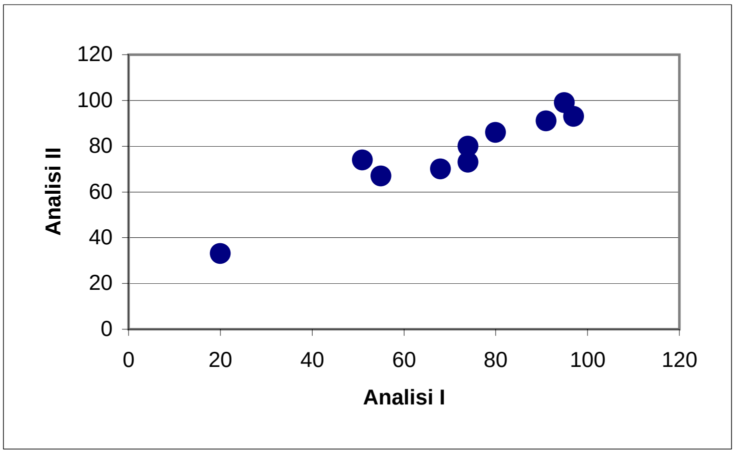


Tabella 1

Analisi I	Analisi II
51	74
68	70
97	93
55	67
95	99
74	73
20	33
91	91
74	80
80	86

Questi dati sono linearmente correlati, perché il valore del coefficiente R è prossimo a 1

Esercizio 31.2

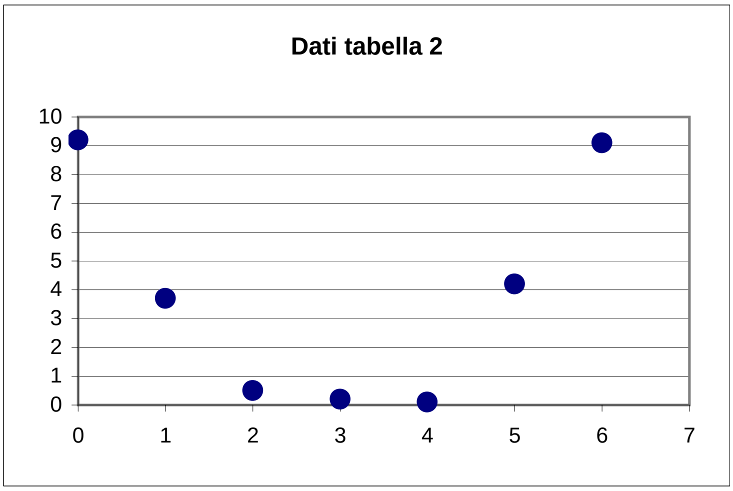
Calcolare il coefficiente di correlazione lineare per i dati della tabella 2
I dati sono linearmente correlati? Perché?
I dati possono avere altri tipi di correlazione?
Disegnare il diagramma di dispersione per stabilire se esiste un altro tipo di correlazione.

Coefficiente di correlazione R	0,0058

Questi dati non sono linearmente correlati perché il coefficiente R è prossimo a 0

Tabella 2

0	9,2
1	3,7
2	0,5
3	0,2
4	0,1
5	4,2
6	9,1

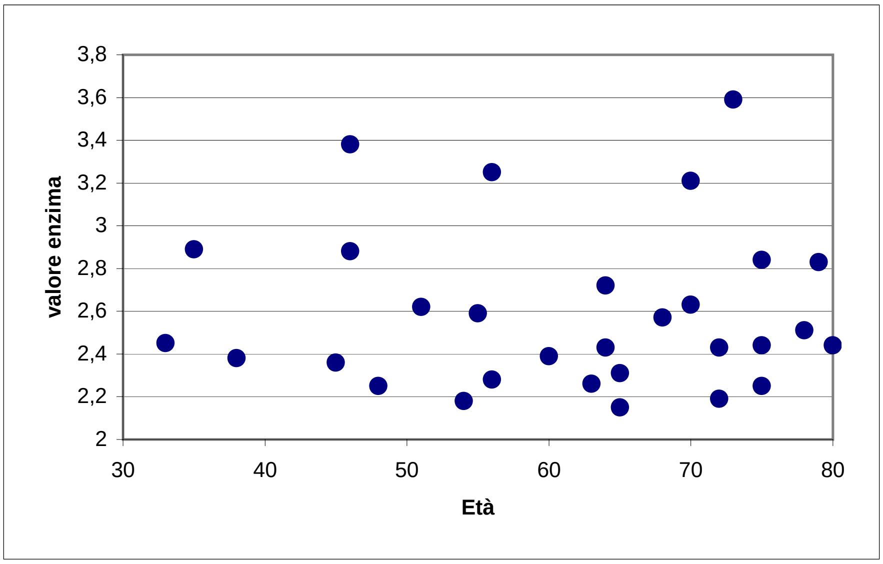


Il grafico suggerisce una correlazione di tipo polinomiale (parabola)
Il coefficiente di correlazione permette solo di stabilire l'eventuale correlazione lineare, ma non di escludere altri tipi di correlazione. Negli esercizi successivi si vedrà come trovare il polinomio di grado superiore a 1 che approssima i dati con il criterio dei minimi quadrati.

Esercizio 31.3

La tabella 3 contiene i valori di un enzima nel sangue di 30 persone di sesso maschile di età compresa fra 30 e 80 anni
Calcolare il coefficiente di correlazione lineare per i dati della tabella 3. I dati sono linearmente correlati?
Disegnare un diagramma a dispersione.

Coefficiente di correlazione R	-0,0093



Questi dati non sono linearmente correlati, perché il valore di R è prossimo a 0.
Il coefficiente di correlazione permette solo di stabilire l'eventuale correlazione lineare, ma non di escludere altri tipi di correlazione; in questo caso il grafico non suggerisce alcun tipo di legame.

Tabella 3

età	valore enzima
63	2,26
75	2,25
46	3,38
64	2,43
72	2,19
64	2,72
79	2,83
60	2,39
45	2,36
56	2,28
80	2,44
73	3,59
33	2,45
65	2,31
78	2,51
75	2,84
35	2,89
70	2,63
75	2,44
51	2,62
38	2,38
56	3,25
55	2,59
68	2,57
65	2,15
70	3,21
72	2,43
48	2,25
54	2,18
46	2,88

Esercizio 31.4

La tabella riporta il peso e l'altezza di un gruppo di 20 studenti di 18 anni

Tabella 4

Peso (kg)	Altezza (cm)	Altezza (m)
72	174	1,74
63	168	1,68
78	183	1,83
60	160	1,6
58	164	1,64
75	170	1,7
80	179	1,79
77	178	1,78
65	170	1,7
69	170	1,7
72	175	1,75
65	170	1,7
80	185	1,85
57	154	1,54
60	165	1,65
77	175	1,75
83	182	1,82
79	178	1,78
67	175	1,75
68	173	1,73

Calcolare la covarianza e il coefficiente di correlazione R.

SUGGERIMENTI

Per il calcolo della covarianza si usa la funzione COVARIANZA

Sintassi

COVARIANZA(**matrice1**; **matrice2**)

matrice1 primo intervallo di celle
matrice2 secondo intervallo di celle

Excel calcola la covarianza dividendo per il numero n dei dati, anziché per n–1
come sarebbe corretto, trattandosi di un campione di dati e non di una popolazione.
Per la formula vedere Guida in linea della funzione
Il valore della covarianza dipende dall'unità di misura, mentre il coefficiente di correlazione
è un numero puro, e il suo valore non cambia cambiando l'unità di misura dei dati.
Questo fatto rende la **covarianza** un **parametro difficile da interpretare**, e si preferisce
usare il coefficiente di correlazione.

Modificare i dati relativi all'altezza trasformandoli in metri e calcolare di nuovo covarianza
e coefficiente di correlazione: quest'ultimo non cambia mentre la covarianza cambia.

covarianza (dati in cm)	54,6
covarianza (dati in m)	0,546
coefficiente di correlazione R (dati in cm)	0,9026
coefficiente di correlazione R (dati in m)	0,9026

Torna su

Soluzione Esercizio 32

Calcolo del coefficiente di correlazione; grafico retta di regressione

Indice

Ritorna Esercizio 32

Esercizio 32.1

La tabella 1 riporta le misure del volume di una quantità di un gas a differenti temperature.

Tabella 1

Temperatura	Volume
10	10,4
20	11,1
30	11,2
40	11,9
50	11,8
60	12,3

Calcolare il coefficiente di correlazione e verificare se esiste una dipendenza lineare del volume dalla temperatura.
(Usare la funzione CORRELAZIONE)
Determinare l'equazione della retta di regressione y=Ax+B e disegnarne il grafico

Il calcolo dei coefficienti della retta di regressione può essere fatto con le funzioni PENDENZA (coefficiente A) e INTERCETTA (coefficiente B)

Sintassi

PENDENZA(**y_nota**;**x_nota**)

INTERCETTA(**y_nota**;**x_nota**)

x_nota insieme dei valori della variabile indipendente X

y_nota insieme dei valori della variabile dipendente Y

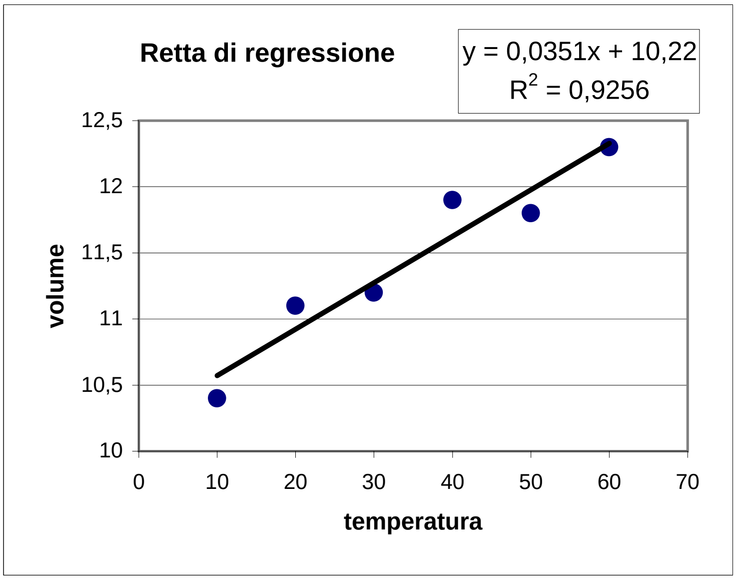
SUGGERIMENTI

Attenzione alla sintassi delle funzioni PENDENZA e INTERCETTA:
x_nota sono i valori di Temperatura; y_nota sono i valori di Volume;
il primo argomento è y_nota, il secondo è x_nota
Per tracciare il grafico della retta di regressione:
Selezionare le celle C13:D18 e tracciare il grafico a dispersione (punti)
Dopo aver terminato il grafico, puntare con il mouse su uno dei punti e premere il tasto destro
Nel menu di scelta rapida che si apre, selezionare **Aggiungi linea di tendenza**
Nella finestra **Tipo** scegliere **Lineare**
Se si vuole anche visualizzare l'equazione della linea, nella finestra Aggiungi linea di tendenza aprire la scheda **Opzioni** e selezionare **Visualizza l'equazione sul grafico**;
selezionando nella stessa scheda **Visualizza il valore R al quadrato sul grafico**, si ottiene anche il valore del coefficiente di correlazione al quadrato
La casella contenente l'equazione e/o il coefficiente di regressione può essere spostata con il mouse in qualunque punto dell'Area del grafico (dentro la figura) con il mouse: selezionare la casella, premere il tasto sinistro e spostare.

Coefficiente di correlazione R	0,962

Il valore del coefficiente di correlazione prossimo a 1 indica la dipendenza di tipo lineare

Equazione retta	y=Ax+B
A	0,0351
B	10,22



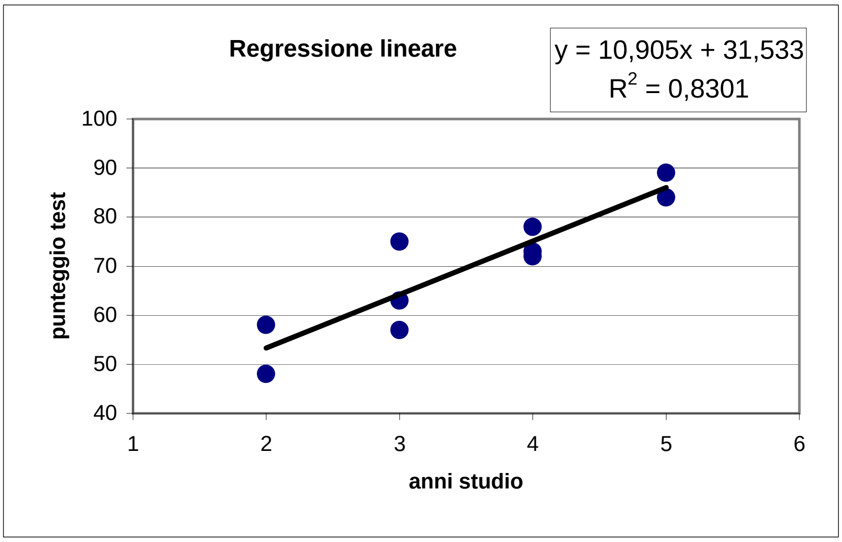
Esercizio 32.2

La tabella 2 mostra la relazione fra il numero di anni di studio di una lingua straniera e il punteggio conseguito in un test di conoscenza della lingua.
Determinare l'equazione della retta di regressione per i dati della tabella usando le funzioni PENDENZA e INTERCETTA
Disegnare il grafico usando Aggiungi linea di tendenza

Tabella 2

N° anni studio	Punteggio
3	57
4	78
4	72
2	58
5	89
3	63
4	73
5	84
3	75
2	48

Equazione retta	y=Ax+B
A	10,905
B	31,533



Torna su

Soluzione Esercizio 33

Retta di regressione: grafico, barre di errore

Indice

Ritorna Esercizio 33

Esercizio 33.1

La tabella 1 riporta 10 misure di velocità in funzione del tempo, e i rispettivi errori di misurazione.
Determinare l'equazione della retta di regressione per i dati della tabella usando le funzioni PENDENZA e INTERCETTA
Disegnare il grafico usando Aggiungi linea di tendenza
Aggiungere le barre di errore per la velocità e per il tempo

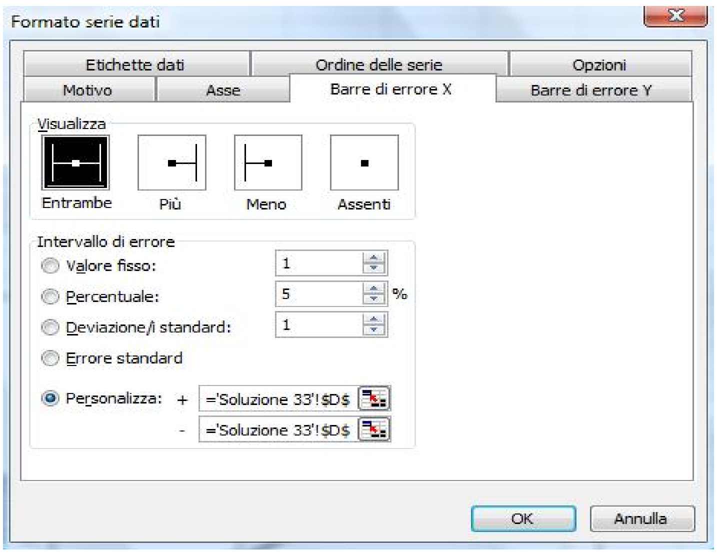
Tabella 1

tempo (s)	velocità (m/s)	errore tempo (s)	errore velocità (m/s)
1	11	0,2	1
2	13	0,2	0,8
3	15	0,2	0,6
4	18	0,2	2
5	22	0,2	2,1
6	23	0,2	2
7	22	0,2	1,5
8	27	0,2	1
9	28	0,2	2,5
10	31	0,2	2

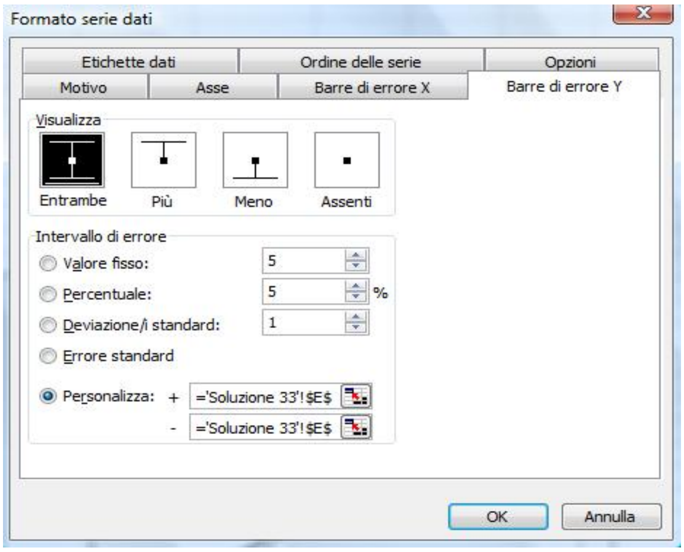
Equazione retta	y=Ax+B
A	2,1697
B	9,0667

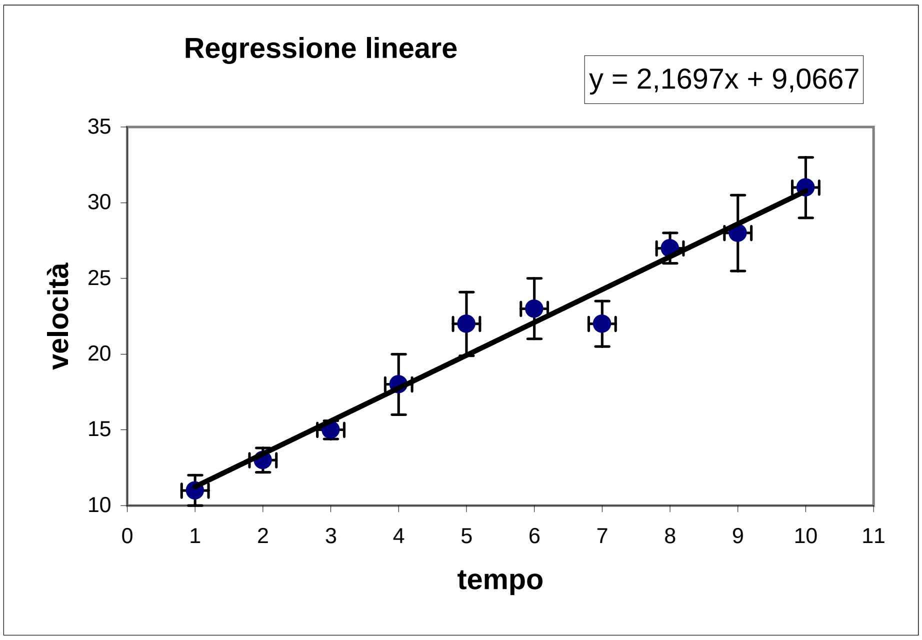
SUGGERIMENTI

Per aggiungere le barre di errore procedere nel modo seguente:
puntare con il mouse su uno dei punti del diagramma a dispersione, premere il tasto destro e scegliere Formato serie dati; nella finestra che si apre operare sulla scheda Barre di errore X
Scegliere **Visualizza entrambe**
Selezionare **Personalizza**: nelle caselle a destra dei segni + e – selezionare con il mouse le celle D16:D25 contenenti gli errori sul tempo



Ripetere il procedimento in modo analogo sulla scheda Barre di errore Y (**Personalizza**: selezionare le celle egli errori sulla velocità)





Torna su

Soluzione Esercizio 34
Retta di regressione: funzioni

Indice

Ritorna Esercizio 34

Esercizio 34.1

La tabella 1 mostra l'età e il valore della pressione sanguigna di un gruppo di 13 persone.
Determinare l'equazione della retta di regressione per i dati della tabella usando le funzioni PENDENZA e INTERCETTA
Disegnare il grafico della retta di regressione (Aggiungi linea di tendenza); visualizzare sul grafico l'equazione della retta di regressione
Stimare il valore della pressione per una persona di 52 anni con la funzione PREVISIONE
Aggiungere sul grafico il punto di ascissa x=52 e ordinata y=valore calcolato con PREVISIONE

Tabella 1

Età	Pressione
55	145
42	125
71	155
36	115
63	145
48	130
50	150
55	145
49	150
38	110
44	140
65	175
69	170

SUGGERIMENTI

La funzione PREVISIONE calcola con il metodo dei minimi quadrati i coefficienti della retta di regressione e fornisce il valore previsto di y per il valore specificato di x.

Sintassi

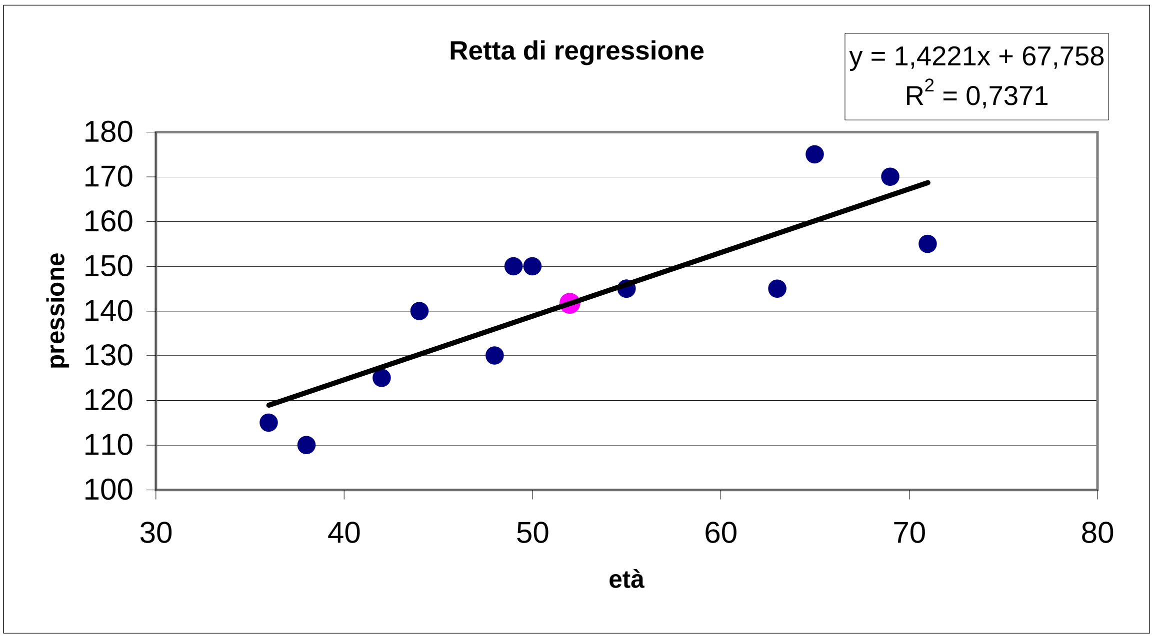
PREVISIONE(x;y_nota;x_nota)

x	valore nel quale si vuole approssimare.
x_nota	insieme dei valori della variabile indipendente X
y_nota	insieme dei valori della variabile dipendente Y

Equazione retta	y=Ax+B
A	1,422
B	67,76

Anni	52
Pressione stimata	142

Per aggiungere il punto sul grafico, dopo aver tracciato il grafico della retta, aggiungere una nuova serie di dati:
Valori X: selezionare la cella E46; Valori Y: selezionare la cella E47



Esercizio 34.2

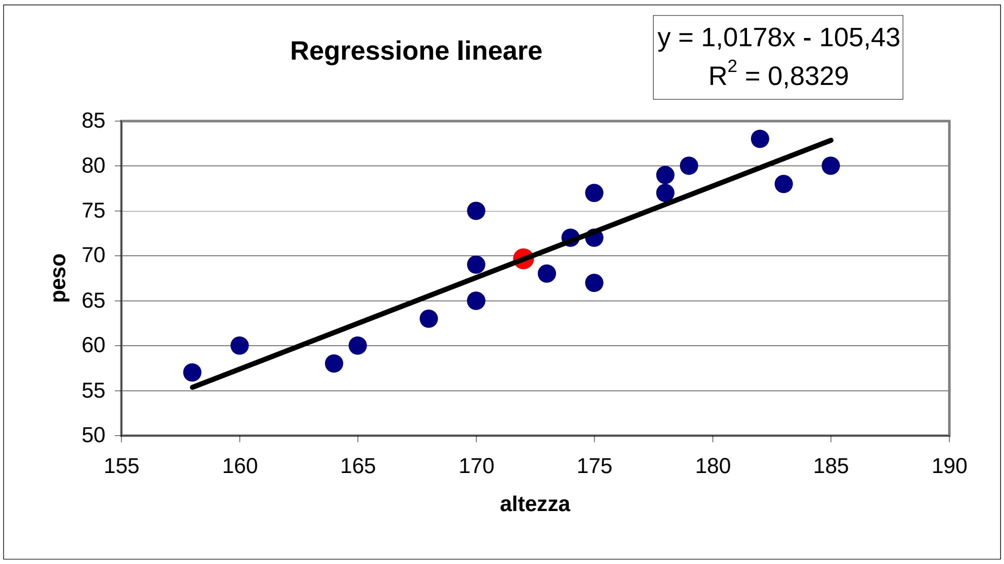
La tabella 2 riporta l'altezza e il peso di un gruppo di 20 studenti di 18 anni
Determinare l'equazione della retta di regressione per i dati della tabella usando le funzioni PENDENZA e INTERCETTA
Disegnare il grafico della retta di regressione (Aggiungi linea di tendenza); visualizzare sul grafico l'equazione della retta di regressione
Stimare il valore del peso di uno studente alto 172 cm con la funzione PREVISIONE

Tabella 2

Altezza (cm)	Peso (kg)
174	72
168	63
183	78
160	60
164	58
170	75
179	80
178	77
170	65
170	69
175	72
170	65
185	80
158	57
165	60
175	77
182	83
178	79
175	67
173	68

Equazione retta	y=Ax+B
A	1,0178
B	-105,43

Altezza cm	172
Peso stimato kg	70



Esercizio 34.3

Sono assegnati i dati della tabella 3, che descrivono l'andamento delle vendite in funzione dell'aumento del prezzo di un bene

Calcolare il coefficiente di correlazione lineare: c'è correlazione lineare?

Trovare i coefficienti della retta di regressione usando le funzioni di Excel

Disegnare in un grafico i dati assegnati e la retta di regressione

Approssimare il valore della quantità venduta y quando il prezzo è x=40

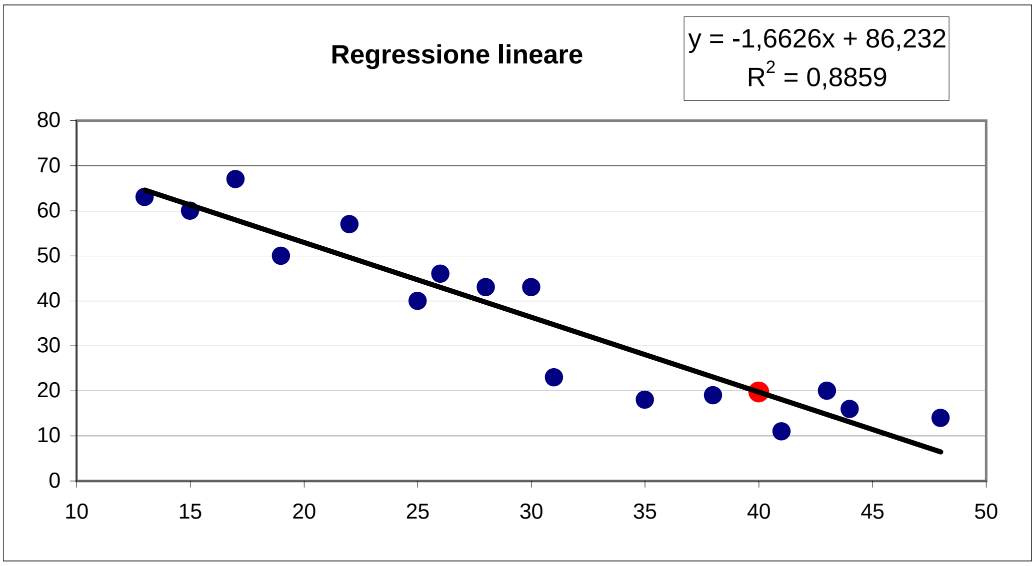
Tabella 3

x	y
13	63
15	60
17	67
19	50
22	57
25	40
26	46
28	43
30	43
31	23
35	18
38	19
41	11
43	20
44	16
48	14

Coeff. correlazione R	-0,941

Equazione retta y=Ax+B	
A	-1,6626
B	86,232

Prezzo	40
Stima quantità venduta	20


Regressione lineare
y = -1,6626x + 86,232
R2 = 0,8859
80
70
60
50
40
30
20
10
0
10
15
20
25
30
35
40
45
50

Torna su

Soluzione Esercizio 35
Serie temporali: grafici e regressione lineare

Indice

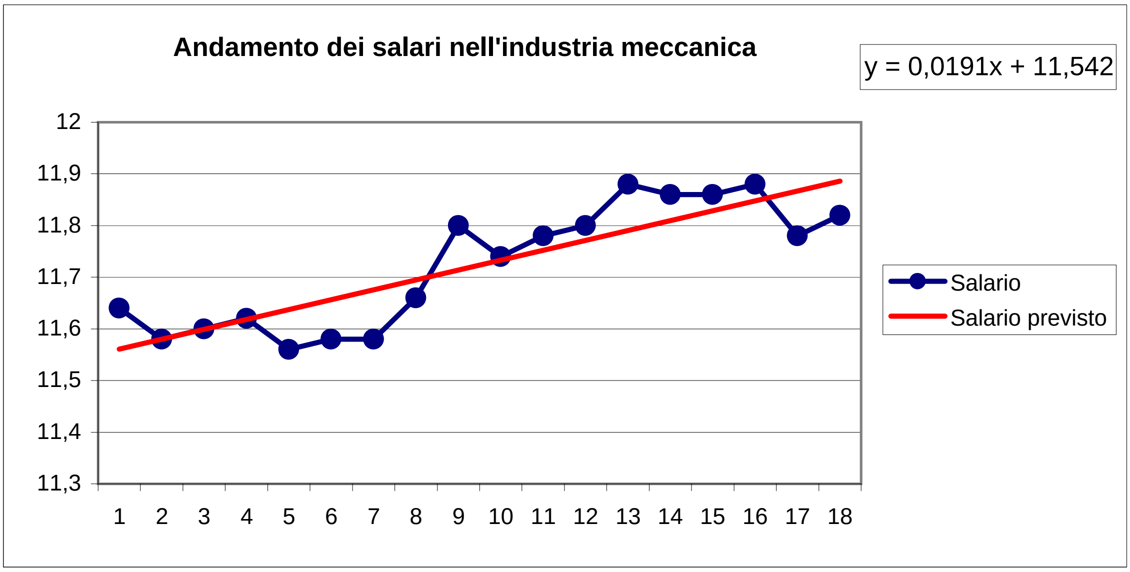
Ritorna Esercizio 35

Esercizio 35.1
Nella tabella 1 sono riportati i dati relativi ai salari orari medi degli operai specializzati dell'industria meccanica per 18 mesi, da gennaio 2003 a giugno 2004
Disegnare un grafico della sequenza temporale dei dati
Aggiungere una linea di tendenza lineare, visualizzando l'equazione sul grafico.

SUGGERIMENTI
Per disegnare un grafico della sequenza temporale dei dati selezionare la colonna Salario cliccare sul pulsante Creazione guidata grafico, scegliere Linee e tracciare il grafico (linea spezzata con indicatori dei valori)

Tabella 1

Mesi	Salario
1	11,64
2	11,58
3	11,6
4	11,62
5	11,56
6	11,58
7	11,58
8	11,66
9	11,8
10	11,74
11	11,78
12	11,8
13	11,88
14	11,86
15	11,86
16	11,88
17	11,78
18	11,82



Esercizio 35.2

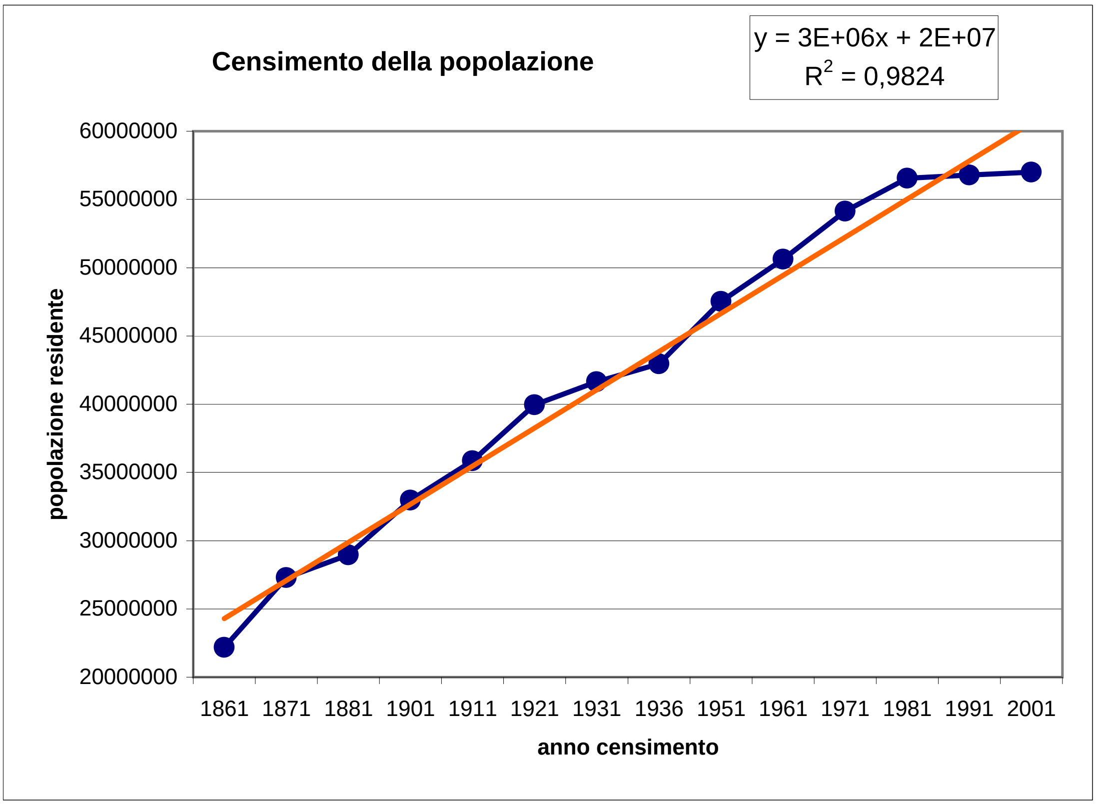
Censimento della popolazione

Nella tabella 2 si riporta la popolazione residente dell'Italia ai confini attuali ai censimenti dal 1861 al 2001 (Fonte: ISTAT Istituto Nazionale di Statistica)

Disegnare il grafico (Tipo di grafico: Linee) e aggiungere la linea di tendenza

Tabella 2

CENSIMENTI	POPOLAZIONE RESIDENTE
1861	22182377
1871	27303509
1881	28953480
1901	32965504
1911	35845048
1921	39943528
1931	41651000
1936	42943602
1951	47515537
1961	50623569
1971	54136547
1981	56556911
1991	56778031
2001	56995744



Torna su

Soluzione Esercizio 36
Regressione polinomiale

Indice

Ritorna Esercizio 36

Esercizio 36.1

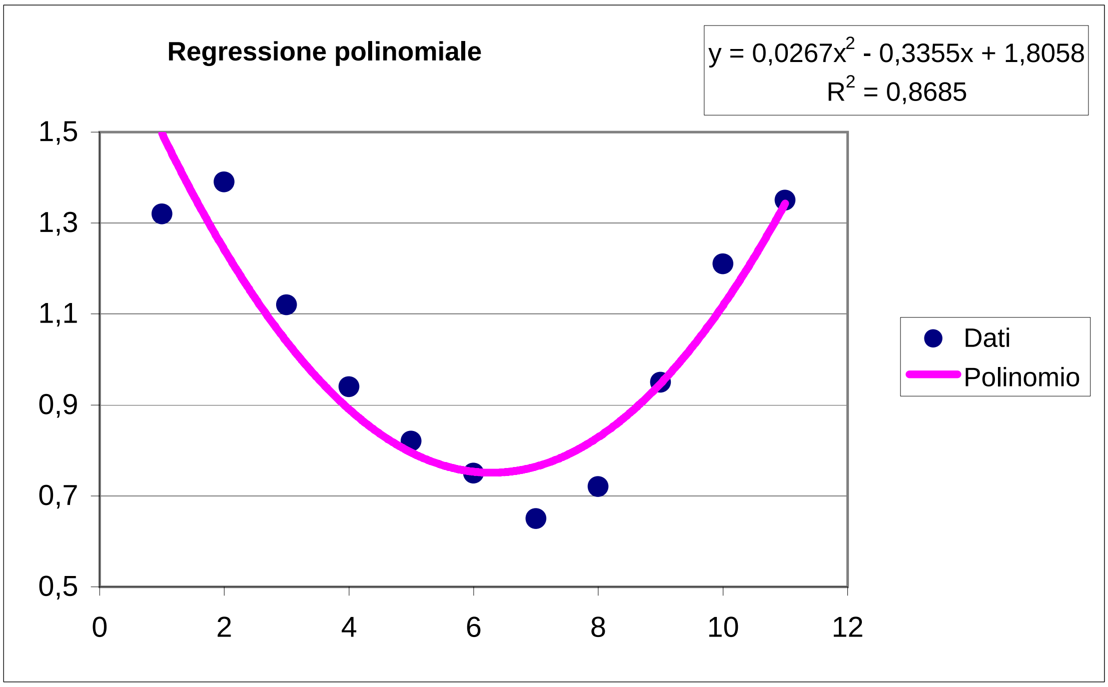
Trovare la parabola che approssima i dati della tabella 1 con il criterio dei minimi quadrati
Usare Aggiungi linea di tendenza, Tipo polinomiale, ordine 2, e far comparire l'equazione della parabola sul grafico

Tabella 1

x	y
1	1,32
2	1,39
3	1,12
4	0,94
5	0,82
6	0,75
7	0,65
8	0,72
9	0,95
10	1,21
11	1,35

SUGGERIMENTI

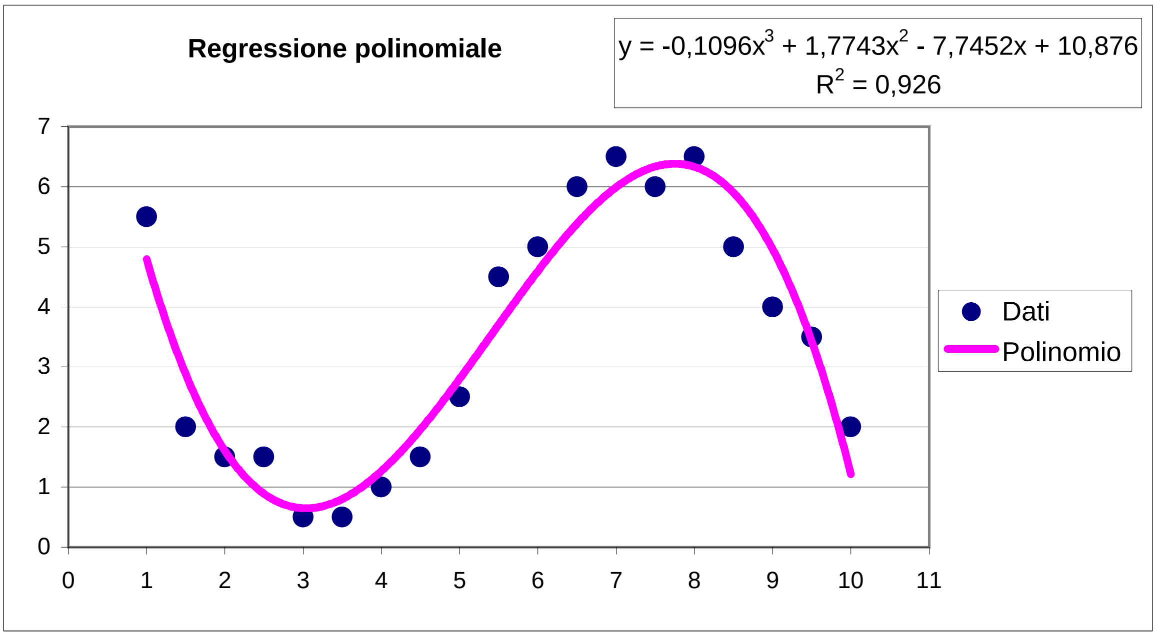
Per copiare l'equazione e incollata in altra zona del foglio di lavoro, cliccare una volta sulla casella dell'equazione, selezionare l'equazione e copiarla, poi incollarla in una cella a scelta nel foglio di lavoro



Esercizio 36.2
Trovare il polinomio di terzo grado che approssima i dati della tabella 2 con il criterio dei minimi quadrati
Usare Aggiungi linea di tendenza, Tipo polinomiale, ordine 3, e far comparire l'equazione del polinomio sul grafico

Tabella 2

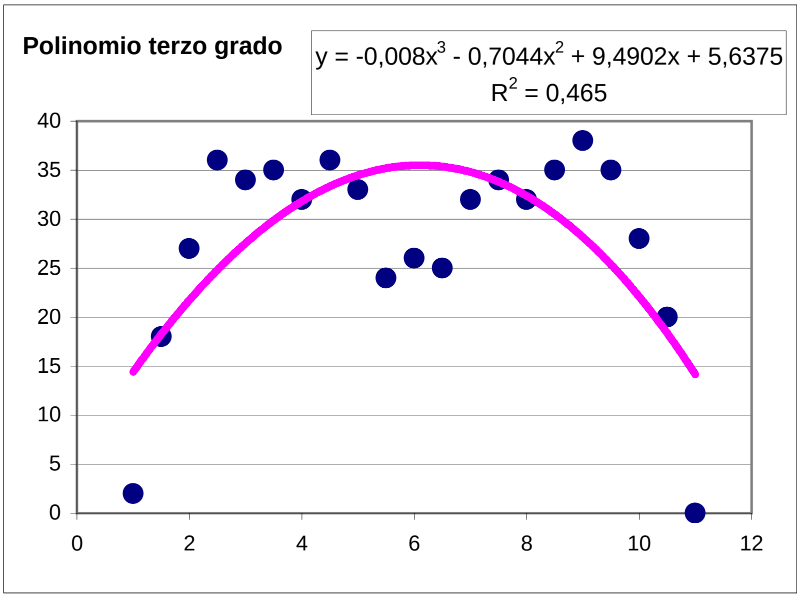
x	y
1	5,5
1,5	2
2	1,5
2,5	1,5
3	0,5
3,5	0,5
4	1
4,5	1,5
5	2,5
5,5	4,5
6	5
6,5	6
7	6,5
7,5	6
8	6,5
8,5	5
9	4
9,5	3,5
10	2

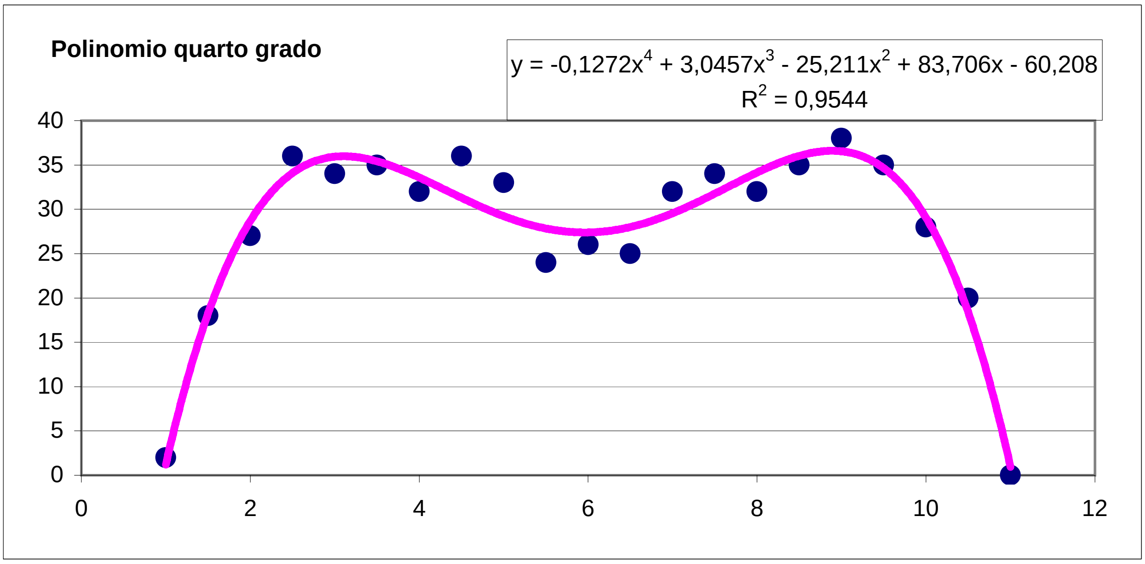


Esercizio 36.3
Trovare i polinomi di terzo e di quarto grado che approssimano i dati della tabella 3 con il criterio dei minimi quadrati
Usare Aggiungi linea di tendenza, Tipo polinomiale, ordine 3, e far comparire l'equazione del polinomio sul grafico; ripetere con ordine 4
Scegliere il polinomio che fornsce l'approssimazione migliore.

Tabella 3

x	y
1	2
1,5	18
2	27
2,5	36
3	34
3,5	35
4	32
4,5	36
5	33
5,5	24
6	26
6,5	25
7	32
7,5	34
8	32
8,5	35
9	38
9,5	35
10	28
10,5	20
11	0





L'approssimazione migliore è data dal polinomio di quarto grado, per il quale il coefficiente R^2 è più vicino a 1

Torna su

Soluzione Esercizio 37
Metodi di linearizzazione

Indice

Ritorna Esercizio 37

Esercizio 37.1
Linea di tendenza: tipo Potenza

Trovare la curva del tipo y=C*x^A che approssima i dati della tabella 1
Disegnare un grafico a dispersione e Aggiungere linea di tendenza, tipo Potenza
Visualizzare l'equazione della curva sul grafico

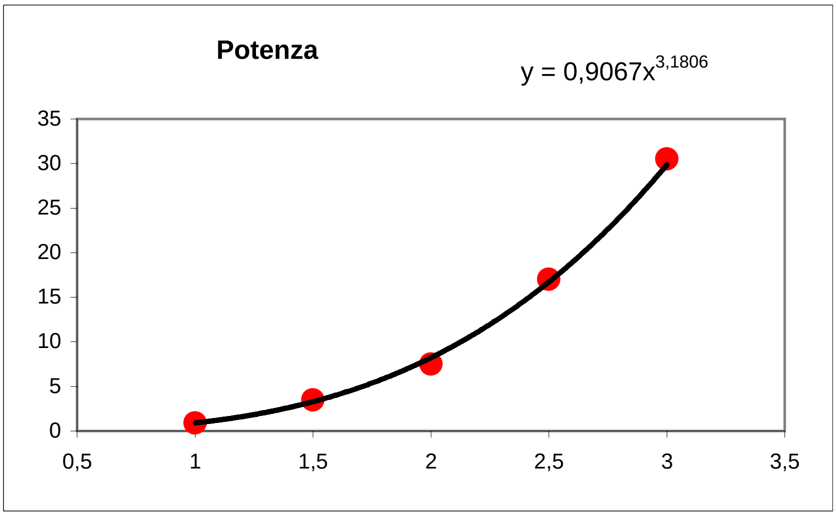


Tabella 1

x	y
1	0,9
1,5	3,5
2	7,5
2,5	17
3	30,5

Esercizio 37.2
Linea di tendenza: tipo Esponenziale

Trovare la curva del tipo y=C*exp(Ax) che approssima i dati della tabella 2
Disegnare un grafico a dispersione e Aggiungere linea di tendenza, tipo Esponenziale
Visualizzare l'equazione della curva sul grafico

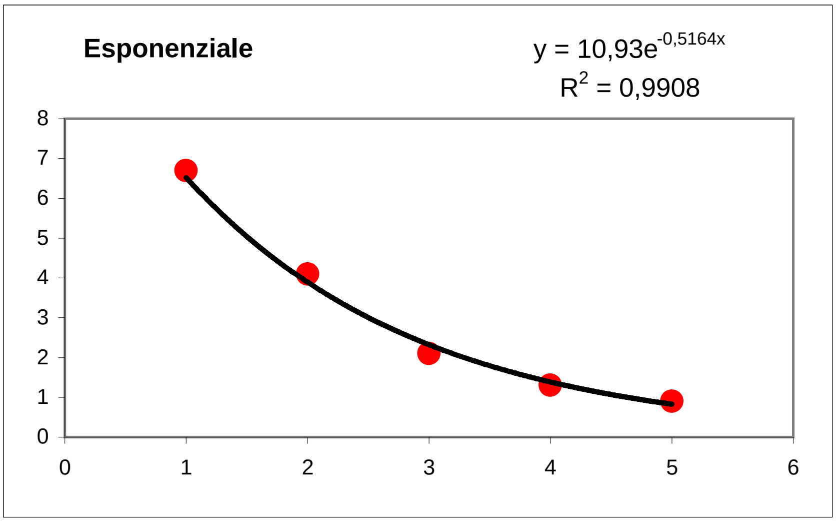


Tabella 2

x	y
1	6,7
2	4,1
3	2,1
4	1,3
5	0,9

Esercizio 37.3

Linea di tendenza: tipo Esponenziale

Il numero y di batteri per unità di volume presenti in una coltura dopo x ore è dato dalla seguente tabella 3.
Trovare la curva del tipo y=C*exp(Ax) che approssima i dati della tabella
Disegnare un grafico a dispersione e Aggiungere linea di tendenza, tipo Esponenziale
Visualizzare l'equazione della curva sul grafico

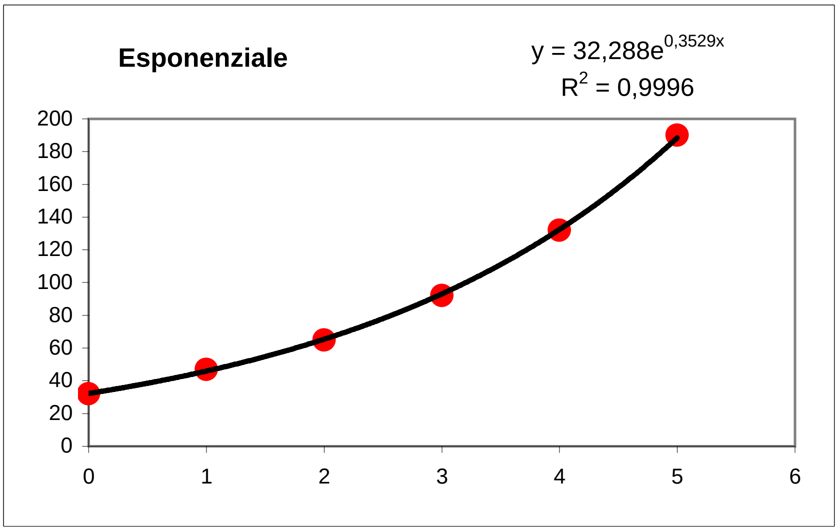


Tabella 3

x	y
0	32
1	47
2	65
3	92
4	132
5	190

Esercizio 37.4

Linea di tendenza: tipo Logaritmo

Trovare la curva del tipo y=Alnx+B che approssima i dati della tabella 4
Disegnare un grafico a dispersione e Aggiungere linea di tendenza, tipo Logaritmo
Visualizzare l'equazione della curva sul grafico

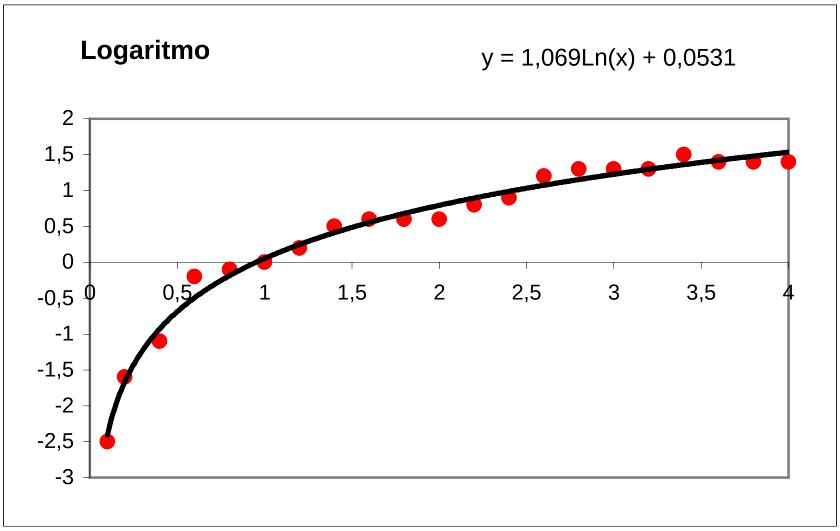


Tabella 4

x	y
0,1	-2,5
0,2	-1,6
0,4	-1,1
0,6	-0,2
0,8	-0,1
1	0
1,2	0,2
1,4	0,5
1,6	0,6
1,8	0,6
2	0,6
2,2	0,8
2,4	0,9
2,6	1,2
2,8	1,3
3	1,3
3,2	1,3
3,4	1,5
3,6	1,4
3,8	1,4
4	1,4

Soluzione Esercizio 38
Confronto fra linee di tendenza

Indice

Ritorna Esercizio 38

Esercizio 38.1

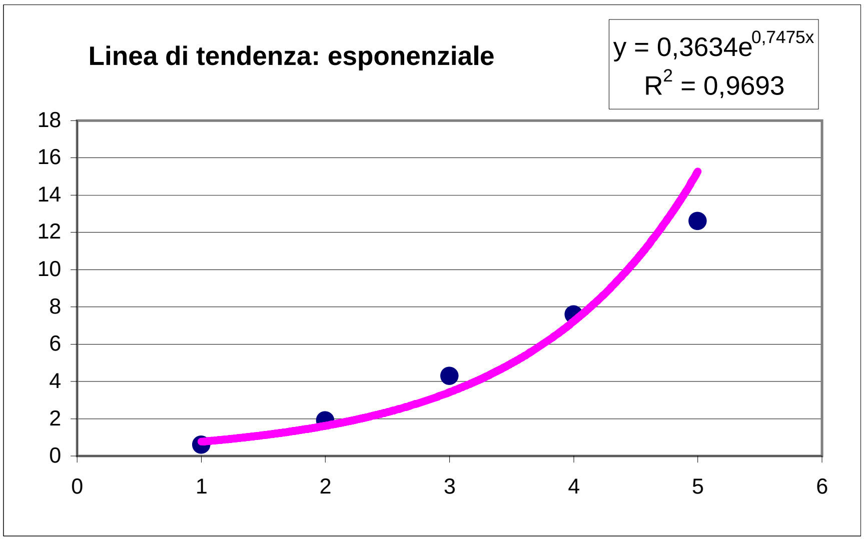
Sono assegnati i dati della tabella 1
Trovare la curva di tipo esponenziale che approssima i dati.
Trovare la curva di tipo potenza che approssima i dati.
Realizzare due grafici; con Aggiungi linea di tendenza
aggiungere la linea del tipo richiesto
Stabilire qual è la curva che approssima meglio i dati.

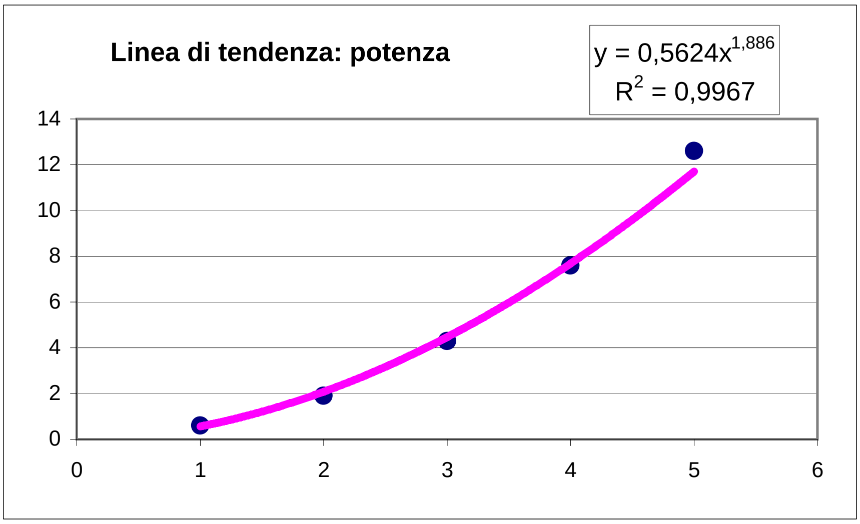
Tabella 1

x	y
1	0,6
2	1,9
3	4,3
4	7,6
5	12,6

SUGGERIMENTI

Per stabilire qual è la curva che approssima meglio i dati,
usando Aggiungi linea di tendenza, nella scheda Opzioni selezionare:
Visualizza equazione sul grafico
Visualizza il valore R al quadrato sul grafico.
Il valore visualizzato è il quadrato del coefficiente di correlazione R:
la linea che approssima meglio i dati è quella per cui il valore di R al quadrato è più vicino a 1



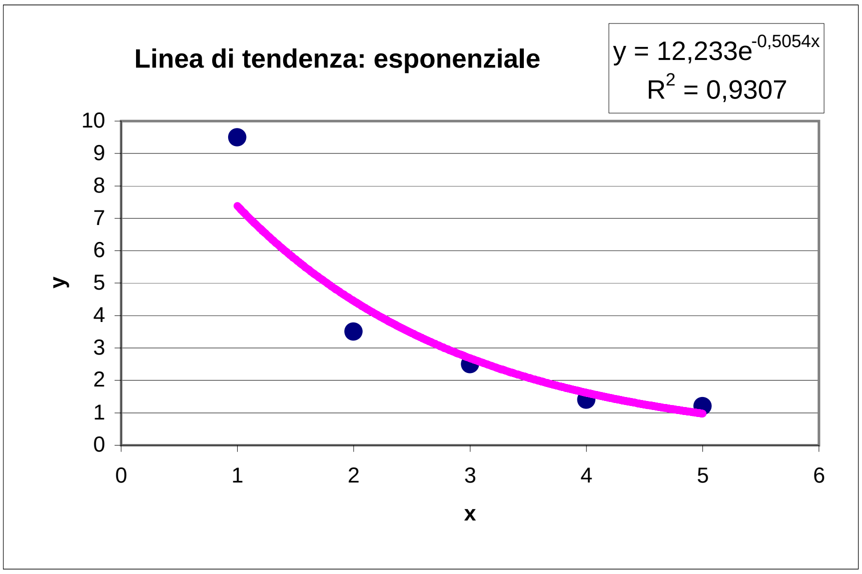


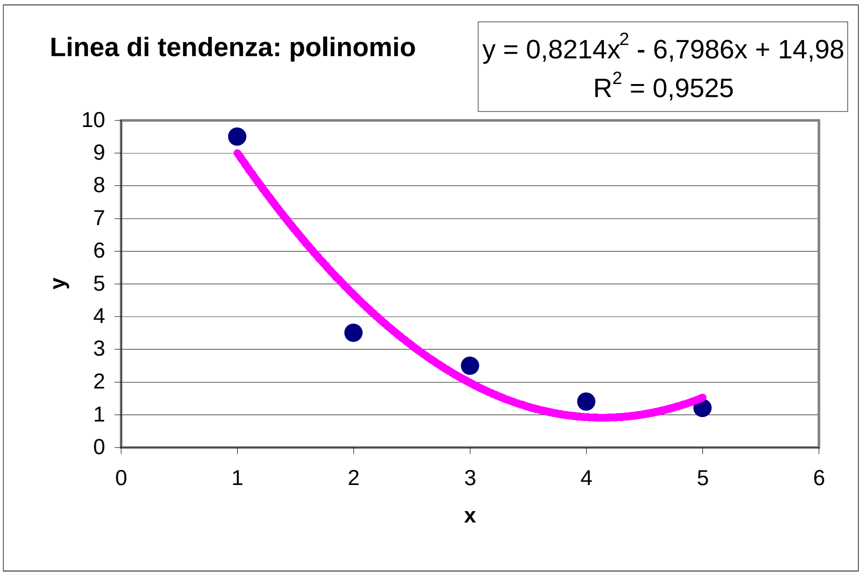
Esercizio 38.2

Sono assegnati i dati della tabella 2
Trovare la curva di tipo esponenziale che approssima i dati.
Trovare la curva d tipo polinomiale (grado 2) che approssima i dati.
Realizzare due grafici; con Aggiungi linea di tendenza
aggiungere la linea del tipo richiesto
Stabilire qual è la curva che approssima meglio i dati.

Tabella 2

x	y
1	9,5
2	3,5
3	2,5
4	1,4
5	1,2





La linea di tendenza migliore è il polinomio di secondo grado (R^2 è più vicino a 1)

Torna su

Soluzione Esercizio 39
Confronto fra linee di tendenza

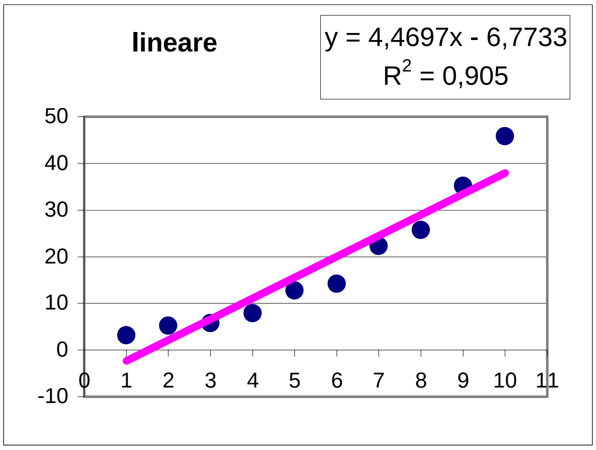
Indice

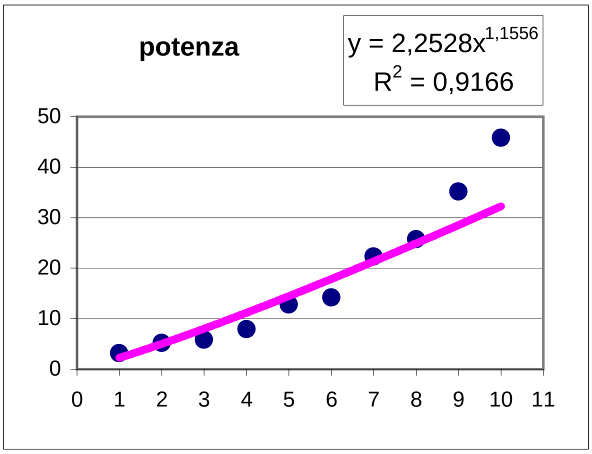
Ritorna Esercizio 39

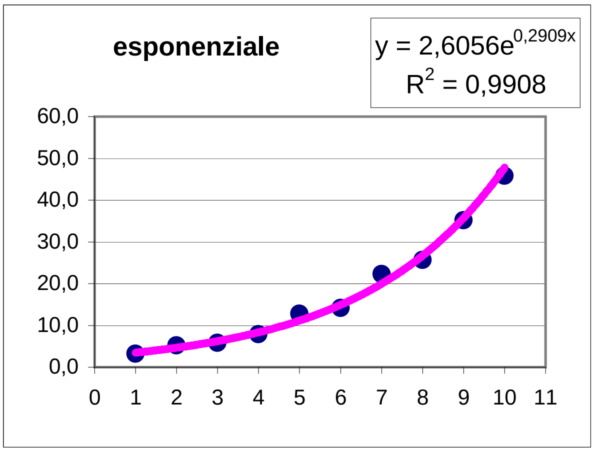
Esercizio 39.1
Sono assegnati i dati della tabella 1
Disegnare in un grafico i dati assegnati; aggiungere una linea di tendenza, confrontando i tipi: lineare, potenza e esponenziale
Realizzare tre grafici, visualizzando l'equazione della linea di tendenza e il coefficiente R^2
Scegliere la linea che approssima meglio i dati

Tabella 1

x	y
1	3,2
2	5,2
3	5,8
4	7,9
5	12,8
6	14,2
7	22,3
8	25,7
9	35,2
10	45,8

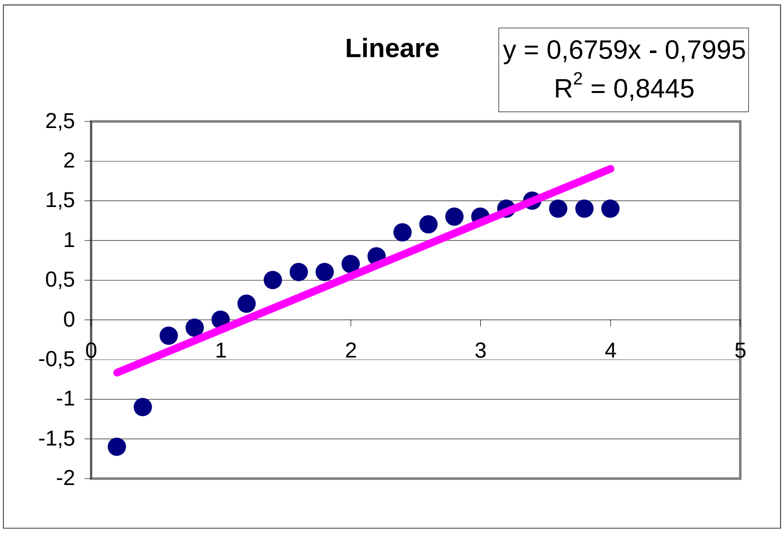






La linea di tendenza migliore è la funzione esponenziale (R^2 è più vicino a 1)

Esercizio 39.2
Sono assegnati i dati della tabella 2
Disegnare in un grafico i dati assegnati; aggiungere una linea di tendenza, confrontando i tipi: lineare e logaritmica
Realizzare due grafici, visualizzando l'equazione della linea di tendenza e il coefficiente R^2
Scegliere la linea che approssima meglio i dati



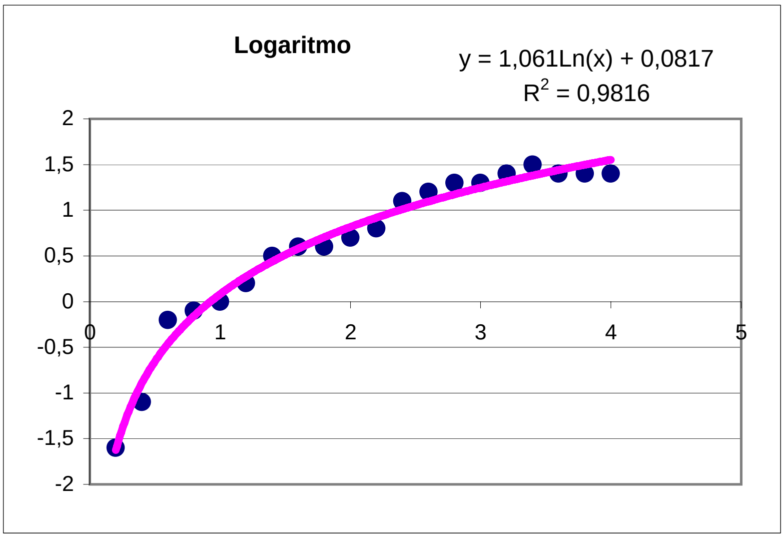


Tabella 2

x	y
0,2	-1,6
0,4	-1,1
0,6	-0,2
0,8	-0,1
1	0
1,2	0,2
1,4	0,5
1,6	0,6
1,8	0,6
2	0,7
2,2	0,8
2,4	1,1
2,6	1,2
2,8	1,3
3	1,3
3,2	1,4
3,4	1,5
3,6	1,4
3,8	1,4
4	1,4

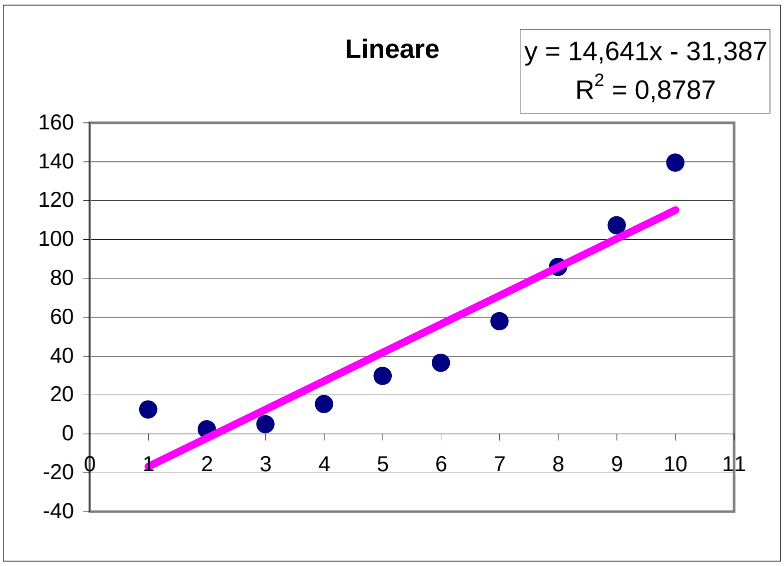
La linea di tendenza migliore è la funzione logaritmo (R^2 è più vicino a 1)

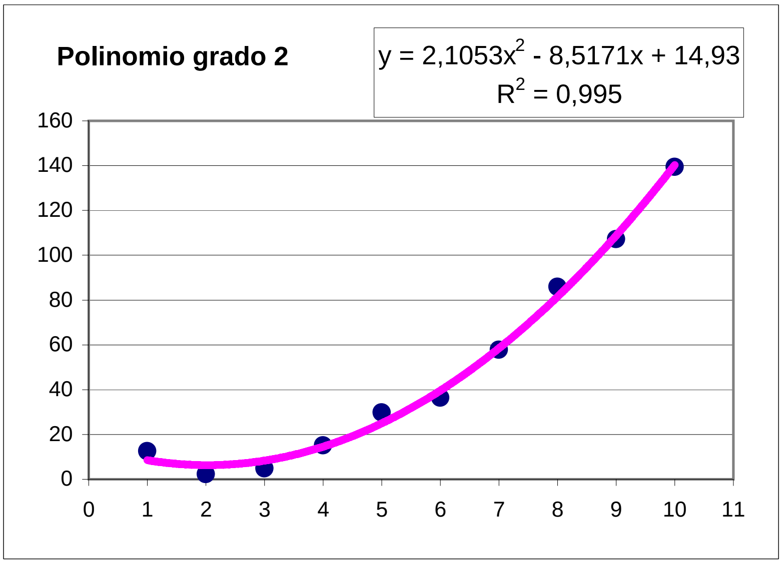
Esercizio 39.3

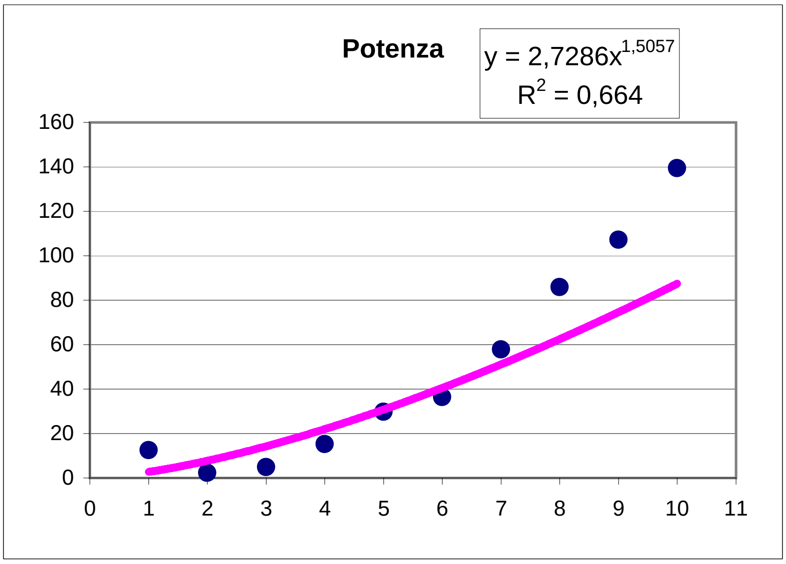
Sono assegnati i dati della tabella 3
Realizzare 4 grafici nel modo seguente:
Disegnare in ogni grafico i dati assegnati e aggiungere una linea di tendenza dei tipi: lineare, polinomio di secondo grado, potenza, esponenziale
Visualizzare l'equazione della linea di tendenza e il coefficiente di correlazione R^2
Scegliere la linea che approssima meglio i dati

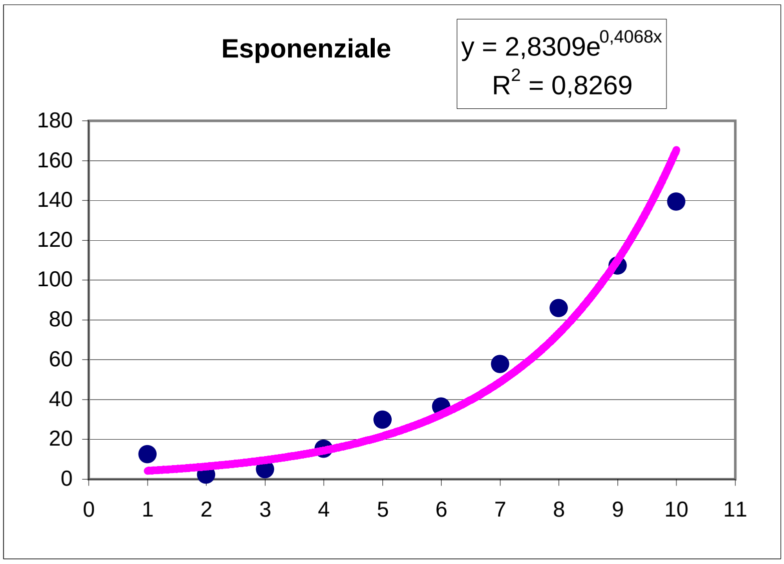
Tabella 3

x	y
1	12,5
2	2,3
3	4,9
4	15,2
5	29,8
6	36,4
7	57,8
8	85,9
9	107,2
10	139,4


Lineare
y = 14,641x - 31,387
R2 = 0,8787
160
140
120
100
80
60
40
20
0
-20
-40
0 1 2 3 4 5 6 7 8 9 10 11


Polinomio grado 2
y = 2,1053x2 - 8,5171x + 14,93
R2 = 0,995
160
140
120
100
80
60
40
20
0
0 1 2 3 4 5 6 7 8 9 10 11


Potenza
y = 2,7286x1,5057
R2 = 0,664
160
140
120
100
80
60
40
20
0
0 1 2 3 4 5 6 7 8 9 10 11



La linea di tendenza che approssima meglio i dati è il polinomio di secondo grado perché il coefficiente R^2 è più vicino a 1.

Torna su

Soluzione Esercizio 40
Curva logistica

Indice

Ritorna Esercizio 40

Esercizio 40.1

I dati y, rilevati a intervalli di tempo x equidistanti, nella tabella 1 caratterizzano un fenomeno di evoluzione studiato in laboratorio.
Disegnare in un grafico i dati assegnati (punti)
Linearizzare i dati scegliendo un tipo di curva adatto ad approssimare i dati

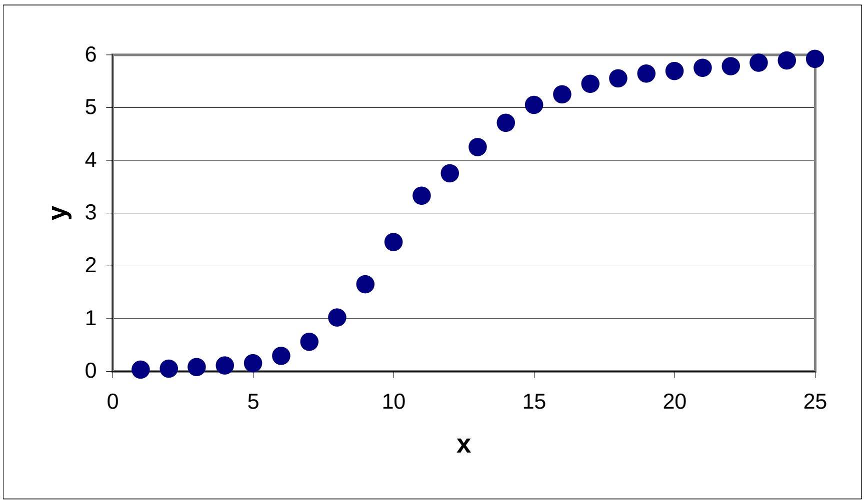


Tabella 1

x	y
1	0,03
2	0,05
3	0,08
4	0,11
5	0,15
6	0,29
7	0,56
8	1,02
9	1,65
10	2,45
11	3,33
12	3,75
13	4,25
14	4,71
15	5,05
16	5,25
17	5,45
18	5,55
19	5,64
20	5,69
21	5,75
22	5,78
23	5,85
24	5,89
25	5,92

SUGGERIMENTI

I dati suggeriscono un andamento descritto dalla **curva logistica**, di equazione

$$y = \frac{L}{1 + Ce^{Ax}}$$

Questo tipo di curva non può essere realizzato in Excel con Aggiungi linea di tendenza;
Occorre linearizzare i dati con i cambiamenti di variabile

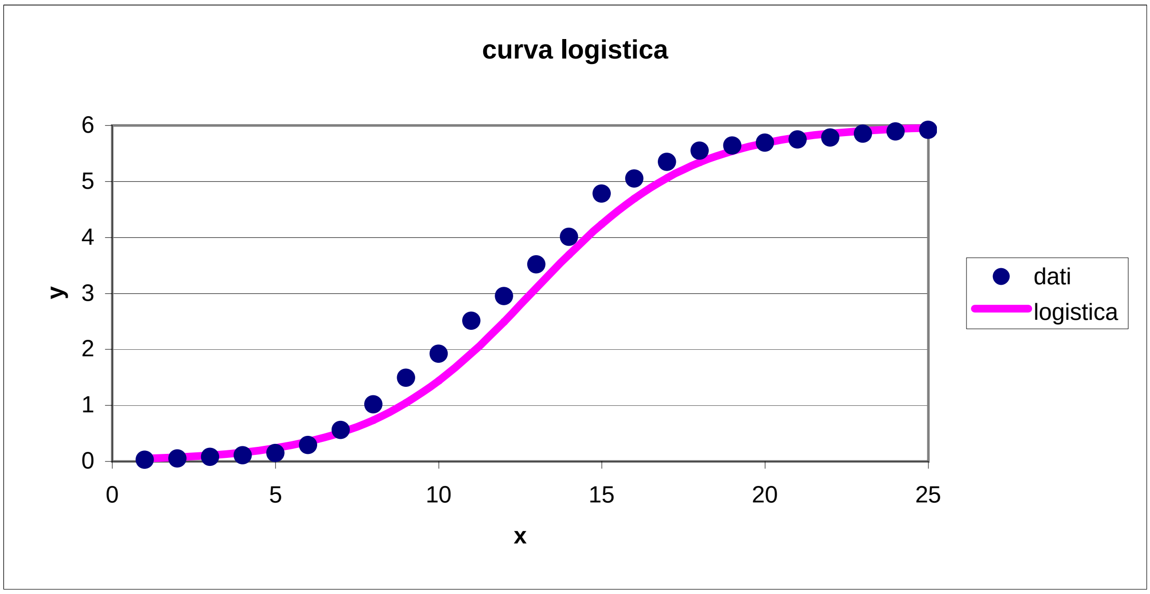
$$X = x \qquad Y = ln\left(\frac{L}{y} - 1\right)$$

$$C = e^{B} \qquad L = \text{costante assegnata}$$

In questo esempio assumere *L* =6.
Sui dati linearizzati si calcolano i coefficienti A e B della retta di regressione, poi si calcola C e infine si traccia il grafico della curva logistica.
Per tracciare il grafico della curva logistica occorre completare la colonna F qui sotto calcolando le ordinate della cuva logistica nei punti x

x	y	X	Y	logistica
1	0,03	1	5,293	0,049
2	0,05	2	4,779	0,073
3	0,08	3	4,304	0,109
4	0,11	4	3,981	0,162
5	0,15	5	3,664	0,240
6	0,29	6	2,980	0,353
7	0,56	7	2,274	0,514
8	1,02	8	1,586	0,740
9	1,49	9	1,108	1,045
10	1,92	10	0,754	1,442
11	2,51	11	0,330	1,931
12	2,95	12	0,033	2,495
13	3,52	13	-0,350	3,098
14	4,01	14	-0,701	3,693
15	4,78	15	-1,366	4,236
16	5,05	16	-1,671	4,696
17	5,35	17	-2,108	5,063
18	5,55	18	-2,512	5,341
19	5,64	19	-2,752	5,544
20	5,69	20	-2,910	5,688
21	5,75	21	-3,135	5,788
22	5,78	22	-3,269	5,857
23	5,85	23	-3,664	5,904
24	5,89	24	-3,981	5,936
25	5,92	25	-4,304	5,957

A	-0,405
B	5,204
C	182,031



Torna su

5. DISTRIBUZIONI DI PROBABILITA'

Soluzione Esercizio 41
Distribuzione binomiale

Indice

Ritorna Esercizio 41

La **distribuzione binomiale** è usata come modello per un processo costituito da un determinato numero di prove; ogni prova del processo ha due soli risultati, indicati con "successo" e "insuccesso".
La probabilità di successo in ogni prova è costante e le prove sono indipendenti
Con la distribuzione binomiale si calcola la probabilità di ottenere un dato numero di successi in un certo numero di prove
Per il calcolo della **distribuzione binomiale** si usa la funzione **DISTRIB.BINOM**

Sintassi
DISTRIB.BINOM(num_successi;prove;probabilità_s;cumulativo)

Num_successi	numero di successi nelle prove effettuate.
Prove	numero di prove indipendenti effettuate.
Probabilità_s	probabilità di successo in ciascuna prova.
Cumulativo	valore logico che determina il tipo di funzione calcolata.

Se il valore cumulativo è VERO, DISTRIB.BINOM restituirà la funzione di ripartizione,
ossia la probabilità di ottenere un numero di successi minore o uguale al valore num_successi.
Se il valore cumulativo è FALSO, verrà restituita la distribuzione di probabilità,
ossia la probabilità di ottenere un numero di successi uguale al valore num_successi.
Nota: in un processo binomiale in cui si effettuano n prove, il numero di successi k varia da 0 a n.

Per il calcolo delle probabilità negli esercizi seguenti sono utili le proprietà della distribuzione binomiale

$$\boxed{\begin{aligned} P(X<k) &= P(X\le k-1) \\ P(X\ge k) &= 1-P(X\le k-1) \\ P(X>k) &= 1-P(X\le k) \end{aligned}}$$

Esempio 41.1
Si effettuano 20 lanci di un dado; il successo sia di ottenere il numero tre.
1 Calcolare la probabilità di ottenere 2 volte il numero tre.

prove	20
probabilità_s	0,1667

Nella finestra della funzione DISTRIB.BINOM per Cumulativo scegliere FALSO

=DISTRIB.BINOM(2;D41;D42;FALSO)

P(X=2)	0,1982

Esempio 41.2
Si effettuano 20 lanci di un dado; il successo sia di ottenere tre.
2 Calcolare la probabilità di ottenere al massimo 2 volte il numero tre.
Nella finestra della funzione per Cumulativo scegliere VERO

=DISTRIB.BINOM(2;D41;D42;VERO)

P(X<=2)	0,3287

3 Calcolare la probabilità di ottenere meno di 2 volte il numero tre

=DISTRIB.BINOM(1;D41;D42;VERO)

P(X<2)	0,1304

4 Calcolare la probabilità di ottenere almeno 2 volte il numero tre

=1-DISTRIB.BINOM(1;D41;D42;VERO)

P(X>=2)	0,8696

5 Calcolare la probabilità di ottenere più di 2 volte il numero tre

=1-DISTRIB.BINOM(2;D41;D42;VERO)

P(X>2)	0,6713

Esercizio 41.3

Costruire la tabella 1, nella quale si riportano le probabilità P(X=k)
e quattro tipi di probabilità cumulative: P(X<=k), P(X<k), P(X>=k), P(X>k)
calcolate con la distribuzione binomiale
Si effettuano n= 6 prove, con probabilità di successo p=0,3

Tabella 1

k	P(X=k)	P(X<=k)	P(X<k)	P(X>=k)	P(X>k)
0	0,1176	0,1176	0	1	0,8824
1	0,3025	0,4202	0,1176	0,8824	0,5798
2	0,3241	0,7443	0,4202	0,5798	0,2557
3	0,1852	0,9295	0,7443	0,2557	0,0705
4	0,0595	0,9891	0,9295	0,0705	0,0109
5	0,0102	0,9993	0,9891	0,0109	0,0007
6	0,0007	1,0000	0,9993	0,0007	0,0000

Esercizio 41.4

Si effettuano 10 lanci di una moneta.

1 Calcolare la probabilità che per metà delle volte esca croce e per metà testa. — P(X=5) 0,2461

2 Calcolare la probabilità di non ottenere mai testa — P(X=0) 0,0010

Esercizio 41.5

Un venditore di auto sa, per esperienza precedente, che il 20% delle persone che visitano il suo punto vendita acquisterà un'auto nuova. Calcolare la probabilità che su 5 clienti

1	3 acquistino un'auto;	P(X=3)	0,0512
2	nessuno acquisti un'auto;	P(X=0)	0,3277
3	al più 2 acquistino un'auto;	P(X<=2)	0,9421
4	meno di 2 acquistino un'auto	P(X<2)	0,7373
5	almeno 2 acquistino un'auto;	P(X>=2)	0,2627
6	più di 2 acquistino un'auto.	P(X>2)	0,0579

Esercizio 41.6

La probabilità che un apparecchio si guasti è p=0,05; calcolare la probabilità che su 16 di tali apparecchi

1	al più 2 si guastino	P(X<=2)	0,9571
2	almeno 2 si guastino	P(X>=2)	0,1892
3	meno di 4 si guastino	P(X<4)	0,9930

Esercizio 41.7

Trovare la probabilità che in 5 lanci di un dado il numero quattro esca

1	2 volte	P(X=2)	0,1608
2	al più 1 volta	P(X<=1)	0,8038
3	almeno 2 volte	P(X>=2)	0,1962

Esercizio 41.8

Se il 5% dei chip di memoria prodotti da una macchina sono difettosi, determinare la probabilità che su 4 chip scelti a caso

1. 1 sia difettoso
2. nessuno sia difettoso
3. meno di 2 siano difettosi
4. più di 2 siano difettosi

P(X=1)	0,1715
P(X=0)	0,8145
P(X<2)	0,9860
P(X>2)	0,0005

Torna su

Soluzione Esercizio 42

Calcolo di probabilità con la distribuzione binomiale e grafico

Indice

Ritorna Esercizio 42

Esercizio 42.1

Un test è composto da 20 domande; ciascuna domanda ha 4 risposte possibili, di cui una sola è corretta.

1 Rispondendo a caso al test, qual è la probabilità di superarlo, se occorrono almeno 12 risposte corrette?

numero prove	20
prob. successo	0,25
P(X>=12)	0,0009354

2 Se si conosce la risposta corretta a 4 domande, qual è la probabilità di superare il test rispondendo a caso alle altre?

numero prove	16
prob. successo	0,25
P(X>=8)	0,0271

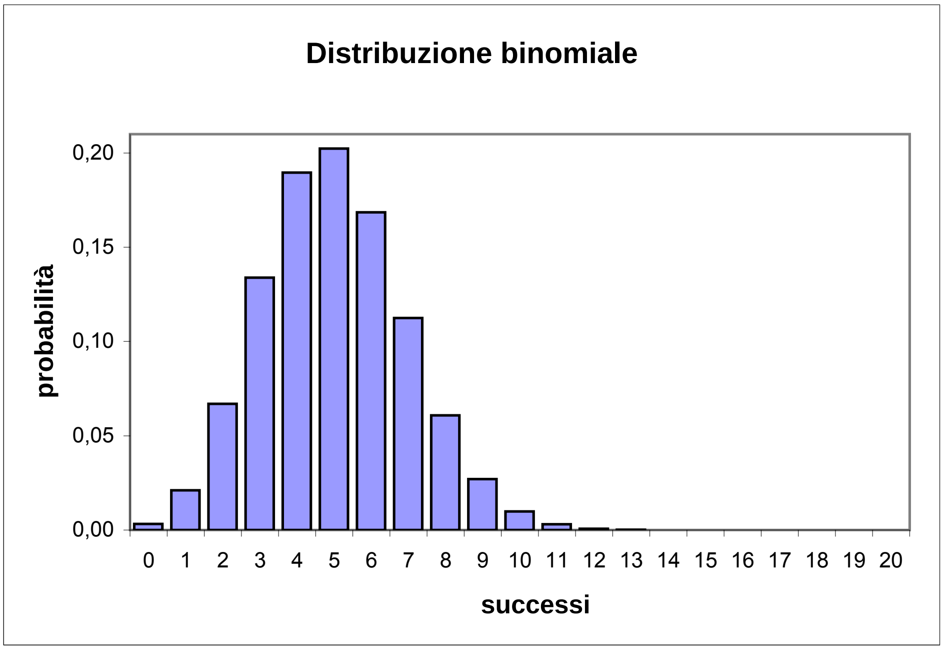
3 Calcolare le probabilità di rispondere (a caso) esattamente a 0, 1, 2, 3, …, 20 domande e disegnare il grafico della distribuzione.

SUGGERIMENTI

La variabile aleatoria X indica il numero di successi e può assumere i valori da 0 a 20; costruire la tabella 1 in cui si indicano tutti i valori della variabile aleatoria (ossia tutti i possibili numeri di successi) e le corrispondenti probabilità.
Servendosi della tabella così costruita si può realizzare il grafico (diagramma a barre o istogramma)
Il grafico non è simmetrico. Da che cosa dipende l'asimmetria?

Tabella 1

Successi	Probabilità
0	0,00317
1	0,02114
2	0,06695
3	0,13390
4	0,18969
5	0,20233
6	0,16861
7	0,11241
8	0,06089
9	0,02706
10	0,00992
11	0,00301
12	0,00075
13	0,00015
14	0,00003
15	0,0000034
16	0,0000004
17	0,00000003
18	0,000000002
19	0,0000000001
20	0,000000000001



Il grafico è asimmetrico positivamente (distribuzione obliqua a destra); l'asimmetria dipende dal valore della probabiltà di successo, che è minore di 0,5.

Torna su

Soluzione Esercizio 43
Grafici della distribuzione binomiale

Indice

Ritorna Esercizio 43

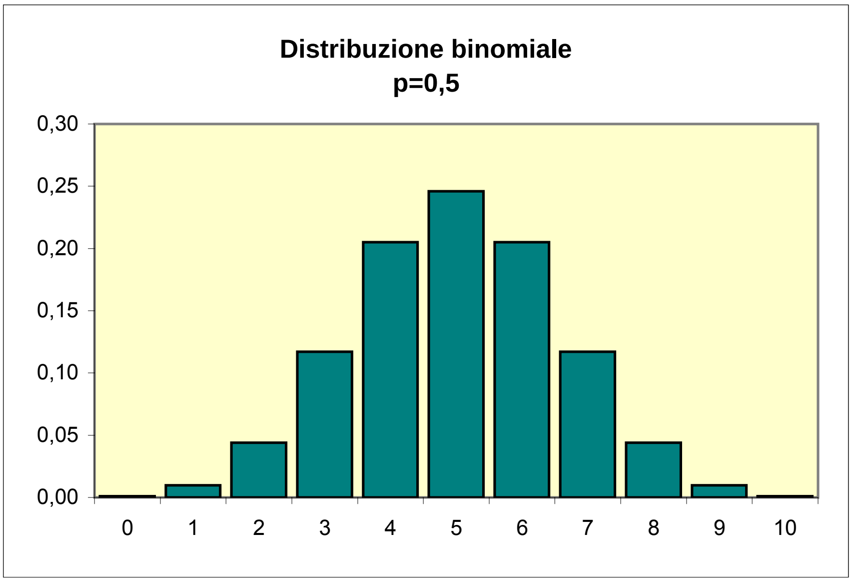
Esercizio 43.1

Si effettuano 10 lanci di una moneta; studiare la distribuzione di probabilità della variabile aleatoria binomiale X=numero di teste uscite nei 10 lanci. Usare la funzione DISTRIB.BINOM Disegnare il grafico. Osservare la simmetria

numero prove	10
probabilità successo	0,5

Tabella 1

Successi k	Probabilità P(X=k)
0	0,0009766
1	0,0097656
2	0,0439453
3	0,1171875
4	0,2050781
5	0,2460938
6	0,2050781
7	0,1171875
8	0,0439453
9	0,0097656
10	0,0009766



Il grafico è simmetrico: la simmetria dipende dal valore della probabilità di successo p=0,5

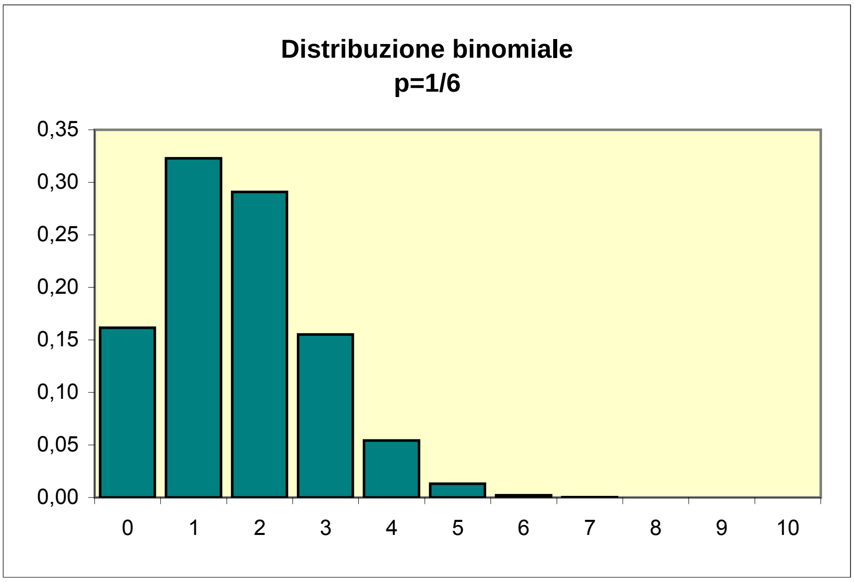
Esercizio 43.2

Si effettuano 10 lanci di un dado. Studiare la distribuzione di probabilità della variabile aleatoria binomiale X = numero di uscite del numero tre (o di un qualunque altro numero fra i sei possibili)
Disegnare il grafico. Osservare l'asimmetria.

numero prove	10
probabilità successo	0,166667

Tabella 2

Successi	Probabilità
0	0,1615
1	0,3230
2	0,2907
3	0,1550
4	0,0543
5	0,0130
6	0,0022
7	0,00025
8	0,000019
9	0,0000008
10	0,00000002



Il grafico non è simmetrico (asimmetria positiva); l'asimmetria positiva dipende dal valore di p=1/6

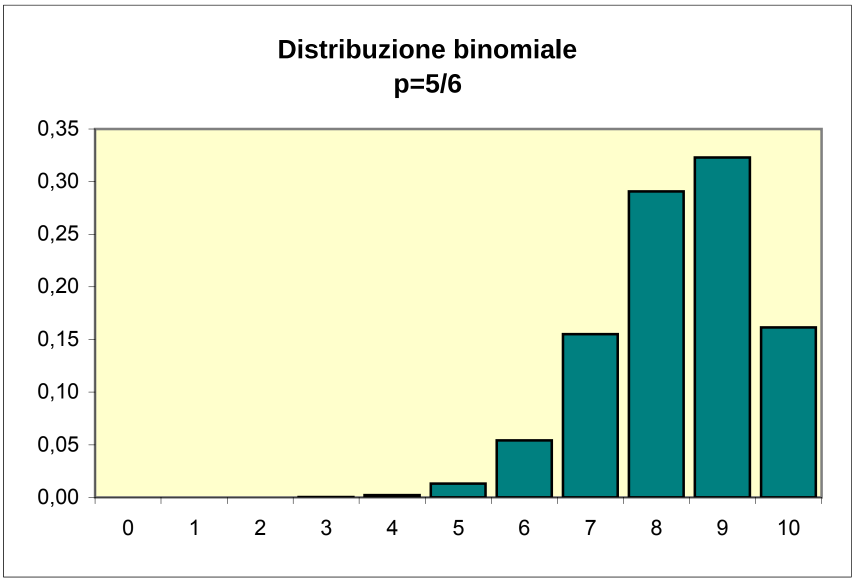
Esercizio 43.3

Si effettuano 10 lanci di un dado. Studiare la distribuzione di probabilità della variabile aleatoria binomiale X = numero di uscite di un numero diverso da tre
Disegnare il grafico. Osservare l'asimmetria.

numero prove	10
probabilità successo	0,833

Tabella 3

Successi	Probabilità
0	0,00000002
1	0,00000083
2	0,0000186
3	0,000248
4	0,0022
5	0,0130
6	0,0543
7	0,1550
8	0,2907
9	0,3230
10	0,1615



Il grafico non è simmetrico (asimmetria negativa); l'asimmetria negativa dipende dal valore di p=5/6

Torna su

Soluzione Esercizio 44
Distribuzione di Poisson

Indice

Ritorna Esercizio 44

La **distribuzione di Poisson** è usata per studiare il numero di eventi rari che si realizzano in un dato intervallo di tempo (o di spazio); gli eventi accadono in modo indipendente l'uno dall'altro.
Il numero di eventi che si realizzano nel dato intervallo varia da 0 a n, e n non è determinabile a priori
Per il calcolo della **distribuzione di probabilità di Poisson** si usa la funzione **POISSON**
Sintassi
POISSON(x;media;cumulativo)

x	numero degli eventi.
media	valor medio della distribuzione di Poisson.
Cumulativo	valore logico che determina il tipo di funzione calcolata.

Se cumulativo è VERO, POISSON restituirà la funzione di ripartizione di Poisson ossia la probabilità che il numero degli eventi casuali sia compreso tra zero e x inclusi
Se il valore cumulativo è FALSO, verrà restituita la distribuzione di probabilità, ossia la probabilità che il numero di eventi sia uguale a x.

Per il calcolo delle probabilità negli esercizi seguenti sono utili le proprietà della distribuzione di Poisson (analoghe alle proprietà della binomiale, esercizio 41)

$$\boxed{\begin{aligned} P(X<k) &= P(X \le k-1) \\ P(X \ge k) &= 1-P(X \le k-1) \\ P(X>k) &= 1-P(X \le k) \end{aligned}}$$

Esempio 44.1
Dalle statistiche degli ultimi cinque anni un'azienda ha calcolato che ogni giorno sono assenti in media 1,8 dipendenti.
1 Calcolare la probabilità che in un giorno siano assenti 3 dipendenti.

media	1,8

Nella finestra della funzione per Cumulativo scegliere FALSO

=POISSON(3;D36;FALSO)

P(X=3)	0,1607

Esempio 44.2
2 Calcolare la probabilità che in un giorno siano assenti al più 4 dipendenti
Nella finestra della funzione per Cumulativo scegliere VERO

=POISSON(4;D36;VERO)

P(X<=4)	0,9636

3 Calcolare la probabilità che in un giorno siano assenti meno di 4 dipendenti

=POISSON(3;D36;VERO)

P(X<4)	0,8913

4 Calcolare la probabilità che in un giorno siano assenti almeno 4 dipendenti

=1-POISSON(3;D36;VERO)

P(X>=4)	0,1087

5 Calcolare la probabilità che in un giorno siano assenti più di 4 dipendenti

P(X>4)	0,0364

=1-POISSON(4;D36;VERO)

Esercizio 44.3

Costruire la tabella 1, nella quale si riportano le probabilità P(X=k)
e quattro tipi di probabilità cumulative: P(X<=k), P(X<k), P(X>=k), P(X>k)
calcolate con la distribuzione di Poisson, con valor medio lambda=3
Nota: in un processo di Poisson il numero di eventi assume i valori k = 0, 1, 2, 3, …;
nella tabella arrestare il calcolo al valore k=10

Tabella 1

k	P(X=k)	P(X<=k)	P(X<k)	P(X>=k)	P(X>k)
0	0,0498	0,0498	0	1	0,9502
1	0,1494	0,1991	0,0498	0,9502	0,8009
2	0,2240	0,4232	0,1991	0,8009	0,5768
3	0,2240	0,6472	0,4232	0,5768	0,3528
4	0,1680	0,8153	0,6472	0,3528	0,1847
5	0,1008	0,9161	0,8153	0,1847	0,0839
6	0,0504	0,9665	0,9161	0,0839	0,0335
7	0,0216	0,9881	0,9665	0,0335	0,0119
8	0,0081	0,9962	0,9881	0,0119	0,0038
9	0,0027	0,9989	0,9962	0,0038	0,0011
10	0,0008	0,9997	0,9989	0,0011	0,0003

Esercizio 44.4

A un servizio di guardia medica ogni ora arrivano in media 3,5 richieste di interventi urgenti a domicilio.

1 Calcolare la probabilità che in una data ora arrivino 3, 4, 5 chiamate urgenti.
2 Calcolare la probabilità che in una data ora arrivi un numero di chiamate urgenti compreso fra 3 e 5.
3 Calcolare la probabilità che in una data ora arrivi un numero di chiamate urgenti maggiore di 3.

	media	3,5
1	**P(X=3)**	0,2158
	P(X=4)	0,1888
	P(X=5)	0,1322
2	**P(3<=X<=5)**	0,5368
3	**P(X>3)**	0,4634

Esercizio 44.5

Un libro di 500 pagine contiene 50 errori di stampa. Calcolare la probabilità di trovare almeno 3 errori in una qualsiasi pagina

media	0,1
P(X>=3)	0,00015

Esercizio 44.6

Il numero di errori che si verificano in un giorno in una rete locale Lan è distribuito secondo la legge di Poisson e il numero medio di errori in un giorno è 2,5.
Calcolare le probabilità che in un giorno:

1 non si verifichino errori nella rete.
2 si verifichi 1 errore.
3 si verifichino almeno 2 errori.
4 si verifichino meno di tre errori.

media	2,5

1	**P(X=0)**	0,0821
2	**P(X=1)**	0,2052
3	**P(X>=2)**	0,7127
4	**P(X<3)**	0,5438

Torna su

Soluzione Esercizio 45
Grafico della distribuzione di Poisson

Indice

Ritorna Esercizio 45

Esercizio 45.1

Disegnare i grafici della distribuzione di Poisson per i seguenti valori della media:

1 lambda = 1
2 lambda = 2
3 lambda = 4
4 lambda = 12

Si osservi che all'aumentare della media il grafico presenta una maggior simmetria.

SUGGERIMENTI

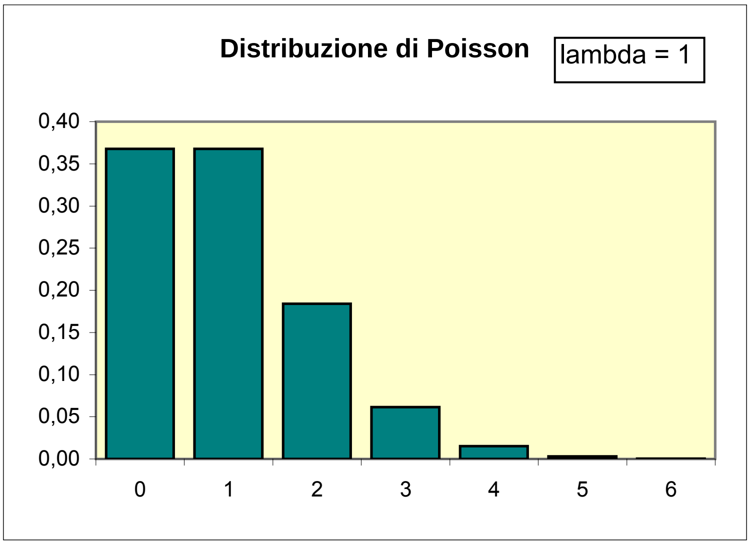
La variabile aleatoria X indica il numero di eventi e può assumere i valori 0, 1, 2, ... ; costruire le tabelle 1-4 in cui si indicano i valori della variabile aleatoria e le corrispondenti probabilità. A seconda del valore della media, si arresta il calcolo a un opportuno numero di eventi, perchè le probabilità diventano sempre più vicine a 0. Servendosi delle tabelle così costruite si possono realizzare i grafici (diagrammi a barre o istogrammi)

1

media	1

Tabella 1

Eventi k	Probabilità P(X=k)
0	0,36788
1	0,36788
2	0,18394
3	0,06131
4	0,01533
5	0,00307
6	0,00051

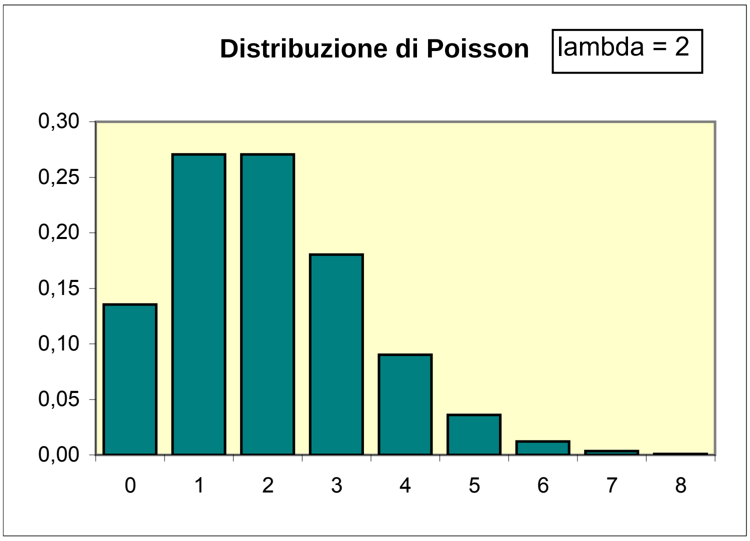


2

media	2

Tabella 2

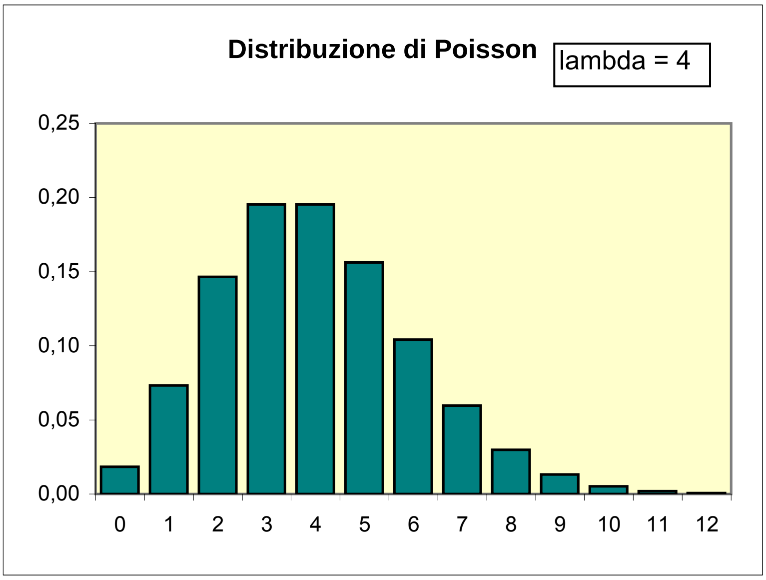
Eventi k	Probabilità P(X=k)
0	0,13534
1	0,27067
2	0,27067
3	0,18045
4	0,09022
5	0,03609
6	0,01203
7	0,00344
8	0,00086



3	media	4

Tabella 3

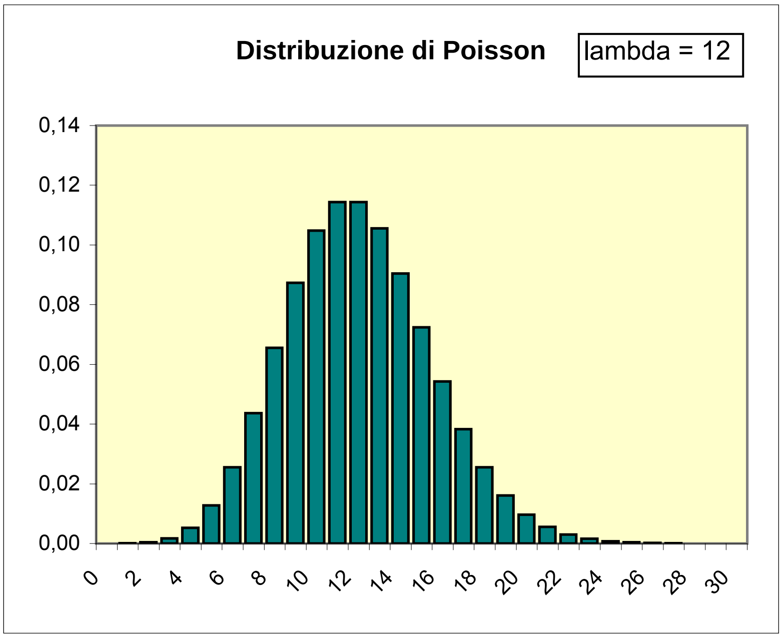
Eventi k	Probabilità P(X=k)
0	0,01832
1	0,07326
2	0,14653
3	0,19537
4	0,19537
5	0,15629
6	0,10420
7	0,05954
8	0,02977
9	0,01323
10	0,00529
11	0,00192
12	0,00064



4	media	12

Tabella 4

Eventi k	Probabilità P(X=k)
0	0,00001
1	0,00007
2	0,00044
3	0,00177
4	0,00531
5	0,01274
6	0,02548
7	0,04368
8	0,06552
9	0,08736
10	0,10484
11	0,11437
12	0,11437
13	0,10557
14	0,09049
15	0,07239
16	0,05429
17	0,03832
18	0,02555
19	0,01614
20	0,00968
21	0,00553
22	0,00302
23	0,00157
24	0,00079
25	0,00038
26	0,00017
27	0,00008
28	0,00003
29	0,00001
30	0,00001



Torna su

Soluzione Esercizio 46

Distribuzione di Poisson e distribuzione binomiale.

Indice

Ritorna Esercizio 46

Quando il numero di prove n è grande e la probabilità di successo p è piccola la distribuzione binomiale può essere approssimata con la distribuzione di Poisson avente media lambda=np.

Regola pratica per ottenere una buona approssimazione: usare la distribuzione di Poisson quando n>=50 e p<=0,1

Esercizio 46.1

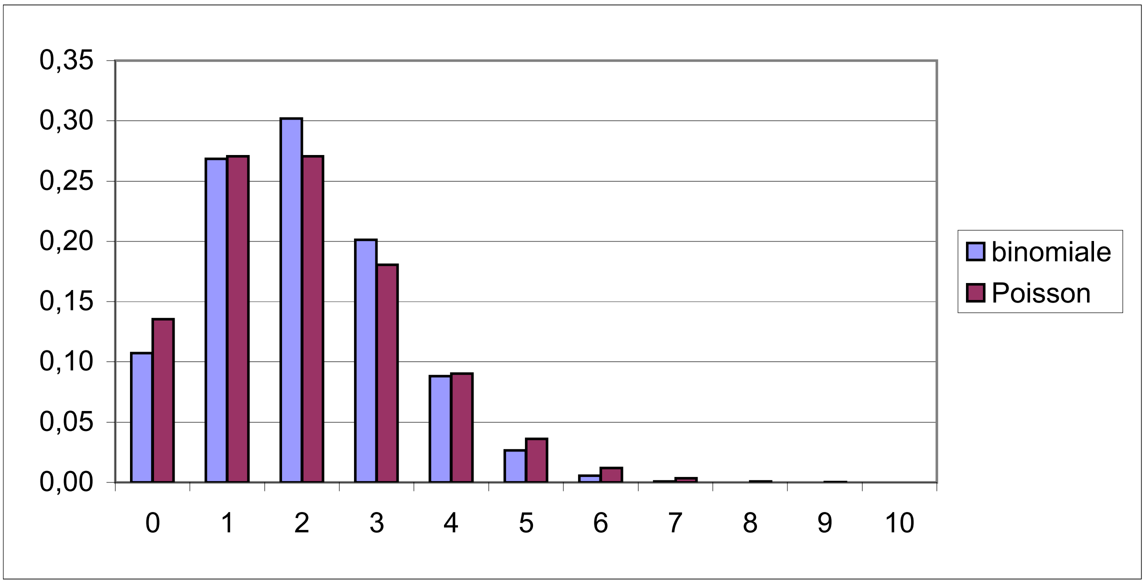
La probabilità che un oggetto prodotto da una macchina sia difettoso è p=0,2.
Calcolare le probabilità che in un campione di 10 oggetti scelti a caso ci siano 0, 1, 2,…, 10 oggetti difettosi usando sia la distribuzione binomiale che la distribuzione di Poisson (Tabella 1).
Confrontare su un grafico i risultati ottenuti.
Ripetere il procedimento nel caso di un campione di 100 oggetti, con probabilità di successo p=0,1 (Tabella 2)

prove	10
probabilità successo	0,2

media	2

Tabella 1

Successi	Probabilità binomiale	Probabilità Poisson
0	0,1074	0,1353
1	0,2684	0,2707
2	0,3020	0,2707
3	0,2013	0,1804
4	0,0881	0,0902
5	0,0264	0,0361
6	0,0055	0,0120
7	0,0008	0,0034
8	0,0001	0,0009
9	0,0000	0,0002
10	0,0000	0,0000



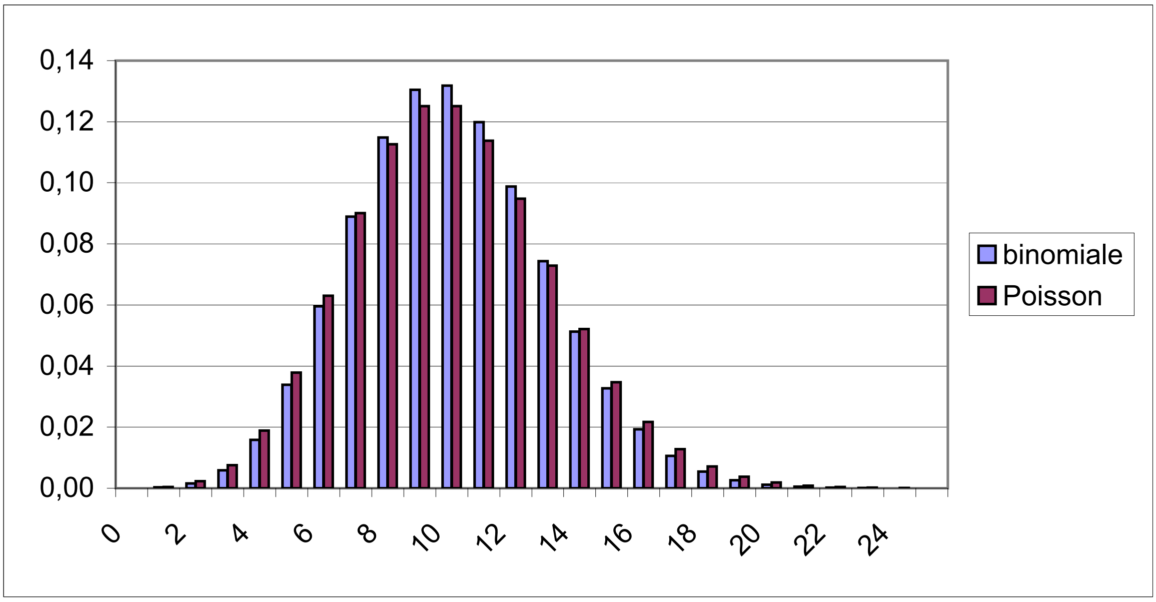
La distribuzione di Poisson non fornisce un'approssimazione molto precisa: la regola pratica suggerita non è soddisfatta.

prove	100
probabilità_successo	0,1000

media	10,0000

Tabella 2

Successi	Probabilità binomiale	Probabilità Poisson
0	0,0000	0,0000
1	0,0003	0,0005
2	0,0016	0,0023
3	0,0059	0,0076
4	0,0159	0,0189
5	0,0339	0,0378
6	0,0596	0,0631
7	0,0889	0,0901
8	0,1148	0,1126
9	0,1304	0,1251
10	0,1319	0,1251
11	0,1199	0,1137
12	0,0988	0,0948
13	0,0743	0,0729
14	0,0513	0,0521
15	0,0327	0,0347
16	0,0193	0,0217
17	0,0106	0,0128
18	0,0054	0,0071
19	0,0026	0,0037
20	0,0012	0,0019
21	0,0005	0,0009
22	0,0002	0,0004
23	0,0001	0,0002
24	0,0000	0,0001
25	0,0000	0,0000



La distribuzione di Poisson fornisce una buona approssimazione: la regola pratica suggerita è soddisfatta.

Esercizio 46.2

La probabilità che una persona sia allergica a un farmaco è p=0,001.
Calcolare la probabilità che su 2000 persone

1 tre siano allergiche
2 meno di due siano allergiche
3 più di due siano allergiche.

prove	2000
probabilità_successo	0,001

media	2

Distribuzione di Poisson	
P(X=3)	0,1804
P(X<2)	0,4060
P(X>2)	0,3233

Torna su

Soluzione Esercizio 47
Distribuzione normale non standardizzata

Indice

Ritorna Esercizio 47

La **distribuzione normale** è usata come modello per molti processi nel modo reale
Ad esempio descrive la distribuzione degli errori casuali nelle misure di una quantità fisica
Il grafico della distribuzione normale è una curva a forma di campana; l'area totale sottesa dalla curva è uguale a 1.
Per individuare una particolare distribuzione normale occorrono due parametri: la media e lo scarto quadratico medio (o deviazione standard)
Per il calcolo della **distribuzione normale non standardizzata** si usa la funzione **DISTRIB.NORM**

Sintassi
DISTRIB.NORM(x;media;dev_standard;cumulativo)

x	valore per il quale si vuole calcolare la distribuzione.
Media	valor medio della distribuzione.
Dev_standard	deviazione standard (scarto quadratico medio) della distribuzione.
Cumulativo	valore logico che determina il tipo di funzione calcolata.

La funzione di solito viene utilizzata con l'argomento Cumulativo uguale a VERO
Se cumulativo è VERO, DISTRIB.NORM restituisce la funzione di ripartizione normale F(x), ossia la probabilità che la variabile aleatoria normale sia minore di x (coda sinistra)
Se è FALSO restituisce l'ordinata della distribuzione di probabilità normale f(x).
L'ordinata non può essere interpretata come probabilità, ma è utile per disegnare il grafico della curva a campana

Osservazione
La terminologia corretta per le variabili aleatorie continue è "densità di probabilità" , mentre il nome "distribuzione di probabilità" è usato per le variabili discrete; è tuttavia ampiamente diffusa nella letteratura statistica la consuetudine di usare il termine "distribuzione" in luogo di "densità" anche per le variabili continue.

Esempio 47.1
E' data una variabile aleatoria X avente distribuzione normale con media uguale a 4,35 e scarto quadratico medio uguale a 0,59

valor medio	4,35
deviazione standard	0,59

Calcolare la probabilità che la variabile aleatoria X assuma valori minori di 5
Nella finestra della funzione per il parametro Cumulativo scegliere VERO

P(X<5)	0,8647	**=DISTRIB.NORM(5;F39;F40;VERO)**

Calcolare la probabilità che la variabile aleatoria X assuma valori compresi fra 4 e 5

P(4<X<5)	0,5882	**=DISTRIB.NORM(5;F39;F40;VERO)-DISTRIB.NORM(4;F39;F40;VERO)**

Calcolare la probabilità che la variabile aleatoria X assuma valori maggiori di 4

P(X>4)	0,7235	**=1-DISTRIB.NORM(4;F39;F40;VERO)**

Esercizio 47.2

E' data una variabile aleatoria X avente distribuzione normale con media uguale a 100 e scarto quadratico medio uguale a 15

valor medio	100
scarto quadratico medio	15

Calcolare la probabilità che la variabile aleatoria assuma valori

1 minori di 118	**P(X<118)**	0,8849
2 maggiori di 112	**P(X>112)**	0,2119
3 compresi fra 110 e 120	**P(110<X<120)**	0,1613

Esercizio 47.3

Il peso delle confezioni di pasta di una data marca è una variabile aleatoria X avente distribuzione normale con valor medio 500 g e scarto quadratico medio 20 g

valor medio	500
scarto quadratico medio	20

Calcolare la probabilità che un pacco scelto a caso abbia peso

1 al più 475 g	**P(X<475)**	0,1056
2 almeno 495 g	**P(X>495)**	0,5987
3 compreso fra 490 g e 510 g	**P(490<X<510)**	0,3829
4 compreso fra 480 g e 520 g	**P(480<X<520)**	0,6827
5 compreso fra 460 g e 540 g	**P(460<X<540)**	0,9545

Esercizio 47.4

Il peso netto delle scatole di cioccolatini di una certa marca si distribuisce normalmente con valor medio 1005 g e scarto quadratico medio 15 g.

valor medio	1005
scarto quadratico medio	15

Calcolare la percentuale di scatole con peso netto

1 compreso tra 990 g e 1020 g	**P(990<X<1020)**	68,3%
2 maggiore di 980 g;	**P(X>980)**	95,2%

3 In un campione di 400 scatole quante pesano più di 980 g?

numero scatole	381

Esercizio 47.5

Il quoziente di intelligenza degli adulti è una variabile aleatoria X avente distribuzione normale con media 100 e scarto quadratico medio 15.

valor medio	100
scarto quadratico medio	15

Calcolare la probabilità che un adulto selezionato a caso abbia un quoziente di intelligenza

1	minore di 90	**P(X<90)**	0,2525
2	compreso fra 90 e 110 (normale)	**P(90<X<110)**	0,4950
3	maggiore di 110 (brillante)	**P(X>110)**	0,2525

Torna su

Soluzione Esercizio 48
Distribuzione normale standardizzata

Indice

Ritorna Esercizio 48

Per il calcolo della **distribuzione normale standardizzata** si usa la funzione **DISTRIB.NORM.ST**

Sintassi
DISTRIB.NORM.ST(z)

z valore per il quale si calcola la funzione di ripartizione F(z), ossia la probabilità P(Z<z)

Osservare che, a differenza della funzione DISTRIB.NORM, non è previsto il calcolo di f(z)

Esempio 48.1

Calcolare la probabilità che una variabile aleatoria Z avente distribuzione normale standardizzata assuma valori

1 minori di 2	**P(Z<2)**	0,9772	**=DISTRIB.NORM.ST(2)**
2 maggiori di 1	**P(Z>1)**	0,1587	**=1-DISTRIB.NORM.ST(1)**
3 compresi fra -1 e 2	**P(-1<Z<2)**	0,8186	**=DISTRIB.NORM.ST(2)-DISTRIB.NORM.ST(-1)**

Si può anche usare la funzione DISTRIB.NORM; in tal caso assegnare valor medio=0 e deviazione standard=1; per Cumulativo usare VERO

P(Z<2)	0,9772

Esercizio 48.2

Calcolare la probabilità che una variabile aleatoria Z avente distribuzione normale standardizzata assuma valori

1 minori di 0,75	**P(Z<0,75)**	0,7734
2 compresi fra 0,87 e 1,28	**P(0,87<Z<1,28)**	0,0919
3 compresi fra -0,34 e 0,62	**P(-0,34<Z<0,62)**	0,3654
4 maggiori di 0,85	**P(Z>0,85)**	0,1977
5 maggiori di -0,65	**P(Z>-0,65)**	0,7422

Torna su

Soluzione Esercizio 49

Grafici della distribuzione normale e della funzione di ripartizione normale
Grafici della distribuzione normale standardizzata e della funzione di ripartizione normale standardizzata

Indice

Ritorna Esercizio 49

Esercizio 49.1

Disegnare i grafici della distribuzione normale e della funzione di ripartizione normale di valor medio 5 e varianza 4

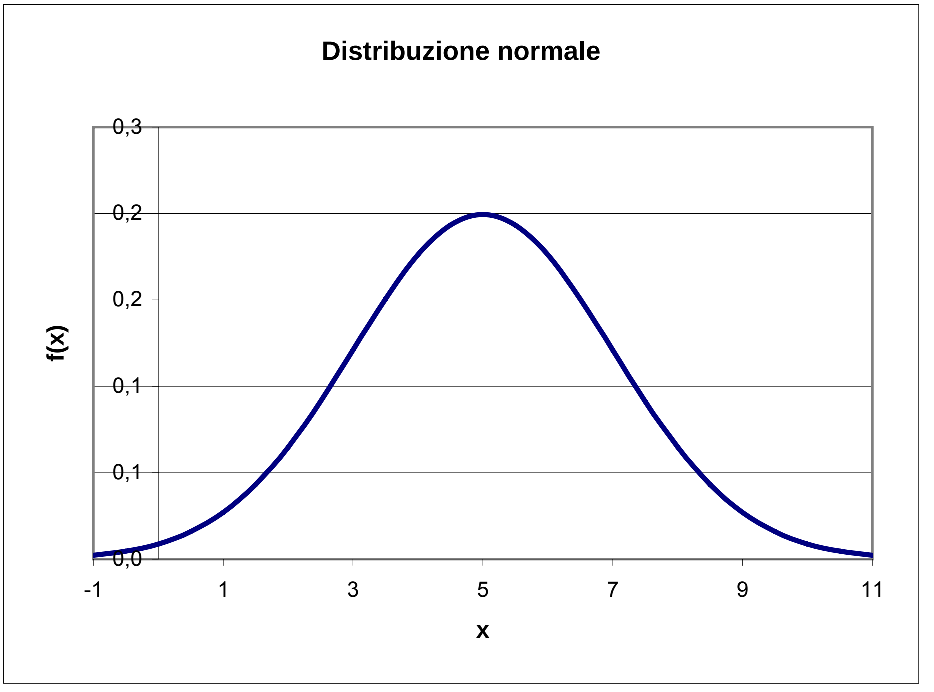
SUGGERIMENTI

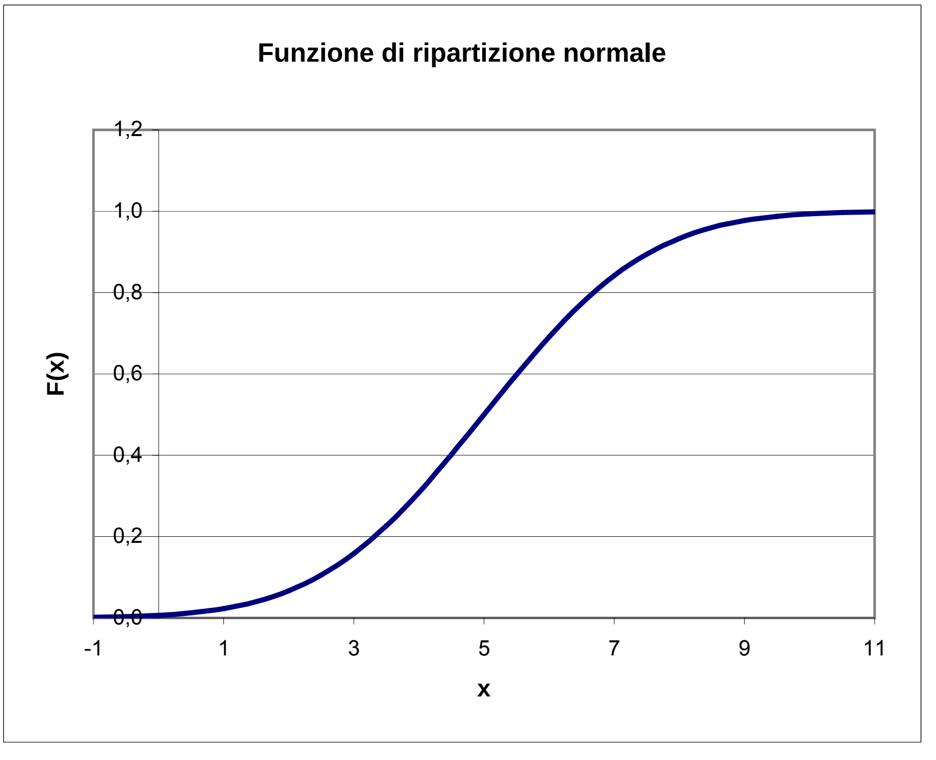
Per disegnare il grafico della distribuzione normale f(x) e della funzione di ripartizione F(x) scegliere un intervallo simmetrico intorno al valor medio (Tabella 1, prima colonna); iniziare da un valore x distante dalla media di tre volte la deviazione standard (in questo caso iniziare da x= –1 e applicare un incremento pari a 0,5, fino a raggiungere il valore x=11)
Usare la funzione DISTRIB.NORM per calcolare i valori di f(x) e di F(x) nell'intervallo scelto, usando per Cumulativo rispettivamente FALSO e VERO (Tabella 1, seconda e terza colonna)

valor medio	5
deviazione standard	2

Tabella 1

x	f(x)	F(x)
-1	0,0022	0,0013
-0,5	0,0045	0,0030
0	0,0088	0,0062
0,5	0,0159	0,0122
1	0,0270	0,0228
1,5	0,0431	0,0401
2	0,0648	0,0668
2,5	0,0913	0,1056
3	0,1210	0,1587
3,5	0,1506	0,2266
4	0,1760	0,3085
4,5	0,1933	0,4013
5	0,1995	0,5000
5,5	0,1933	0,5987
6	0,1760	0,6915
6,5	0,1506	0,7734
7	0,1210	0,8413
7,5	0,0913	0,8944
8	0,0648	0,9332
8,5	0,0431	0,9599
9	0,0270	0,9772
9,5	0,0159	0,9878
10	0,0088	0,9938
10,5	0,0045	0,9970
11	0,0022	0,9987





Esercizio 49.2

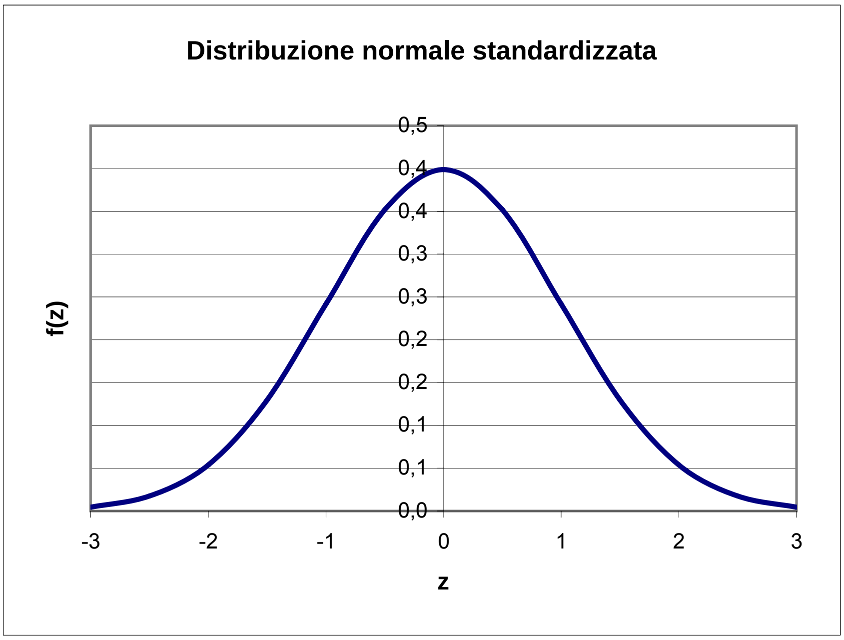
Disegnare i grafici della distribuzione normale standardizzata f(z) e della funzione di ripartizione normale standardizzata F(z)

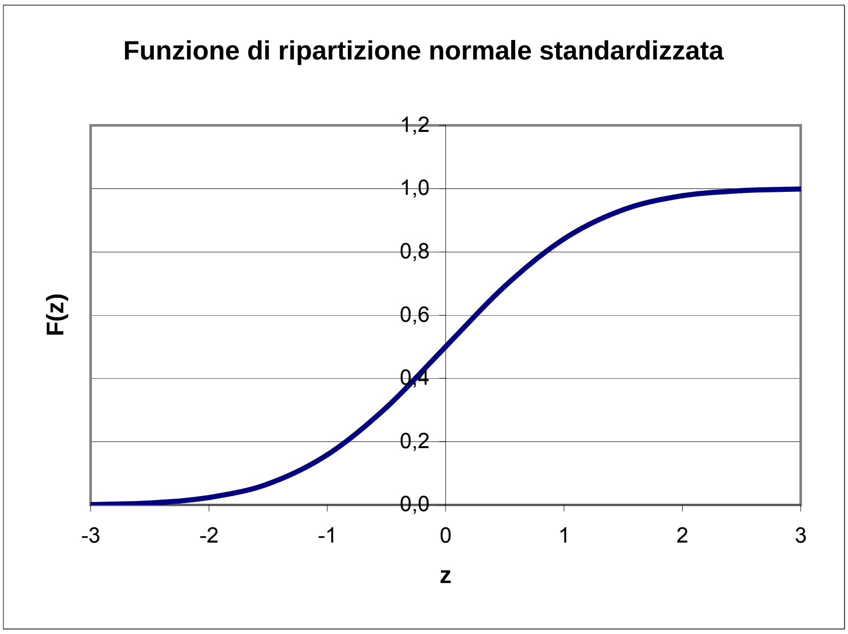
SUGGERIMENTI

Usare la funzione DISTRIB.NORM per calcolare i valori di f(z) e la funzione DISTRIB.NORM.ST per calcolare i valori di F(z) nell'intervallo (-3,3) con incremento=0,5

valor medio	0
deviazione standard	1

z	f(z)	F(z)
-3	0,0044	0,0013
-2,5	0,0175	0,0062
-2	0,0540	0,0228
-1,5	0,1295	0,0668
-1	0,2420	0,1587
-0,5	0,3521	0,3085
0	0,3989	0,5000
0,5	0,3521	0,6915
1	0,2420	0,8413
1,5	0,1295	0,9332
2	0,0540	0,9772
2,5	0,0175	0,9938
3	0,0044	0,9987

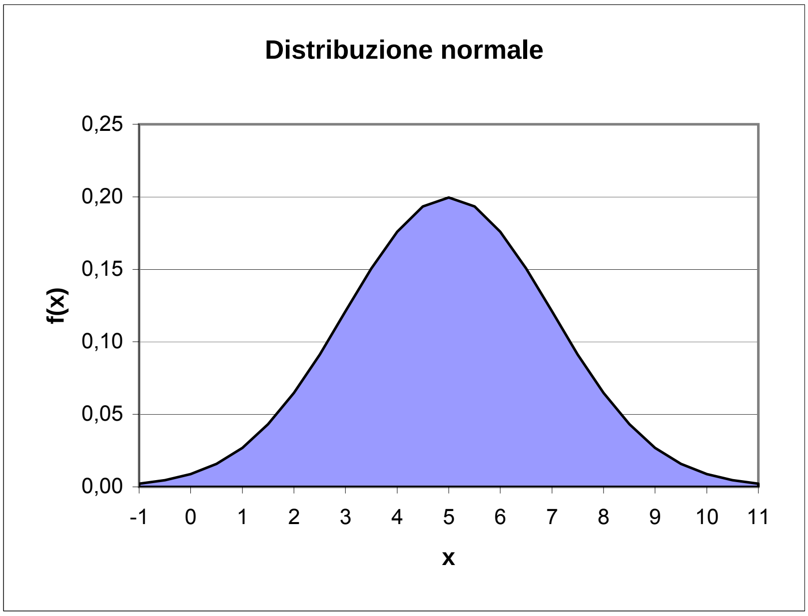




SUGGERIMENTI

Il grafico della distribuzione normale può anche essere realizzato colorando l'area sottesa dalla curva normale
Selezionare nella tabella 1 le celle delle ordinate f(x) (seconda colonna); premere il pulsante Creazione guidata grafico, scegliere Tipo di grafico>Area
Nella scheda Serie>Etichette Asse Categorie X: inserire le celle delle ascisse x (Tabella 1, prima colonna); concludere il grafico.

Fare doppio clic su Asse Categorie x; nella scheda Scala scegliere:
Numero di categorie tra le etichette di graduazione: 2
Numero di categorie tra i segni di graduazione: 2
Fare doppio cli su Asse valori y; nella scheda Scala scegliere:
Valore minimo   0
Valore massimo 0,25
Unità principale 0,05
Unità secondaria   0,05
Nella scheda Numero scegliere: Posizioni decimali 2



Torna su

Soluzione Esercizio 50

Confronto fra distribuzioni normali con parametri diversi

Indice

Ritorna Esercizio 50

Esercizio 50.1

1 Disegnare nello stesso grafico due distribuzioni normali aventi valor medio diverso e uguale varianza. Scegliere ad esempio come valori medi 3 e 5 e come varianza 4

2 Disegnare in un altro grafico due distribuzioni aventi la stessa media e due varianze diverse. Scegliere ad esempio come valore medio 3 e come varianze 4 e 9

SUGGERIMENTI

f1(x): distribuzione con valor medio = 3 e varianza = 4 (deviazione standard = 2).
f2(x): distribuzione con valor medio = 5 e varianza = 4 (deviazione standard = 2).
f3(x): distribuzione con valor medio = 3 e varianza = 9 (deviazione standard = 3).
Realizzare il grafico di f1(x) nell'intervallo (-3,9), con incremento = 0,5
il grafico di f2(x) nell'intervallo (-1,11), con incremento = 0,5
il grafico di f3(x) nell'intervallo (-6,12), con incremento = 0,5
La scelta degli intervalli dipende dai valori medi delle distribuzioni
e dalla rispettiva deviazione standard (vedi SUGGERIMENTI, esercizio 49).
Primo grafico:
Realizzare il grafico della prima distribuzione selezionando le celle contenenti i valori
di x e di f1(x) nella Tabella 1
Aggiungere il grafico della seconda distribuzione agendo sulla scheda Serie>Aggiungi:
per i valori di x e di f2(x) selezionare le celle corrispondenti nella Tabella 2
Secondo grafico:
Realizzare il grafico della prima distribuzione selezionando le celle della Tabella 1
Aggiungere il grafico della seconda distribuzione agendo sulla scheda Serie>Aggiungi:
per i valori di x e di f3(x) selezionare le celle corrispondenti nella Tabella 3

	f1(x)	f2(x)	f3(x)
valore medio	3	5	3
deviazione standard	2	2	3

Tabella 1

x	f1(x)
-3	0,0022
-2,5	0,0045
-2	0,0088
-1,5	0,0159
-1	0,0270
-0,5	0,0431
0	0,0648
0,5	0,0913
1	0,1210
1,5	0,1506
2	0,1760
2,5	0,1933
3	0,1995

Tabella 2

x	f2(x)
-1	0,0022
-0,5	0,0045
0	0,0088
0,5	0,0159
1	0,0270
1,5	0,0431
2	0,0648
2,5	0,0913
3	0,1210
3,5	0,1506
4	0,1760
4,5	0,1933
5	0,1995

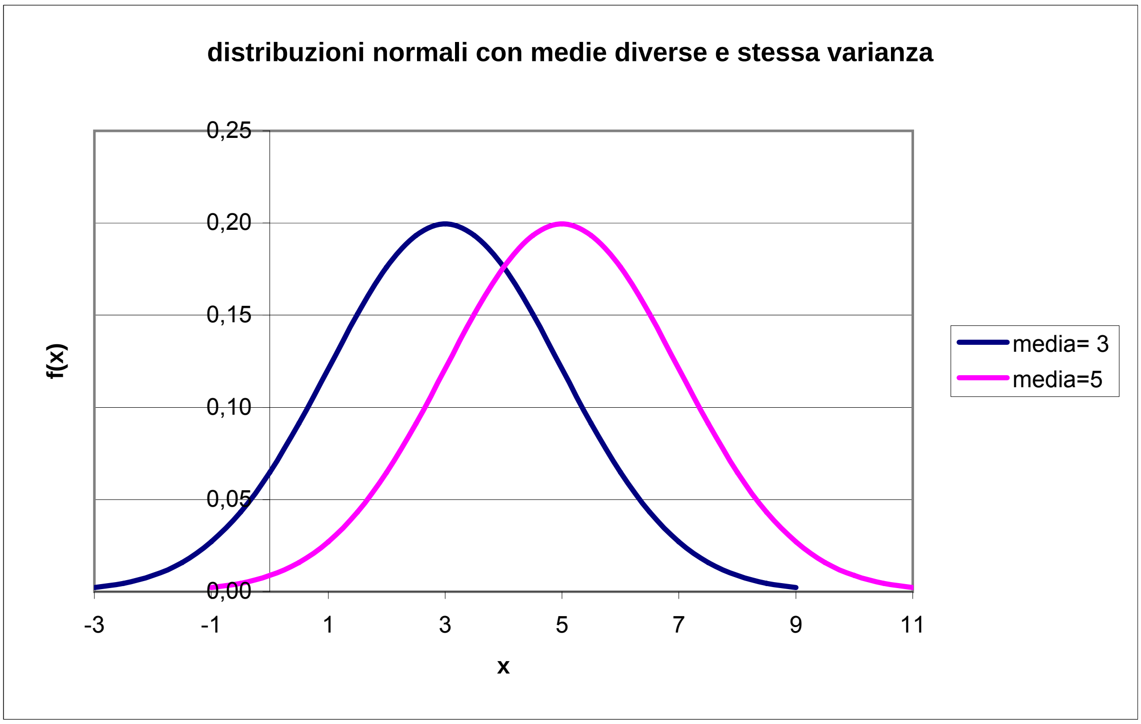
Tabella 3

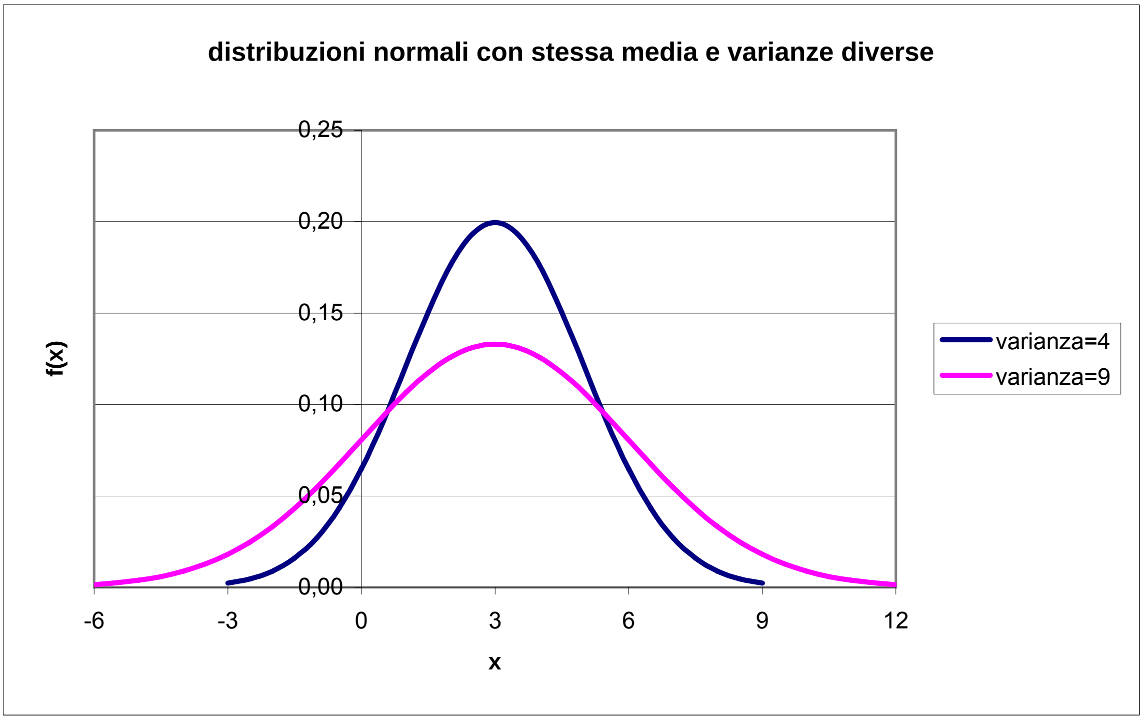
x	f3(x)
-6	0,0015
-5,5	0,0024
-5	0,0038
-4,5	0,0058
-4	0,0087
-3,5	0,0127
-3	0,0180
-2,5	0,0248
-2	0,0332
-1,5	0,0432
-1	0,0547
-0,5	0,0673
0	0,0807

3,5	0,1933
4	0,1760
4,5	0,1506
5	0,1210
5,5	0,0913
6	0,0648
6,5	0,0431
7	0,0270
7,5	0,0159
8	0,0088
8,5	0,0045
9	0,0022

5,5	0,1933
6	0,1760
6,5	0,1506
7	0,1210
7,5	0,0913
8	0,0648
8,5	0,0431
9	0,0270
9,5	0,0159
10	0,0088
10,5	0,0045
11	0,0022

0,5	0,0940
1	0,1065
1,5	0,1174
2	0,1258
2,5	0,1311
3	0,1330
3,5	0,1311
4	0,1258
4,5	0,1174
5	0,1065
5,5	0,0940
6	0,0807
6,5	0,0673
7	0,0547
7,5	0,0432
8	0,0332
8,5	0,0248
9	0,0180
9,5	0,0127
10	0,0087
10,5	0,0058
11	0,0038
11,5	0,0024
12	0,0015




distribuzioni normali con stessa media e varianze diverse
f(x)
0,25
0,20
0,15
0,10
0,05
0,00
-6
-3
0
3
6
9
12
x
varianza=4
varianza=9

Torna su

Soluzione Esercizio 51
Distribuzione normale e distribuzione normale standardizzata
Funzioni inverse

Indice

Ritorna Esercizio 51

Per il calcolo della funzione inversa della distribuzione normale si usano le funzioni
INV.NORM (inversa della distribuzione normale non standardizzata)
INV.NORM.ST (inversa della distribuzione normale standardizzata)

Data la variabile aleatoria X avente distribuzione normale, la funzione **INV.NORM** calcola il valore x tale che la probabilità P(X<x) assume un valore assegnato (probabilità della coda sinistra).
Data la variabile aleatoria Z avente distribuzione normale standardizzata, la funzione **INV.NORM.ST** calcola il valore z tale che la probabilità P(Z<z) assume un valore assegnato (probabilità della coda sinistra).

Sintassi
INV.NORM(probabilità;media;dev_standard)

Probabilità	probabilità assegnata (distribuzione normale).
Media	valor medio della distribuzione.
Dev_standard	scarto quadratico medio della distribuzione.

INV.NORM.ST(probabilità)

Probabilità	probabilità assegnata (distribuzione normale standardizzata).

Esempio 51.1

1 La variabile aleatoria X ha distribuzione normale con valor medio 19 e varianza 49

valor medio	19
scarto quadr. medio	7

Trovare il valore x tale che P(X<x)=0,8

x	24,89

=INV.NORM(0,8;F32;F33)

2 La variabile aleatoria X ha distribuzione normale con valor medio 19 e varianza 49

valor medio	19
scarto quadr. medio	7

Trovare il valore x tale che P(X>x)=0,3
(attenzione alla coda sinistra!)

x	22,67

=INV.NORM(0,7;F40;F41)

3 La variabile aleatoria X ha la distribuzione normale standardizzata.
Trovare il valore z tale che P(Z<z)=0,4

z	-0,25

=INV.NORM.ST(0,4)

Esercizio 51.2

La variabile aleatoria X ha distribuzione normale con media=19 e varianza=49.

valor medio	19
scarto quadr. medio	7

Trovare i valori x per cui

$P(X<x) = 0.9$	**x**	27,97
$P(X>x) = 0,65$	**x**	16,30
$P(0<X<x) = 0,42$	**x**	17,65

Attenzione: alla probabilità 0,42 bisogna aggiungere la probabilità P(X<0) che si calcola con la formula DISTRIB.NORM(0;F57;F58;VERO)

Esercizio 51.3

La variabile aleatoria Z ha distribuzione normale standardizzata.
Trovare i valori z per cui

$P(Z<z) = 0,9953$	**z**	2,597
$P(Z>z) = 0,2743$	**z**	0,600
$P(0<Z<z) = 0,3770$	**z**	1,160
$P(-z<Z<z) = 0,5762$	**z**	0,800

la probabilità P(Z<0) vale 0,5

Esercizio 51.4

La variabile aleatoria Z ha la distribuzione normale standardizzata.
Trovare i valori z per cui

$P(-z<Z<z) = 90\%$	**z**	1,645
$P(-z<Z<z) = 95\%$	**z**	1,960
$P(-z<Z<z) = 99\%$	**z**	2,576

Osservare che questi sono i valori tradizionalmente usati nella statistica inferenziale.

Torna su

Soluzione Esercizio 52

Approssimazione della distribuzione binomiale con la distribuzione normale

Indice

Ritorna Esercizio 52

Quando il numero di prove n è grande e la probabilità di successo p è prossima a 0,5 la distribuzione binomiale può essere approssimata con la distribuzione normale avente media=np e varianza=np(1–p).
REGOLA PRATICA per ottenere una buona approssimazione: usare la distribuzione normale quando np>=5 e n(1–p)>=5

Esercizio 52.1

Disegnare il grafico della distribuzione binomiale con n=12 e p=0,2.
Approssimare la distribuzione binomiale con la distribuzione normale e verificare che non si ottiene una buona approssimazione (la regola pratica non è rispettata)

SUGGERIMENTI

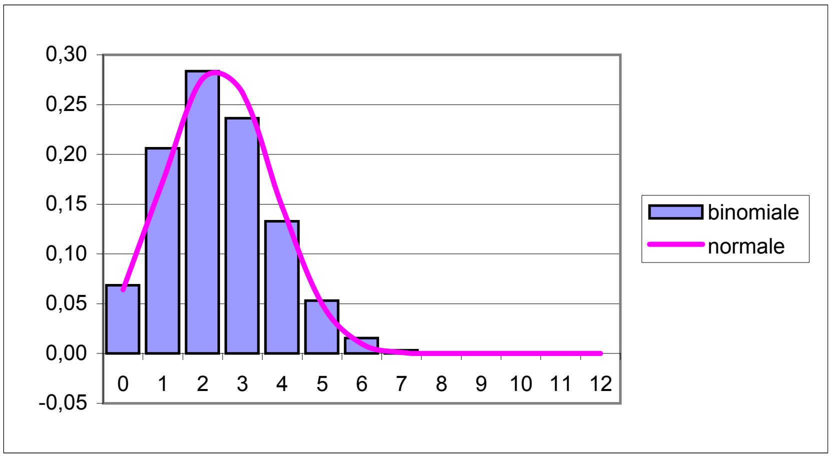
Per realizzare il grafico della normale occorre completare la colonna H della Tabella 1, calcolando i valori della distribuzione normale avente il valor medio e lo scarto quadratico medio indicati nelle celle D29 e D31

n	12
p	0,2

valor medio	2,4
varianza	1,92
scarto quadratico medio	1,39

Tabella 1

Successi	Probabilità P(X=x)	Distribuzione normale
0	0,0687	0,0642
1	0,2062	0,1728
2	0,2835	0,2762
3	0,2362	0,2621
4	0,1329	0,1478
5	0,0532	0,0495
6	0,0155	0,0099
7	0,0033	0,0012
8	0,0005	0,0001
9	0,0001	0,000003
10	0,000004	0,0000001
11	0,0000002	0,000000001
12	0,000000004	0,00000000001



Regola pratica: np>=5 e n(1-p)>=5

np	2,4
n(1-p)	9,6

La regola pratica non è rispettata: l'approssimazione non è buona

Esercizio 52.2

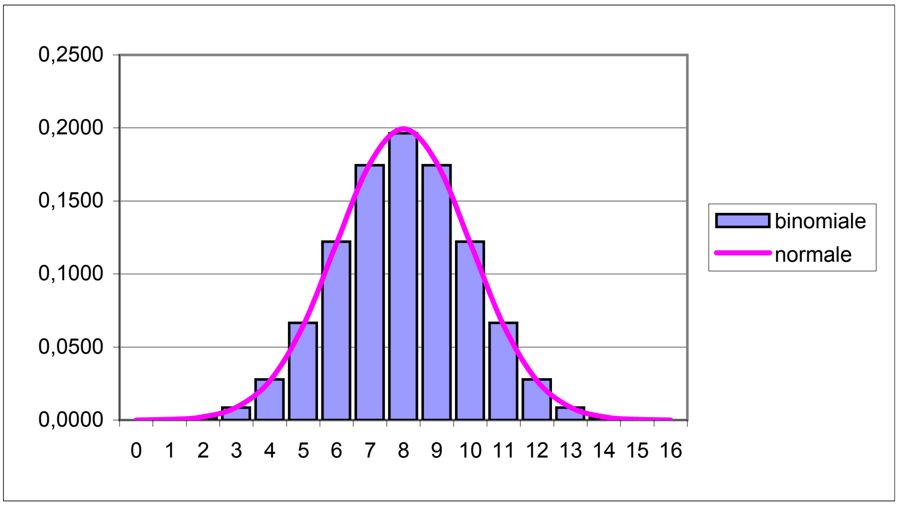
Disegnare il grafico della distribuzione binomiale con n=16 e p=0,5.
Approssimare la distribuzione binomiale con la distribuzione normale e verificare che si ottiene una buona approssimazione (la regola pratica è rispettata)

n	16
p	0,5

valor medio	8
varianza	4
scarto quadratico medio	2

Tabella 2

Successi	**Probabilità P(X=x)**	**Distribuzione normale**
0	0,0000	0,0001
1	0,0002	0,0004
2	0,0018	0,0022
3	0,0085	0,0088
4	0,0278	0,0270
5	0,0667	0,0648
6	0,1222	0,1210
7	0,1746	0,1760
8	0,1964	0,1995
9	0,1746	0,1760
10	0,1222	0,1210
11	0,0667	0,0648
12	0,0278	0,0270
13	0,0085	0,0088
14	0,0018	0,0022
15	0,0002	0,0004
16	0,00002	0,0001



Regola pratica: 8
np 8
n(1-p)
La regola pratica è rispettata: l'approssimazione è buona

Esercizio 52.3

Un test è composto da 20 domande con risposta Vero/Falso.

n	20

Rispondendo a caso alle domande, la probabilità p di rispondere correttamente (successo) é

p	0,5

1 Calcolare (Tabella 3) con la distribuzione binomiale le probabilità di rispondere correttamente a X=0, 1, 2, 3, …., 20 domande e disegnare l'istogramma.
La variabile aleatoria X rappresenta il numero di risposte esatte possibili (successi)

2 Dopo aver realizzato l'istogramma, aggiungere il grafico della distribuzione normale che approssima la binomiale

Il valor medio e la varianza della distribuzione normale che approssima la binomiale sono

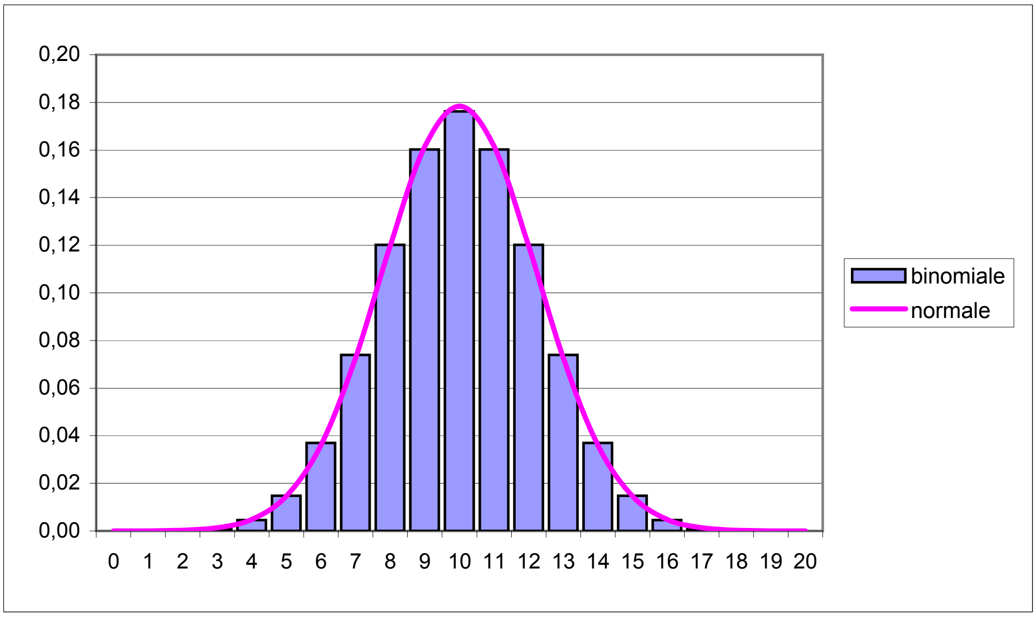
valor medio	10
varianza	5
scarto quadratico medio	2,236

SUGGERIMENTI

Per realizzare il grafico della normale occorre completare la Tabella 3 (distribuzione normale), calcolando (colonna H) i valori della distribuzione normale avente il valor medio e lo scarto quadratico medio sopra indicati, nei punti X=0, 1, 2, 3, …., 20 dell'intervallo (0,20)

Tabella 3

Successi	Probabilità P(X=x)	Distribuzione normale
0	0,000001	0,000008
1	0,00002	0,0001
2	0,0002	0,0003
3	0,0011	0,0013
4	0,0046	0,0049
5	0,0148	0,0146
6	0,0370	0,0360
7	0,0739	0,0725
8	0,1201	0,1196
9	0,1602	0,1614
10	0,1762	0,1784
11	0,1602	0,1614
12	0,1201	0,1196
13	0,0739	0,0725
14	0,0370	0,0360
15	0,0148	0,0146
16	0,0046	0,0049
17	0,0011	0,0013
18	0,00018	0,0003
19	0,000019	0,00005
20	0,000001	0,000008



3 Calcolare con la distribuzione binomiale la probabilità di rispondere esattamente ad almeno 12 domande

binomiale

P(X>=12)	0,25172

4 Calcolare la stessa probabilità usando l'approssimazione della binomiale con la distribuzione normale; usare la correzione di continuità

Il valor medio e la varianza della distribuzione normale che approssima la binomiale sono

valor medio	10
varianza	5
scarto quadratico medio	2,2361

normale

P(X>=11,5)	0,25117

L'approssimazione è molto buona, la regola pratica è rispettata.

Esercizio 52.4

Calcolare la probabilità che, in 10 lanci di una moneta, si presenti T un numero di volte compreso fra 3 e 6. Usare:

1 la distribuzione binomiale

2 la distribuzione normale con la correzione di continuità

Verificare che si ottiene una buona approssimazione (la regola pratica è rispettata)

n	10
p	0,5

binomiale

P(3<=X<=6)	0,7734

normale

valor medio della normale	5
varianza della normale	2,5
scarto quadratico medio	1,58

P(3<=X<=6)	0,7717

Esercizio 52.5

Calcolare la probabilità che, in 1000 lanci di una moneta, si presenti T un numero di volte

1 compreso fra 450 e 550.

2 uguale a 500

Usare la distribuzione normale con la correzione di continuità

Verificare che si ottiene una buona approssimazione (la regola pratica è rispettata)

n	1000
p	0,5

normale

valor medio della normale	500
varianza della normale	250
scarto quadratico medio	15,81

P(450<=X<=550)	0,9986

P(X=500)	0,0252

Torna su

Soluzione Esercizio 53

Approssimazione della distribuzione di Poisson con la distribuzione normale

Indice

Ritorna Esercizio 53

Al crescere del valor medio lambda, la distribuzione di Poisson può essere approssimata con una distribuzione normale avente media μ=lambda e varianza σ^2=lambda.
REGOLA PRATICA per ottenere una buona approssimazione: usare la distribuzione normale quando lambda>=10

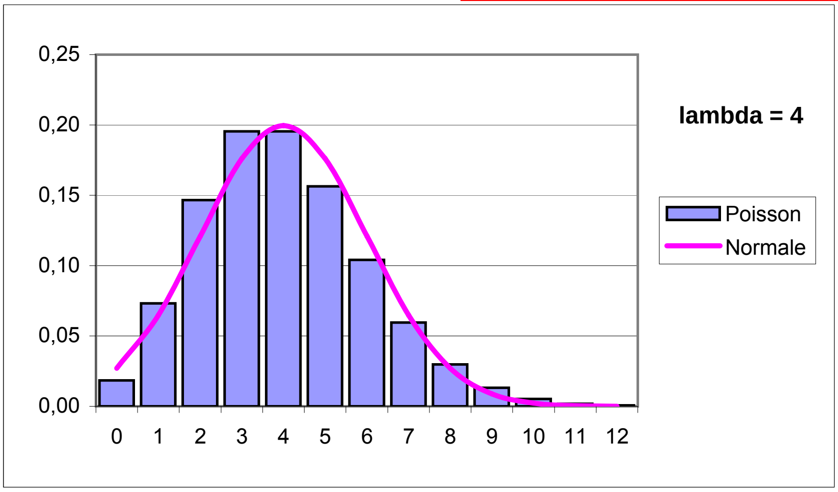
Esercizio 53.1
Disegnare il grafico della distribuzione di Poisson con lambda=4.
Approssimare la distribuzione di Poisson con la distribuzione normale e verificare che non si ottiene una buona approssimazione (la regola pratica non è rispettata)

SUGGERIMENTI
Per realizzare il grafico della normale occorre completare la colonna H della Tabella 1, calcolando i valori della distribuzione normale avente il valor medio e lo scarto quadratico medio indicati nelle celle D25 e D27

valor medio	4
varianza	4
scarto quadratico medio	2

Tabella 1

Eventi	Probabilità P(X=k)	Distribuzione normale
0	0,0183	0,0270
1	0,0733	0,0648
2	0,1465	0,1210
3	0,1954	0,1760
4	0,1954	0,1995
5	0,1563	0,1760
6	0,1042	0,1210
7	0,0595	0,0648
8	0,0298	0,0270
9	0,0132	0,0088
10	0,0053	0,0022
11	0,0019	0,0004
12	0,0006	0,0001

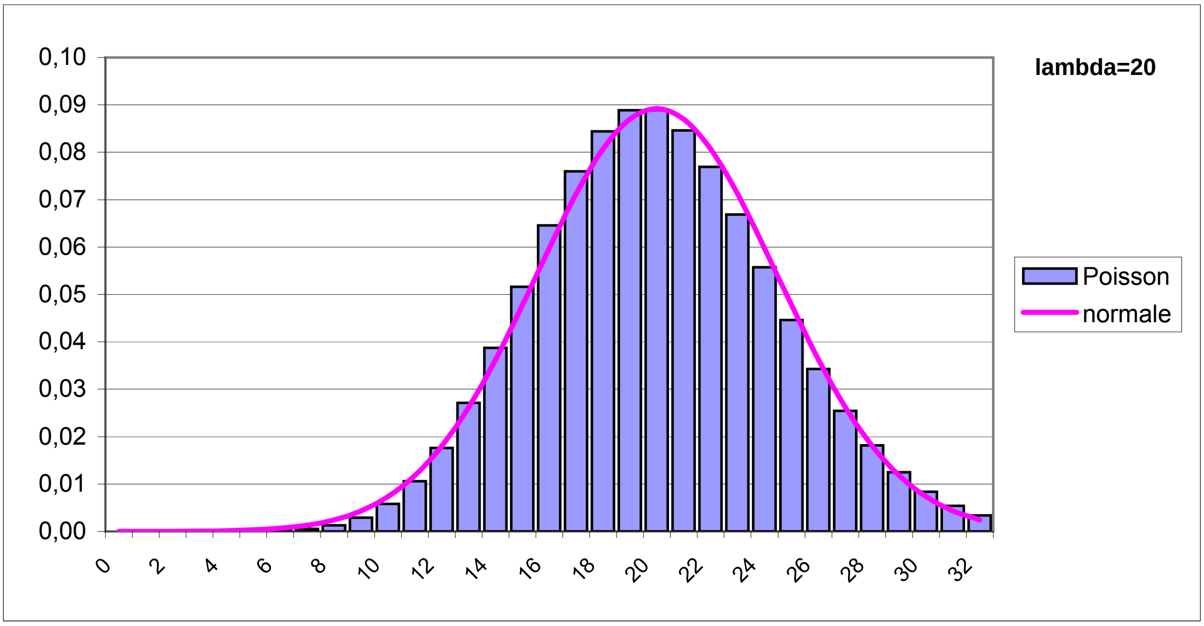


Esercizio 53.2
Disegnare il grafico della distribuzione di Poisson con lambda=20.
Approssimare la distribuzione di Poisson con la distribuzione normale e verificare che si ottiene una buona approssimazione (la regola pratica è rispettata)

valor medio	20
varianza	20
scarto quadratico medio	4,47

Tabella 2

Eventi	Probabilità P(X=k)	Distribuzione normale
0	0,000000002	0,000004
1	0,00000004	0,000011
2	0,0000004	0,000027
3	0,000003	0,000065
4	0,000014	0,0001
5	0,000055	0,0003
6	0,0002	0,0007
7	0,0005	0,0013
8	0,0013	0,0024
9	0,0029	0,0043
10	0,0058	0,0073
11	0,0106	0,0118
12	0,0176	0,0180
13	0,0271	0,0262
14	0,0387	0,0363
15	0,0516	0,0477
16	0,0646	0,0598
17	0,0760	0,0712
18	0,0844	0,0807
19	0,0888	0,0870
20	0,0888	0,0892
21	0,0846	0,0870
22	0,0769	0,0807
23	0,0669	0,0712
24	0,0557	0,0598
25	0,0446	0,0477
26	0,0343	0,0363
27	0,0254	0,0262
28	0,0181	0,0180
29	0,0125	0,0118
30	0,0083	0,0073
31	0,0054	0,0043
32	0,0034	0,0024



Esercizio 53.3

Il numero di errori di stampa su una pagina scelta a caso in un libro è una variabile distribuita secondo la distribuzione di Poisson con media lambda=0,4.

1 Calcolare la probabilità che il numero totale di errori nelle prime 10 pagine sia 3
2 Calcolare la probabilità che il numero totale di errori nelle prime 10 pagine sia maggiore di 3
3 Il libro ha 250 pagine. Se ci sono più di 110 errori la casa editrice corregge tali errori e ristampa il libro: calcolare la probabilità che questo accada usando l'approssimazione normale.

SUGGERIMENTI

Il numero medio di errori su 10 pagine è uguale a lambda*10
Il numero medio di errori su 250 pagine è uguale a lambda*250
Osservare che, se X indica la variabile di Poisson, si ha

P(X>110)=1–P(X<=110)

Usare la correzione di continuità.

lambda	0,4
valor medio	4

P(X=3)	0,1954

P(X>3)	0,5665

valor medio	100
scarto quadratico medio	10

P(X>110)	0,1469

Torna su

Soluzione Esercizio 54

Approssimazione di una distribuzione di frequenza con una distribuzione normale

Indice

Ritorna Esercizio 54

Esercizio 54.1

Nella tabella 1 è riportato un campione di misure (lunghezze in cm)
Raccogliere i dati in una distribuzione di frequenza assoluta usando come estremi destri delle classi i numeri indicati nella Tabella 2 (colonna C)
Approssimare la distribuzione di frequenza dei dati con una distribuzione normale (distribuzione teorica della popolazione), confrontando su un grafico i dati sperimentali e la distribuzione teorica.

Tabella 1

128	152	87	118	97	87
138	102	106	100	74	118
134	76	109	138	123	121
81	115	104	57	71	79
111	142	99	74	82	102
142	114	59	80	70	137
73	119	108	154	126	101
105	102	96	110	82	104
150	96	52	88	86	110
67	151	114	100	120	67
76	128	130	92	136	109
119	101	108	124	116	112
99	86	128	100	103	58
123	82	91	59	78	75
144	143	130	117	81	117
55	70	146	94	130	135
80	106	105	97	128	107
95	98	85	80	120	49
95	98	85	80	120	113
79	63	87	94	105	108

SUGGERIMENTI

Ricordare che la funzione **FREQUENZA** è una **funzione matrice**; occorre quindi selezionare tutte le celle in cui devono comparire le frequenze e premere poi Ctrl+Maiuscolo+Invio (vedi esercizio 23)
La prima classe e l'ultima risulteranno vuote e vengono aggiunte solo per migliorare l'aspetto del grafico della distribuzione normale
Calcolare le frequenze relative dividendo le frequenze assolute per il numero di dati
Poiché si devono approssimare i dati con una distribuzione continua e confrontare i dati sperimentali con la distribuzione teorica, bisogna costruire un istogramma in cui la somme delle aree dei rettangoli sia uguale a 1; a tale scopo si devono normalizzare le frequenze relative dividendo ogni frequenza relativa per l'ampiezza della classe corrispondente
Calcolare il valor medio, la varianza e la deviazione standard dei dati assegnati.
Calcolare il valore della distribuzione teorica (distribuzione normale), usando come valor medio e deviazione standard i valori calcolati
La distribuzione normale f(x) deve essere calcolata nei punti x usati come valori centrali delle classi
Realizzare un grafico in cui compaiono l'istogramma delle frequenze relative normalizzate e la distribuzione teorica

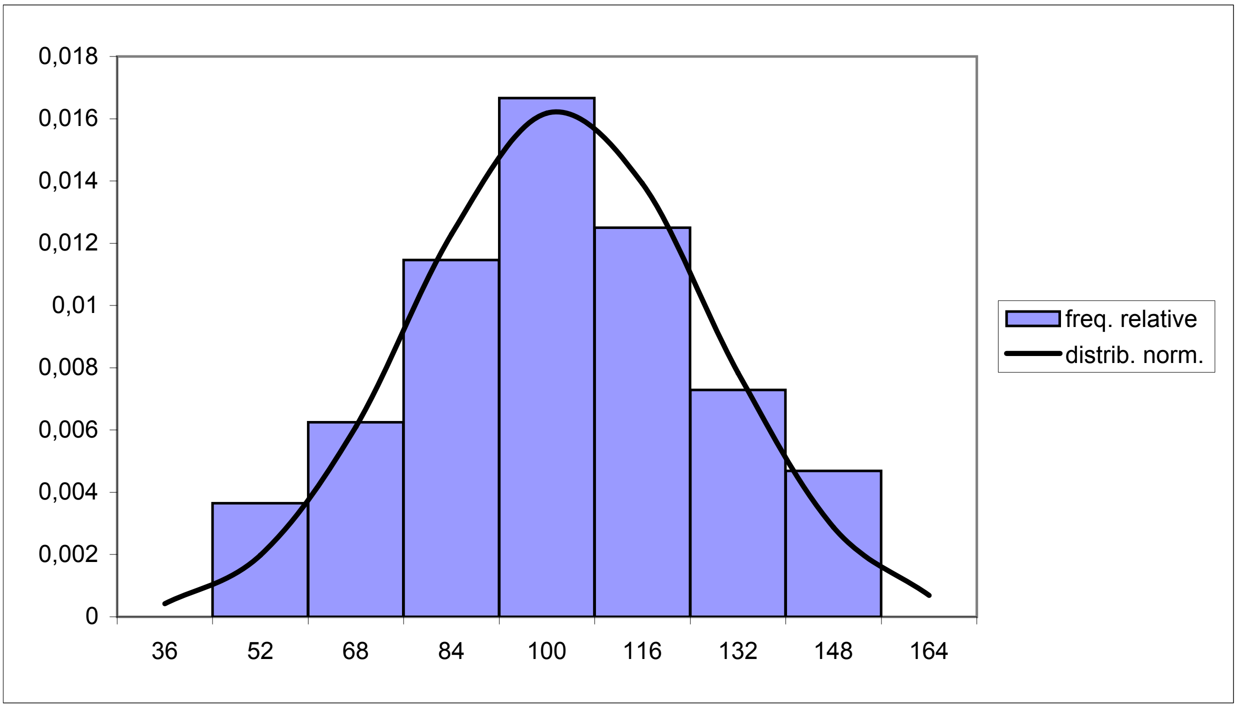
Realizzare prima l'istogramma delle frequenze relative normalizzate: nella scheda Serie, per Etichette Asse categorie X usare le celle dei valori centrali.
Nella scheda Serie aggiungere una nuova serie; come valori Y selezionare le celle della distribuzione normale; terminare il grafico che si presenta come un istogramma multiplo
Cliccare una volta sulle barre del secondo istogramma, premere il tasto destro, selezionare Tipo di grafico, modificare il tipo di grafico scegliendo Dispersione, linee continue
Modificare la larghezza delle barre dell'istogramma: cliccare una volta su una barra, con il tasto destro selezionare Formato Serie dati, Opzioni, Distanza tra le barre (distanza = 0)

Tabella 2

classi	estremi destri	valori centrali	frequenze assolute	frequenze relative	frequenze relative normalizzate	distribuzione normale
1	44	36	0	0	0	0,0004
2	60	52	7	0,0583	0,0036	0,0020
3	76	68	12	0,1	0,0063	0,0061
4	92	84	22	0,1833	0,0115	0,0123
5	108	100	32	0,2667	0,0167	0,0162
6	124	116	24	0,2	0,0125	0,0139
7	140	132	14	0,1167	0,0073	0,0078
8	156	148	9	0,075	0,0046875	0,0029
9	172	164	0	0	0	0,0007
Totali			120	1		

ampiezza delle classi	16

valor medio	102,333
varianza	602,594
scarto quadratico medio	24,548



Esercizio 54. 2

Nella tabella 3 è riportato un campione di misure

Costruire una distribuzione di frequenza assoluta; approssimare i dati con una distribuzione normale (distribuzione teorica della popolazione), confrontando su un grafico i dati sperimentali e la distribuzione teorica.

Tabella 3

68	73	61	66	80
84	79	65	78	78
75	88	75	82	89
82	73	87	75	61
86	60	94	94	75
90	93	62	77	95
59	71	95	69	60
88	59	78	74	79
76	85	63	68	83
81	75	78	60	71
79	87	86	61	66
62	80	67	65	78
59	80	73	75	82
97	57	81	87	75
78	88	72	74	82
85	78	63	62	77
76	62	76	95	69
87	76	75	78	74
71	53	85	63	68
75	74	96	72	60

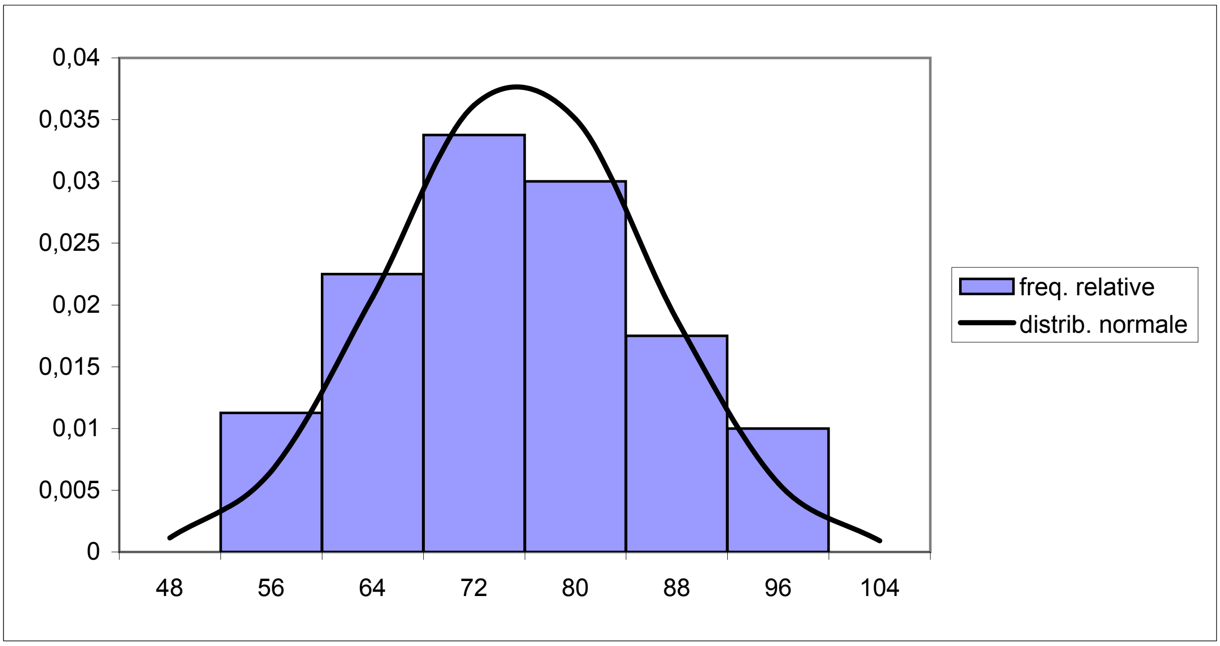
Come estremi destri delle classi si possono usare i valori indicati nella tabella 4 (colonna C)
Questa scelta non è l'unica possibile: per esercizio si può provare con altri valori,
scelti in modo che tutti i dati siano compresi nelle classi

Tabella 4

classi	estremi destri	valori centrali	frequenze assolute	frequenze relative	frequenze relative normalizzate	distribuzione normale
1	52	48	0	0	0	0,00114
2	60	56	9	0,09	0,01125	0,00651
3	68	64	18	0,18	0,0225	0,02064
4	76	72	27	0,27	0,03375	0,03616
5	84	80	24	0,24	0,03	0,03505
6	92	88	14	0,14	0,0175	0,01880
7	100	96	8	0,08	0,01	0,00558
8	108	104	0	0	0	0,00092
Totali			100	1		

ampiezza delle classi	8

valor medio	75,580
varianza	108,105
scarto quadratico medio	10,397


0,04
0,035
0,03
0,025
0,02
0,015
0,01
0,005
0
48
56
64
72
80
88
96
104
freq. relative
distrib. normale

Torna su

Soluzione Esercizio 55
Distribuzione t di Student

Indice

Ritorna Esercizio 55

Per il calcolo della **funzione di ripartizione t di Student** si usa la funzione **DISTRIB.T**
Con la funzione DISTRIB.T si può calcolare la probabilità P(X>x) che la variabile aleatoria. X avente la distribuzione di Student con un dato grado di libertà sia maggiore di un valore assegnato x (distribuzione a una coda), oppure la probabilità P(|X|>x)= P(X > x or X < –x) (distribuzione a due code)

Sintassi
DISTRIB.T(x;gradi_libertà;coda)

X	valore in cui si vuole calcolare la distribuzione; x deve essere positivo.
Gradi_libertà	grado di libertà della distribuzione t.
Coda	specifica il numero di code di distribuzione da restituire. Se coda = 1, DISTRIB.T restituisce la distribuzione a una coda (destra). Se coda = 2, DISTRIB.T restituisce la distribuzione a due code.

Esempio 55.1
Calcolare la probabilità che la v. a. X avente distribuzione di Student con grado di libertà 9 assuma valori maggiori di 1,833
Scegliere Code = 1

=DISTRIB.T(1,833;9;1)

P(X>1,833)	0,0500

Questo risultato significa che l'area sottesa dalla distribuzione a destra di 1,833 vale 0,05

Scegliere Code = 2

=DISTRIB.T(1,833;9;2)

P(|X|>1,833)	0,1000

Questo risultato significa che la somma delle aree a destra di 1,833 e a sinistra di –1,833 vale 0,1 (due code di uguale area)

Per ogni esercizio spiegare qual è il significato del risultato trovato

Esercizio 55.2
Trovare la probabilità che la variabile aleatoria X avente la distribuzione t di Student con grado di libertà 9 assuma valori maggiori di 2,262

1 coda	**P(X>2,262)**	0,025

Questo risultato significa che l'area a destra del valore 2,262 vale 0,025

2 code	**P(|X|>2,262)**	0,050

Questo risultato significa che la somma dell'area a destra di 2,262 e a sinistra di –2,262 vale 0,050

Esercizio 55.3
Trovare la probabilità che la variabile aleatoria X avente la distribuzione t di Student con grado di libertà 9 assuma valori maggiori di –2,821
Nota: il valore del parametro x deve essere positivo, quindi
P(X>–2,821)=P(X<2,821)=1–P(X>2,281)

1 coda	**P(X>-2,821)**	0,990

Questo risultato significa che l'area a destra di –2,821 vale 0,99 (quest'area è uguale all'area a sinistra di 2,821)

Esercizio 55.4

Trovare la probabilità che la variabile aleatoria X avente la distribuzione di Student con grado di libertà 27 assuma valori maggiori di 1,703

1 coda	**P(X>1,703)**	0,050

Questo risultato significa che l'area a destra di 1,703 vale 0,050

2 code	**P(|X|>1,703)**	0,100

Questo risultato significa che l'area a destra di 1,703 e a sinistra di –1,703 vale 0,100

Funzione **INV.T** (inversa della distribuzione di Student)

La funzione **INV.T** calcola l'inversa della distribuzione t di Student per il grado di libertà specificato. **Usa due code**

Sintassi

INV.T(probabilità;gradi_libertà)

Probabilità probabilità associata alla distribuzione t di Student a due code.

Gradi_libertà grado di libertà della distribuzione t.

La funzione INV.T restituisce sempre un valore positivo

Esempio 55.5

Data la distribuzione t con grado di libertà 9, trovare il valore x tale che l'area delle **due code** (a destra di x e a sinistra di –x) sia uguale a 0,05

x	2,262

=INV.T(0,05;9)

Esercizio 55.6

Data la distribuzione t con grado di libertà 9, trovare il valore x tale che l'area della coda a destra di x è uguale a 0,05

x	1,833

Attenzione: la funzione INV.T lavora su due code: l'area delle due code è 0,1

Esercizio 55.7

Data la distribuzione t con grado di libertà 10, trovare il valore x tale che l'area compresa fra –x e x è uguale a 0,90

In questo caso l'area delle due code vale 0,10

x	1,812

Esercizio 55.8

Data la distribuzione t con grado di libertà 9, trovare il valore x tale che l'area a destra di x è uguale a 0,99

In questo caso l'area delle due code vale 0,02

x	-2,821

Torna su

Soluzione Esercizio 56
Distribuzione chi quadro

Indice

Ritorna Esercizio 56

Per il calcolo della **funzione di ripartizione chi quadro** si usa la funzione **DISTRIB.CHI**
La funzione DISTRIB.CHI calcola la probabilità P(X>x) che la variabile aleatoria X che ha la distribuzione chi quadro con un dato grado di libertà sia maggiore di un valore assegnato x

Sintassi
DISTRIB.CHI(x;gradi_libertà)

x	valore in cui si vuole calcolare la distribuzione
Gradi_libertà	grado di libertà della distribuzione t

Usa **una sola coda (destra)**

Esempio 56.1
Calcolare la probabilità che una variabile aleatoria avente distribuzione chi quadro con grado di libertà 15 assuma valori maggiori di 25

=DISTRIB.CHI(25;15)

P(X>25)	0,050

Questo risultato significa che l'area a destra di 25 è uguale a 0,050

Esercizio 56.2
Calcolare la probabilità che una variabile aleatoria avente distribuzione chi quadro con grado di libertà 5 assuma valori minori di 1,145

P(X<1,145)	0,050

Esercizio 56.3
Calcolare la probabilità che una variabile aleatoria X avente la distribuzione chi quadro con grado di libertà 5 assuma valori compresi fra 0,831 e 12,832

P(0,831<X<12,832)	0,950

Funzione **INV.CHI** (inversa della distribuzione chi quadro **a una coda**)
La funzione INV.CHI calcola l'inversa della distribuzione chi quadro (**a una coda, destra)**
per il grado di libertà specificato.

Sintassi
INV.CHI(probabilità;gradi_libertà)

Probabilità	probabilità associata alla distribuzione chi quadro.
Gradi_libertà	grado di libertà della distribuzione t

Esempio 56.4
Data la variabile aleatoria X avente la distribuzione chi quadro con grado di libertà 5, trovare il valore x tale che l'area a destra di x è uguale a 0,05

=INV.CHI(0,05;5)

x	11,070

Esercizio 56.5
Data la variabile aleatoria X avente la distribuzione chi quadro con grado di libertà 5, trovare il valore x tale che l'area a sinistra di x vale 0,05

x	1,145

Esercizio 56.6

Data la variabile aleatoria X avente la distribuzione chi quadro con grado di libertà 5, trovare i valori x1 e x2 tali che la somma delle aree delle due code (a sinistra di x1 e a destra di x2) sia uguale a 0,05 (code di uguale area)

x1	0,831
x2	12,833

Torna su

Soluzione Esercizio 57
Distribuzione F di Fisher

Indice

Ritorna Esercizio 57

Per il calcolo della **funzione di ripartizione F** si usa la funzione **DISTRIB.F**
La funzione **DISTRIB.F** calcola la probabilità P(X>x) che la variabile aleatoria X che ha la distribuzione F con dati gradi di libertà di numeratore e denominatore sia maggiore di un valore assegnato x

Sintassi
DISTRIB.F(x;gradi_libertà1;gradi_libertà2)

x	valore in cui si vuole calcolare la distribuzione (x>0).
Gradi_libertà1	grado di libertà del numeratore.
Gradi_libertà2	gradio di libertà del denominatore.

Usa una sola coda (destra)

Esempi0 57.1
Calcolare la probabilità che una variabile aleatoria avente distribuzione F con gradi di libertà rispettivamente 15 e 25 assuma valori maggiori di 1,77

=DISTRIB.F(1,77;15;25)

P(X>1,77)	0,100

Calcolare la probabilità che una variabile aleatoria X avente la distribuzione F con gradi di libertà rispettivamente 5 e 10 assuma valori minori di 3,326

=1-DISTRIB.F(3,326;5;10)

P(X<3,326)	0,950

Esercizio 57.2
Calcolare la probabilità che una variabile aleatoria X avente la distribuzione F con gradi di libertà rispettivamente 10 e 25 assuma valori maggiori di 1,865

P(X>1,865)	0,100

Esercizio 57.3
Calcolare la probabilità che una variabile aleatoria avente la distribuzione F con gradi di libertà rispettivamente 10 e 25 assuma valori minori di 2,24

P(X<2,24)	0,950

Esercizio 57.4
Calcolare la probabilità che una variabile aleatoria avente distribuzione F con gradi di libertà rispettivamente 10 e 20 assuma valori maggiori di 0,455

P(X>0,455)	0,900

Funzione **INV.F** (inversa della distribuzione F a una coda)
La funzione **INV.F** calcola l'inversa della distribuzione F (a una coda) per i gradi di libertà specificati. **Usa una sola coda (destra)**

Sintassi
INV.F(probabilità;gradi_libertà1;gradi_libertà2)

Probabilità	è la probabilità associata alla distribuzione chi quadro.
Gradi_libertà1	grado di libertà del numeratore.
Gradi_libertà2	grado di libertà del denominatore.

Esempio 57.5

Data la variabile aleatoria X avente la distribuzione F con gradi di libertà 15 e 25, trovare il valore x tale che l'area a destra di x è uguale a 0,05

x	2,089

=INV.F(0,05;15;25)

Esercizio 57.6

Data la variabile aleatoria X avente la distribuzione F con gradi di libertà 10 e 20, trovare il valore di x tale che l'area a destra di x vale 0,90

x	0,454

Esercizio 57.7

Data la variabile aleatoria X avente la distribuzione F con gradi di libertà 15 e 25, trovare il valore di x tale che l'area a destra di x vale 0,10

x	1,771

Esercizio 57.8

Data la variabile aleatoria X avente la distribuzione F con gradi di libertà 15 e 10, trovare i valori x1 e x2 tali che l'area compresa fra essi vale 0,90

x1	0,393
x2	2,845

Torna su

Soluzione Esercizio 58

Generazione di numeri casuali. Campionamento

Strumenti Analisi Dati: Generazione di un numero casuale, Campionamento

Indice

Ritorna Esercizio 58

Excel può essere usato come generatore di numeri casuali, utili ad esempio per simulare esperimenti.
Lo Strumento di Analisi **Generazione di un numero casuale** genera campioni di numeri casuali distribuiti secondo una distribuzione di probabilità da scegliere fra più distribuzioni disponibili.
Le distribuzioni disponibili sono:
Uniforme, Normale, Bernoulli, Binomiale, Poisson, Discreta
(La distribuzione del tipo A schema, che qui non sarà usata, genera campioni di valori non casuali)

Esempio 58.1

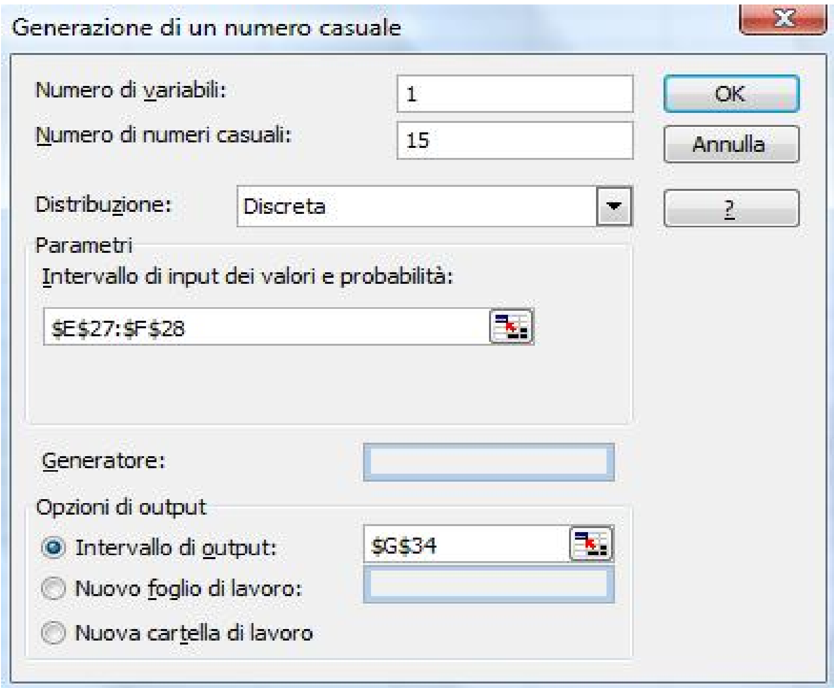
Distribuzione discreta

Una distribuzione di probabilità discreta è specificata da un elenco di valori possibili e dalla probabilità associata a ciascun valore.

Simulazione dei risultati del lancio di una moneta
Excel restituisce dati numerici, quindi si assegna 1 per "testa" e 0 per "croce"
Entrambe le facce della moneta hanno la stessa probabilità di presentarsi, e tale probabilità è 0,5

	valori	probabilità
testa	1	0,5
croce	0	0,5

Dal Menu **Strumenti Analisi Dati** selezionare **Generazione di un numero casuale** e cliccare su OK
Si apre la seguente finestra di dialogo



Numeri casuali
1
1
0
1
0
1
0
1
1
1
0
1
1
1
1

Nella casella **Numero di variabili** scegliere il numero di campioni da generare (in questo esempio 1)
Nella casella **Numero di numeri casuali** scegliere l'ampiezza del campione
(se si genera più di un campione, tutti i campioni hanno la stessa ampiezza)
Scegliere il tipo di distribuzione nel menu **Distribuzione** (in questo esempio Discreta)
Nella casella dei **Parametri** selezionare le celle dei valori e delle rispettive probabilità.
Nelle O**pzioni di output** scegliere la posizione della tabella di output:
in **Intervallo di output**, selezionare la cella in cui sarà collocato il primo elemento del campione
Se si generano più campioni, gli elementi di ciascun campione vengono elencati in colonne adiacenti
Quando si lancia per la prima volta lo strumento Generazione di un numero casuale, nella casella **Generatore** si può scegliere il primo numero causale del campione, in modo da poter rigenerare successivamente lo stesso campione; per cambiare/eliminare questa scelta in modo da generare numeri diversi occorre riavviare Excel.
NOTA: eseguendo gli esercizi, i dati dei campioni possono essere diversi da quelli riportati in questa soluzione perché sono casuali

Esercizio 58.2

Simulare le vendite mensili di un prodotto in un anno, supponendo che i possibili valori delle vendite (numero di esemplari venduti) e le corrispondenti probabilità siano quelli della tabella 1 seguente
Generare tre campioni

Tabella 1

n° esemplari venduti	probabilità
500	0,2
600	0,4
800	0,3
1000	0,1

SUGGERIMENTI

Nella casella **Numero di variabili** inserire 3
Nella casella **Numero di numeri casuali** inserire 12 (il numero dei mesi)

	Vendite		
Mese	**1° campione**	**2° campione**	**3° campione**
Gen	500	500	600
Feb	600	800	600
Mar	600	800	800
Apr	600	600	1000
Mag	1000	600	800
Giu	500	600	600
Lug	500	800	600
Ago	1000	1000	600
Set	800	1000	500
Ott	800	500	800
Nov	600	600	500
Dic	600	800	800

Esercizio 58.3

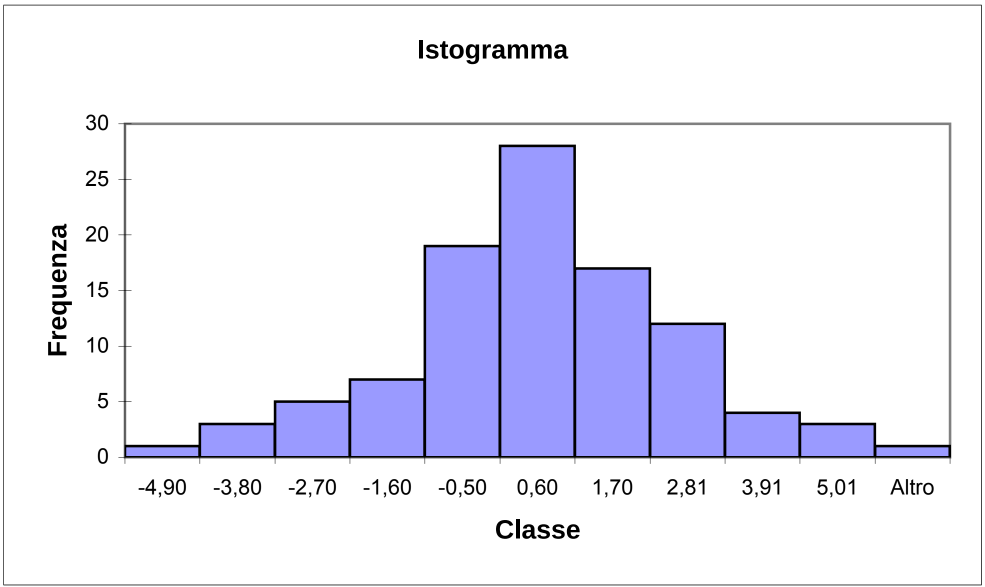
Distribuzione normale

Con questa opzione si generano numeri casuali distribuiti secondo la distribuzione normale con media e deviazione standard specificate nelle caselle dei **Parametri**
Per generare numeri aventi la distribuzione normale standardizzata scegliere media = 0 e deviazione standard = 1

Generare un campione di 100 numeri casuali con distribuzione normale, media = 0, deviazione standard = 2
Con lo Strumento Istogramma raccogliere i dati in classi e disegnare un istogramma per mostrare che i dati del campione hanno (approssimativamente) la distribuzione normale

campione
-1,2686
-0,3902
2,1408
0,3473
0,5491
1,7256
-1,2495
-1,4973
-0,9637
0,7804
1,6818
-0,0053
0,5909
-1,7679
-0,0172
0,5501
-0,8719
4,3062
-3,0038
0,9079
2,0724
-2,2844
4,3738
0,3085
-1,5700
-2,7123
0,0980
2,8532
0,5135
-0,3194
-1,6014
2,2836
3,0968
-3,1867
2,2012
-1,9811
-4,5546
-4,5775
0,2327
-1,1842
1,4067
1,2958
0,7998
-0,5765
-2,4986
1,1362
0,3568
0,0070
-2,9410
-1,5830
2,8587

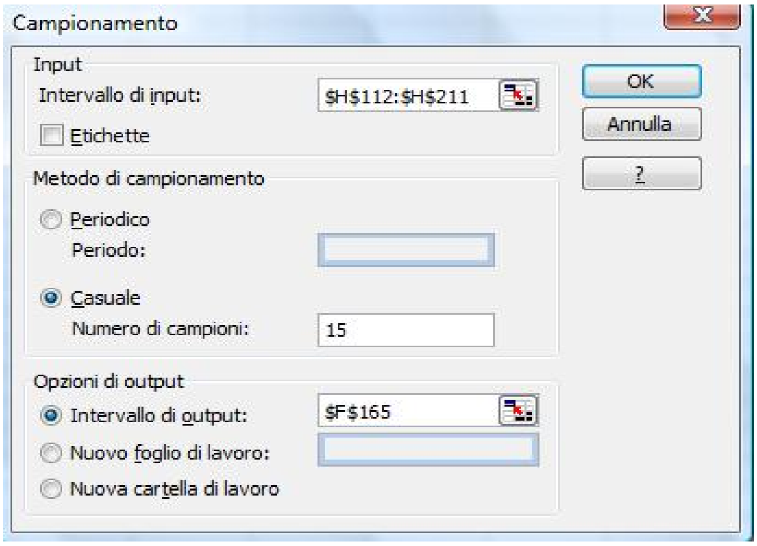
Classe	*Frequenza*
-4,90	1
-3,80	3
-2,70	5
-1,60	7
-0,50	19
0,60	28
1,70	17
2,81	12
3,91	4
5,01	3
Altro	1



Esempio 58.4
Generazione di campioni da un insieme di dati
Lo strumento di analisi **Campionamento** seleziona un campione casuale da una data popolazione di valori; il campionamento è effettuato con ripetizione

Selezionare un campione di 15 valori dalla popolazione di 100 valori generata nell'esercizio precedente

Nella casella **Intervallo di input** selezionare i valori della popolazione
Come **Metodo di campionamento** selezionare casuale
Nella casella **Numero di campioni** indicare il numero di elementi del campione



campione
-0,2144
0,3568
-1,8957
-1,9811
2,3318
2,5588
0,2327
1,3218
2,2012
1,6818
0,7998
-0,9637
1,5287
-0,8719
-0,5899

1,6230
-0,2287
-0,4827
-4,2273
-0,2628
2,3318
-0,8189
-0,3465
6,1080
-0,2144
-2,9222
-0,8499
-0,0518
2,4744
-0,0311
0,2550
-0,3678
-1,3246
-2,4058
2,3873
2,5588
-1,3353
-0,6033
1,4663
1,5287
-1,2934
1,2232
1,3218
-0,5106
2,3075
-4,9001
1,7763
0,7650
-0,9229
-0,3322
-1,8957
1,9513
-0,3575
1,3497
-0,4555
-1,2019
0,0706
1,0992
0,7433
4,1387
2,8953
0,3661
0,8307
-0,5899

Esercizio 58.5

Selezionare un campione casuale di 5 clienti da un elenco di 20 clienti; individuare i clienti numerandoli da 1 a 20.

clienti
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20

campione
3
3
5
14
1

Notare che il campione è ottenuto con ripetizione: il cliente 3 compare due volte.

Torna su

6. STIMA DEI PARAMETRI

Soluzione Esercizio 59
Intervallo di confidenza per la media
(varianza della popolazione nota - grandi campioni)

Indice

Ritorna Esercizio 59

Dato un campione di ampiezza n (n≥30) estratto da una popolazione con varianza nota σ^2 , l'intervallo di confidenza per la media μ della popolazione con grado di fiducia $(1 - \alpha)$*100% è dato dalla formula

$$\boxed{\bar{x} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}}$$

$\bar{x}$ valor medio del campione

$z_{\frac{\alpha}{2}}$ valore critico della distribuzione normale

L'ampiezza dell'intervallo di confidenza è uguale a $\boxed{2\, z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}}$

La funzione CONFIDENZA calcola l'ampiezza della metà dell'intervallo di confidenza

Sintassi
CONFIDENZA(**alfa,dev_standard,dimensioni**)
Alfa livello di significatività
Nota: il grado di fiducia è (1 – alfa)*100%; ad esempio per il grado di fiducia del 95% si ha 1-alfa=0,95 , quindi alfa = 0,05
Dev_standard deviazione standard (o scarto quadratico medio) della popolazione
Dimensioni numero di elementi del campione.
La funzione CONFIDENZA deve essere usata solo per grandi campioni

Esempio 59.1
Sia dato un campione di ampiezza n = 100 estratto da una popolazione avente deviazione standard σ=5,1; il valor medio del campione sia 21,6.
Trovare l'intervallo di confidenza con grado di fiducia del 95% per la media della popolazione.

Numero elementi del campione	100
media campione	21,6
deviazione standard popolazione	5,1
grado di fiducia 95%	0,95

ampiezza metà intervallo	1,00

Intervallo di confidenza per la media

estremo sinistro = media – metà intervallo	20,60
estremo destro = media + metà intervallo	22,60

Esercizio 59.2
Da un'indagine svolta su un campione di 200 giovani risulta che i ragazzi del Nord America dedicano alla televisione un numero medio di 26 ore settimanali
La deviazione standard della popolazione (calcolata da una precedente indagine) è σ=7 ore.
Trovare l'intervallo di confidenza con grado di fiducia del 99% per il numero medio di ore dedicate

alla televisione dall'intera popolazione dei ragazzi americani.

Numero dati del campione	200
media campione	26
deviazione standard popolazione	7
grado di fiducia 99%	0,99
Metà intervallo	1,275

Intervallo di confidenza per la media

estremo sinistro = media – metà intervallo	24,725
estremo destro = media + metà intervallo	27,275

Esercizio 59.3

Sia dato un campione di 100 studenti tratto da una popolazione di studenti di sesso maschile iscritti all'università

Trovare un intervallo di confidenza per il peso medio della popolazione da cui è tratto il campione di studenti, sapendo che il peso medio degli studenti del campione è 67,45 Kg e la varianza della popolazione da cui è tratto il campione è σ^2=9 Kg^2

Numero elementi del campione	100
media campione	67,45
deviazione standard popolazione	3
grado di fiducia 95%	0,95
ampiezza metà intervallo	0,59

Intervallo di confidenza per la media

estremo sinistro	66,86
estremo destro	68,04

Esercizio 59.2

Si vuole stimare il numero medio di battiti cardiaci al minuto per una data popolazione.
Il numero medio di battiti al minuto per un campione di 50 individui è uguale a 90.
La popolazione è distribuita normalmente con una deviazione standard σ=10.
Trovare gli intervalli di confidenza per la media della popolazione con i gradi di fiducia del 90%, 95%, 98% e 99%.

SUGGERIMENTI

Sfruttare l'aggiornamento automatico di Excel: cambiando il valore del grado di fiducia devono automaticamente aggiornarsi sia l'ampiezza della metà dell'intervallo, che gli estremi dell'intervallo.
Nella soluzione dell'esercizio, per comodità si riportano tutti i risultati relativi ai vari gradi di fiducia
Osservare che, restando invariata l'ampiezza del campione (n=50), all'aumentare del grado di fiducia, cresce l'ampiezza dell'intervallo di confidenza, ossia la stima della media della popolazione è meno precisa.

Numero dati del campione	50
media campione	90
deviazione standard popolazione	10

Grado di fiducia 90 %

grado di fiducia 90%	0,9
ampiezza metà intervallo	2,33

Intervallo di confidenza per la media (grado di fiducia 90%)

estremo sinistro = media – metà intervallo	87,67
estremo destro = media + metà intervallo	92,33

Grado di fiducia 95%

grado di fiducia 95%	0,95
ampiezza metà intervallo	2,77

Intervallo di confidenza per la media (grado di fiducia 95%)

estremo sinistro = media – metà intervallo	87,23
estremo destro = media + metà intervallo	92,77

Grado di fiducia 98%

grado di fiducia 98%	0,98
ampiezza metà intervallo	3,29

Intervallo di confidenza per la media (grado di fiducia 98%)

estremo sinistro = media – metà intervallo	86,71
estremo destro = media + metà intervallo	93,29

Grado di fiducia 99%

grado di fiducia 99%	0,99
ampiezza metà intervallo	3,64

Intervallo di confidenza per la media (grado di fiducia 99%)

estremo sinistro = media – metà intervallo	86,36
estremo destro = media + metà intervallo	93,64

Torna su

Soluzione Esercizio 60

Distribuzione di frequenza - Intervallo di confidenza per la media (varianza della popolazione incognita - grandi campioni)

Indice

Ritorna Esercizio 60

Se la varianza della popolazione non è nota, per grandi campioni può essere sostituita con la varianza del campione. Con questa sostituzione si commette un errore di approssimazione.

Esercizio 60.1

Sia dato un campione di 100 studenti tratto da una popolazione di studenti di sesso maschile iscritti al primo anno di università. La tabella 1 contiene i pesi in Kg degli studenti

Tabella 1

66	67	69	66	62
62	70	63	70	68
68	65	73	74	64
68	75	73	70	72
64	72	65	72	65
72	68	72	73	69
72	71	72	72	68
67	68	62	63	74
69	67	63	67	60
68	65	69	67	67
75	68	69	70	69
70	72	67	70	68
69	71	66	71	65
73	61	66	64	66
70	72	62	68	69
65	71	67	68	64
69	70	68	64	70
64	70	72	65	69
67	68	73	71	65
67	70	66	67	67

Trovare gli intervalli di confidenza al 95% e al 99% per il peso medio di tutti gli studenti.
La varianza della popolazione non è nota e può essere sostituita con la varianza del campione, perché il campione è grande

Numero dati del campione	100
media campione	68,16
deviazione standard campione	3,34

grado di fiducia 95%	0,95
ampiezza metà intervallo	0,65

Intervallo di confidenza per la media (grado di fiducia 95%)

estremo sinistro	67,51
estremo destro	68,81

grado di fiducia 99%	0,99
ampiezza metà intervallo	0,86

Intervallo di confidenza per la media (grado di fiducia 99%)

estremo sinistro	67,30
estremo destro	69,02

Esercizio 60.2

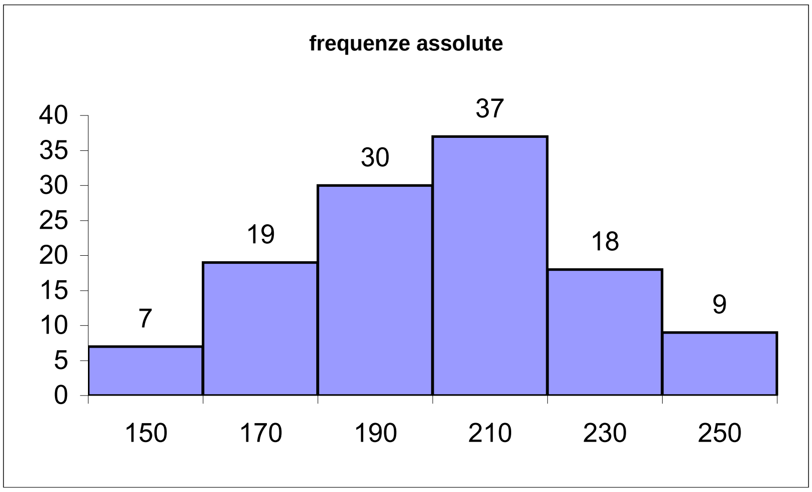
Sono assegnati i dati della tabella 2
Costruire una distribuzione di frequenza assoluta, raggruppando i dati in sei classi
Disegnare l'istogramma della distribuzione di frequenza assoluta
Trovare l'intervallo di confidenza per la media con grado di fiducia del 95%.

Tabella 2

228	252	187	218	197
238	202	206	200	174
234	176	209	238	223
181	215	204	157	171
211	242	199	174	182
242	214	159	180	170
173	219	208	254	226
205	202	196	210	182
250	196	152	188	186
167	251	214	200	220
176	228	230	192	236
219	201	208	224	216
199	186	228	200	203
223	182	191	159	178
244	243	230	217	181
155	170	246	194	230
180	206	205	197	228
195	198	185	180	220
195	198	185	180	220
179	163	187	194	205
187	202	210	158	207
218	237	167	175	149
221	201	209	217	213
179	204	212	235	208

numero dati	120
minimo	149
massimo	254
range	105
numero classi	6
ampiezza classi	20

classi	**estremo destro**	**frequenza assoluta**	**valori centrali**
140<x<=160	160	7	150
160<x<=180	180	19	170
180<x<=200	200	30	190
200<x<=220	220	37	210
220<x<=240	240	18	230
240<x<=260	260	9	250



Intervallo di confidenza per la media

grado di fiducia	0,95
media	202,33
deviazione standard	24,55
ampiezza metà intervallo	4,39
estremo sinistro	197,94
estremo destro	206,73

Esercizio 60.3

Data la distribuzione di frequenza della tabella 3, trovare l'intervallo di confidenza per la media della popolazione con grado di fiducia del 95%

Tabella 3

classi	frequenze assolute
6<x<=10	3
10<x<=14	15
14<x<=18	24
18<x<=22	34
22<x<=26	18
26<x<=30	6

La varianza della popolazione non è nota e si sostituisce con la varianza del campione.
Poiché non è nota la tabella dei dati grezzi, ma solo la distribuzione di frequenza, occorre calcolare media e varianza con le formule per i dati raggruppati.
Vedere le formule nell'esercizio 27 (Suggerimenti)

classi	frequenze assolute fi	valori centrali mi	fi*mi	fi*mi^2	
6<x<=10	3	8	24	192	
10<x<=14	15	12	180	2160	
14<x<=18	24	16	384	6144	
18<x<=22	34	20	680	13600	
22<x<=26	18	24	432	10368	
26<x<=30	6	28	168	4704	
			1868	37168	totali

Numero dati del campione	100
media campione	18,68
varianza campione	22,9673
deviazione standard campione	4,79

grado di fiducia 95%	0,95
ampiezza metà intervallo	0,94

Intervallo di confidenza per la media

estremo sinistro	17,74
estremo destro	19,62

Esercizio 60.4

Sia dato un campione di 120 dati avente la seguente distribuzione di frequenza

Tabella 4

Classi	Frequenze assolute
30 < *x* <= 40	8
40 < *x* <= 50	28
50 < *x* <= 60	38
60 < *x* <= 70	30
70 < *x* <= 80	12
80 < *x* <= 90	4

Calcolare la media campionaria e la varianza campionaria (dati raggruppati).
Determinare l'intervallo di confidenza al 99% per la media della popolazione da cui proviene il campione.

classi	frequenze assolute fi	valori centrali mi	fi*mi	fi*mi^2	
30 < *x* <= 40	8	35	280	9800	
40 < *x* <= 50	28	45	1260	56700	
50 < *x* <= 60	38	55	2090	114950	
60 < *x* <= 70	30	65	1950	126750	
70 < *x* <= 80	12	75	900	67500	
80 < *x* <= 90	4	85	340	28900	
	120		6820	404600	totali

media	56,83
varianza	142,83
deviazione standard	11,95

Intervallo di confidenza per la media

grado fiducia	0,99
ampiezza metà intervallo	2,81
estremo sinistro	54,02
estremo destro	59,64

Torna su

Soluzione Esercizio 61
Intervallo di confidenza per la media
(varianza della popolazione incognita - grandi campioni)

Indice

Ritorna Esercizio 61

Se la varianza della popolazione non è nota, per grandi campioni può essere sostituita con la varianza del campione, ma con questa sostituzione si commette un errore di approssimazione. Tale errore può essere evitato usando la distribuzione t di Student per calcolare il valore critico, invece della distribuzione normale.

Esempio 61.1

Le misure dei diametri di un campione di 200 sferette prodotte in una settimana hanno una media uguale a 0,824 cm e una deviazione standard campionaria s = 0,042 cm

Trovare l'intervallo di confidenza per la media della popolazione con grado di fiducia del 95%

Calcolo dell'intervallo di confidenza con la distribuzione normale (intervallo approssimato)
Si può usare la funzione CONFIDENZA perché si tratta di un grande campione, e si ottiene un intervallo approssimato

numero dati del campione	200
media campionaria	0,824
deviazione standard campionaria	0,042
grado di fiducia	0,95
ampiezza metà intervallo	0,00582

Intervallo di confidenza per la media	
estremo sinistro	0,8182
estremo destro	0,8298

Calcolo dell'intervallo di confidenza con la distribuzione t di Student (intervallo esatto)

Per trovare il valore critico per il fissato grado di fiducia si usa la funzione **INV.T**

Per trovare l'intervallo di confidenza si usa la formula

$$\overline{x} - t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} < \mu < \overline{x} + t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$$

$t_{\frac{\alpha}{2}}$ indica il valore critico della distribuzione t di Student

grado di libertà	199
valore critico della distrib. di Student	1,972

Intervallo di confidenza per la media	
estremo sinistro	0,8181
estremo destro	0,8299

=E25-E44*E26/RADQ(E24)

=E25+E44*E26/RADQ(E24)

Esercizio 61.2

Da una popolazione di studenti universitari maschi viene scelto un campione di 50 studenti
Il peso medio calcolato in base al campione è di 67,45 Kg; la varianza campionaria è s^2=8,6 Kg^2
Trovare l'intervallo di confidenza con grado di fiducia del 95% usando sia la distribuzione normale che la distribuzione di Student

Intervallo approssimato con CONFIDENZA (distribuzione normale)

numero dati del campione	50
media campionaria	67,45
varianza campionaria	8,6
deviazione standard campionaria	2,9326
grado di fiducia	0,95
ampiezza metà intervallo	0,813

Intervallo di confidenza per la media	
estremo sinistro	66,637
estremo destro	68,263

Intervallo esatto (distribuzione di Student)

grado di libertà	49
valore critico distrib. Student	2,0096

Intervallo di confidenza per la media	
estremo sinistro	66,617
estremo destro	68,283

Torna su

Soluzione Esercizio 62

Intervalli di confidenza per la media
(varianza della popolazione incognita - piccoli campioni)

Indice

Ritorna Esercizio 62

Se il campione è piccolo e proviene da una popolazione con distribuzione normale di varianza incognita si deve usare la distribuzione t di Student

Dato un campione di ampiezza n (n<30) estratto da una popolazione con varianza incognita, l'intervallo di confidenza per la media μ della popolazione con grado di fiducia $(1-\alpha)$*100% è dato dalla formula

$$\boxed{\overline{x} - t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} < \mu < \overline{x} + t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}}$$

$\overline{x}$	valor medio del campione
s	deviazione standard del campione
$t_{\frac{\alpha}{2}}$	valore critico della distribuzione t di Student

L'ampiezza dell'intervallo di confidenza è uguale a $\boxed{2\, t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}}$

Esempio 62.1

E' dato un campione di 16 oggetti di cui si misura il peso, trovando un peso medio di 3,42g e una deviazione standard di 0,68g.

Trovare un intervallo di confidenza per la media della popolazione con grado di fiducia del 95%

Si suppone che la popolazione abbia distribuzione normale.

Numero dati del campione	16
media campionaria	3,42
deviazione standard campionaria	0,68
grado di libertà	15
grado di fiducia	0,95
valore critico	2,1314

Intervallo di confidenza per la media	
estremo sinistro	3,058
estremo destro	3,782

=E34-E38*E35/RADQ(E33)

=E34+E38*E35/RADQ(E33)

Esercizio 62.2

Un campione di 10 misure del diametro di una sferetta ha una media campionaria di 4,38cm e una deviazione standard di 0,06cm

Si suppone che la popolazione abbia distribuzione normale.

Determinare gli intervalli di confidenza per il diametro medio con gradi di fiducia del 95%, 98%, e 99%

Usare la funzionalità di Excel di aggiornamento automatico per il grado di fiducia

Numero dati del campione	10
media campionaria	4,38
deviazione standard campionaria	0,06
grado di libertà	9
grado di fiducia	0,95
valore critico	2,2622

Grado di fiducia 95%

intervallo di confidenza per la media

estremo sinistro	4,337
estremo destro	4,423

Per rendere possibile il controllo dei risultati, si riportano di seguito i risultati per gli altri gradi di fiducia ottenibili con la funzionalità Excel di aggiornamento automatico, cambiando il valore del grado di fiducia nella cella E57

Grado di fiducia 98%

Intervallo di confidenza per la media	
estremo sinistro	4,326
estremo destro	4,434

Grado di fiducia 99%

Intervallo di confidenza per la media	
estremo sinistro	4,318
estremo destro	4,442

Esercizio 62.3

Le misure in Kg del peso di un campione di 10 studenti sono le seguenti

60	63	60	68	70
72	65	61	69	67

Determinare l'intervallo di confidenza per la media della popolazione con grado di fiducia 95%
Si suppone che la popolazione abbia distribuzione normale

Calcolo della media e della deviazione standard campionaria con le funzioni MEDIA e DEV.ST

media campionaria	65,5
deviazione standard campionaria	4,35

numero dati	10
grado di libertà	9
grado di fiducia	0,95
valore critico	2,262

intervallo di confidenza per la media

estremo sinistro	62,39
estremo destro	68,61

Esercizio 62.4

8 misure in g effettuate in laboratorio forniscono i seguenti dati

3,12	3,16	2,94	3,33	3	3,11	3,5	2,81

Determinare l'intervallo di confidenza per la media della popolazione con grado di fiducia 95%
Si suppone che la popolazione abbia distribuzione normale

Calcolo della media e della deviazione standard campionaria con le funzioni MEDIA e DEV.ST

media campionaria	3,121
deviazione standard campionaria	0,2181

numero dati	8
grado di libertà	7
grado di fiducia	0,95
valore critico	2,365

intervallo di confidenza per la media

estremo sinistro	2,939
estremo destro	3,304

Torna su

Soluzione Esercizio 63
Intervalli di confidenza per la media (varianza della popolazione incognita)
Strumento Analisi Dati: Statistica descrittiva

Indice

Ritorna Esercizio 63

Se sono disponibili i dati del campione, e non solo le statistiche campionarie, oltre al metodo illustrato nell'esercizio precedente, si può anche usare lo strumento **Statistica Descrittiva**

Se il menu **Strumenti** di Excel non contiene l'opzione **Analisi dati**, selezionare **Componenti aggiuntivi** dal menu Strumenti; nella finestra di dialogo **Componenti aggiuntivi** attivare la casella **Strumenti di analisi** (Vedere Esempio 8 per maggiori dettagli).

Esempio 63.1

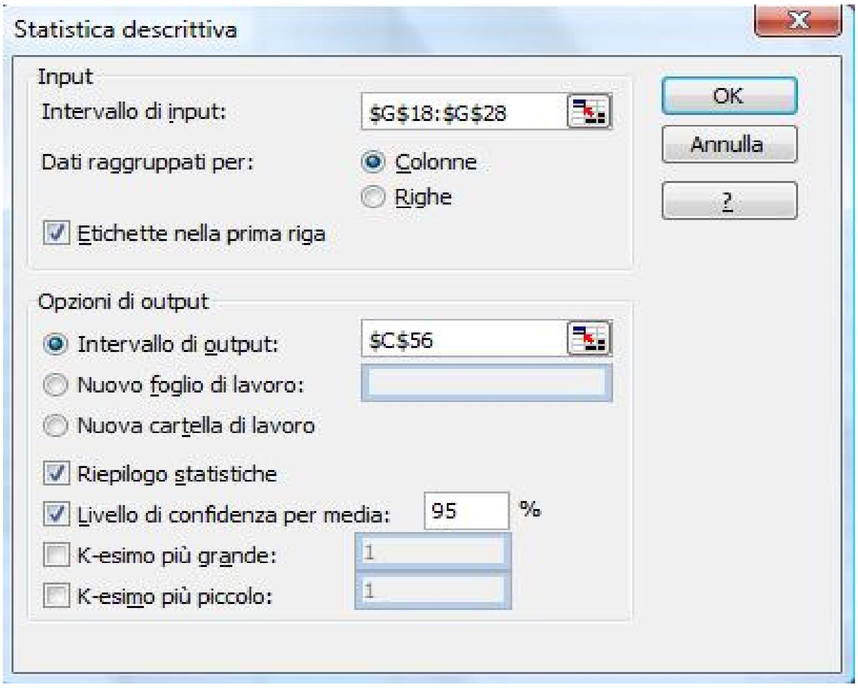
I pesi in Kg del peso di un campione di 10 studenti sono riportati nella Tabella 1
(**i dati devono essere disposti in colonna**)

Determinare l'intervallo di confidenza per la media della popolazione, che si suppone sia normale, usando lo strumento Statistica descrittiva

Tabella 1

Pesi
60
63
60
68
70
72
65
61
69
67

Dal menu **Strumenti** selezionare **Analisi dati, Statistica descrittiva**
Osservare nell'immagine le scelte da effettuare



Pesi	
Media	65,5
Errore standard	1,3764
Mediana	66
Moda	60
Deviazione standard	4,3525
Varianza campionaria	18,9444
Curtosi	-1,5001
Asimmetria	-0,0202
Intervallo	12
Minimo	60
Massimo	72
Somma	655
Conteggio	10
Livello di confidenza(95,0%)	3,1136

In questa cella compare la metà dell'ampiezza dell'intervallo di confidenza

Intervallo di confidenza per la media

estremo sinistro	-3,114
estremo destro	3,114

Se nella finestra di dialogo Statistica descrittiva riprodotta nell'immagine non si seleziona Riepilogo statistiche, in uscita si ottiene una tabella più sintetica con il solo risultato relativo all'intervallo di confidenza, del tipo seguente

Pesi	
Livello di confidenza(95,0%)	3,1136

Esercizio 63.2

8 misure in g effettuate in laboratorio forniscono i dati della Tabella 2
Determinare l'intervallo di confidenza per la media della popolazione, che si suppone sia normale, usando lo strumento Statistica descrittiva

Tabella 2

Misure
3,12
3,16
2,94
3,33
3
3,11
3,5
2,81

Misure	
Media	3,1213
Errore standard	0,0771
Mediana	3,115
Moda	#N/D
Deviazione standard	0,2181
Varianza campionaria	0,0476
Curtosi	0,1212
Asimmetria	0,4697
Intervallo	0,69
Minimo	2,81
Massimo	3,5
Somma	24,97
Conteggio	8
Livello di confidenza(95,0%)	0,1824

Intervallo di confidenza per la media

estremo sinistro	2,939
estremo destro	3,304

Torna su

Soluzione Esercizio 64
Intervalli di confidenza per la media
(varianza della popolazione incognita - grandi campioni)

Indice

Ritorna Esercizio 64

Lo strumento **Statistica descrittiva** può essere utilizzato anche per grandi campioni, nel caso in cui la varianza della popolazione non sia nota; occorre disporre della tabella completa dei dati, disposti in colonna, e si procede poi come nell'esercizio precedente
L'intervallo viene ricavato usando la distribuzione t di Student ed è quindi più preciso di quello che si potrebbe trovare con la funzione CONFIDENZA (che comunque può essere usata)

Esempio 64.1
I seguenti dati (tabella 1) sono il risultato di 80 determinazioni, in una data unità di misura, dell'emissione giornaliera di un gas inquinante da un impianto industriale
Trovare le statistiche e l'intervallo di confidenza per la media con grado di fiducia 95% con lo Strumento di Analisi Statistica descrittiva

Tabella 1

15,8	24,6	24,8	13,5
22,7	19,4	26,1	24,6
26,8	12,3	20,9	20
19,1	15,9	21,4	24,1
18,5	11,2	18	9
14,4	14,7	24,3	17,6
8,3	20,5	11,8	16,7
25,9	26,6	17,9	16,9
26,4	20,1	18,7	23,5
9,8	17	12,8	18,4
22,7	22,3	15,5	25,7
15,2	27,5	19,2	20,1
23	23,9	7,7	13,2
29,6	17,5	22,5	23,7
21,9	11	19,3	10,7
10,5	20,4	9,4	19
17,3	16,2	13,9	14,5
6,2	20,8	28,6	18,1
18	13,3	19,4	31,8
22,9	18,1	21,6	28,5

Tabella 1B

Emissioni gas
15,8
22,7
26,8
19,1
18,5
14,4
8,3
25,9
26,4
9,8
22,7
15,2
23
29,6
21,9
10,5
17,3
6,2
18
22,9
24,6
19,4
12,3
15,9
11,2
14,7
20,5
26,6
20,1
17
22,3

SUGGERIMENTI
Ricordare che per utilizzare lo strumento Statistica descrittiva
i **dati** devono essere disposti **in colonna** (Tabella 1B)

Emissioni gas	
Media	18,8963
Errore standard	0,6324
Mediana	19,05
Moda	22,7
Deviazione standard	5,6565
Varianza campionaria	31,9956
Curtosi	-0,4983
Asimmetria	-0,1025
Intervallo	25,6
Minimo	6,2
Massimo	31,8
Somma	1511,7
Conteggio	80
Livello di confidenza(95,0%)	1,2588

Intervallo di confidenza per la media	
estremo sinistro	17,64
estremo destro	20,16

Questo intervallo è più preciso di quello che si trova con la funzione CONFIDENZA, perché il valore critico viene calcolato usando la distribuzione di Student anziché la distribuzione normale

Esercizio 64.2

Sono assegnate le seguenti misure (tabella 2); trovare i valori delle statistiche campionarie e l'intervallo di confidenza per la media con grado di fiducia 95%, con lo Strumento di Analisi **Statistica descrittiva**

Tabella 2

35	68	29	48
13	52	45	47
41	24	45	55
32	55	56	40
28	45	45	46
51	36	50	46
51	48	43	49
39	48	29	49
43	47	55	30
51	52	39	42
38	46	45	42
34	51	36	29
34	27	36	32
61	39	47	50
38	38	25	38
41	23	37	43
50	29	54	33
40	29	43	33
39	54	54	34
51	31	56	44
42	55	40	30
47	42	50	47
47	30	50	45
56	19	21	52
33	39	44	37

27,5
23,9
17,5
11
20,4
16,2
20,8
13,3
18,1
24,8
26,1
20,9
21,4
18
24,3
11,8
17,9
18,7
12,8
15,5
19,2
7,7
22,5
19,3
9,4
13,9
28,6
19,4
21,6
13,5
24,6
20
24,1
9
17,6
16,7
16,9
23,5
18,4
25,7
20,1
13,2
23,7
10,7
19
14,5
18,1
31,8
28,5

Misure	
Media	41,77
Errore standard	0,9945
Mediana	43
Moda	47
Deviazione standard	9,9452
Varianza campionaria	98,9062
Curtosi	-0,0037
Asimmetria	-0,3077
Intervallo	55
Minimo	13
Massimo	68
Somma	4177
Conteggio	100
Livello di confidenza(95,0%)	1,9733

Intervallo di confidenza per la media	
estremo sinistro	39,80
estremo destro	43,74

Tabella 2B

Misure
35
13
41
32
28
51
51
39
43
51
38
34
34
61
38
41
50
40
39
51
42
47
47
56
33
68
52
24
55
45
36
48
48
47
52
46
51
27
39
38
23
29
29
54
31
55
42
30
19
39
29
45
45
56

45
50
43
29
55
39
45
36
36
47
25
37
54
43
54
56
40
50
50
21
44
48
47
55
40
46
46
49
49
30
42
42
29
32
50
38
43
33
33
34
44
30
47
45
52
37

Torna su

Soluzione Esercizio 65
Intervalli di confidenza per la varianza

Indice

Ritorna Esercizio 65

Per calcolare l'intervallo di confidenza per la varianza di una popolazione occorre che
la popolazione da cui viene estratto il campione abbia distribuzione normale
Si usa la distribuzione chi-quadro; il metodo viene usato sia per piccoli che per grandi campioni
Dato un campione di ampiezza n estratto da una popolazione normale
l'intervallo di confidenza per la varianza σ^2 della popolazione con grado di fiducia
$(1 - \alpha)$*100% è dato dalla formula

$$\frac{(n-1)\, s^2}{\chi^2_{\frac{\alpha}{2}}} < \sigma^2 < \frac{(n-1)\, s^2}{\chi^2_{1-\frac{\alpha}{2}}}$$

grado di libertà della distribuzione chi quadro $\nu = n - 1$
Per calcolare i valori critici usare la funzione **INV.CHI**

Esempio 65.1
In una scuola superiore si sceglie un campione di 16 studenti dell'ultimo anno e si misura
l'altezza degli studenti. La varianza campionaria è 37,09cm^2
Trovare gli intervalli di confidenza per la varianza della popolazione con grado di fiducia del 95%
e del 99%
Si suppone che la popolazione sia normale.

numero dati	16		
varianza campionaria	37,09		
grado di libertà	15		
grado di fiducia	0,95		
alfa	0,05		
alfa/2	0,025	valori critici	27,488
1-alfa/2	0,975		6,262

Intervallo di confidenza per la varianza

estremo sinistro	20,239
estremo destro	88,843

grado di fiducia	0,99		
alfa	0,01		
alfa/2	0,005	valori critici	32,801
1-alfa/2	0,995		4,601

Intervallo di confidenza per la varianza

estremo sinistro	16,961
estremo destro	120,922

Esercizio 65.2

Le misure della durata in ore di 10 batterie sono le seguenti (tabella 1)
Trovare media e varianza campionarie e l'intervallo di confidenza per la media con lo Strumento Statistica descrittiva
Trovare l'intervallo di confidenza per la varianza con grado di fiducia 95%

Tabella 1

Durata (ore)
140
136
150
144
148
152
138
141
143
151

Durata (ore)	
Media	144,3
Errore standard	1,7954
Mediana	143,5
Moda	#N/D
Deviazione standard	5,6774
Varianza campionaria	32,2333
Curtosi	-1,4913
Asimmetria	0,0407
Intervallo	16
Minimo	136
Massimo	152
Somma	1443
Conteggio	10
Livello di confidenza(95,0%)	4,0614

numero dati	10		
grado di libertà	9		
grado di fiducia	0,95		
alfa	0,05		
alfa/2	0,025	valori critici	19,023
1-alfa/2	0,975		2,700

Intervallo di confidenza per la varianza

estremo sinistro	15,25
estremo destro	107,43

Esercizio 65.3

La tabella 2 riporta la distribuzione di frequenza dei pesi in kg di 200 studenti di sesso maschile scelti fra gli iscritti al primo anno di università.
Trovare gli intervalli di confidenza per media e varianza della popolazione di tutti gli studenti del primo anno, con grado di fiducia del 95%

Tabella 2

classi (peso kg)	numero studenti
58 < x ≤ 62	15
62 < x ≤ 66	31
66 < x ≤ 70	50
70 < x ≤ 74	54
74 < x ≤ 78	36
78 < x ≤ 82	14

Poiché non sono noti i dati grezzi, ma solo la distribuzione di frequenza, per la media e la varianza campionarie si devono usare le formule per i dati raggruppati

Calcolo di media e varianza con le formule per dati raggruppati

classi (peso kg)	numero studenti fi	valori centrali mi	fi*mi	fi*mi^2
58 < x ≤ 62	15	60	900	54000
62 < x ≤ 66	31	64	1984	126976
66 < x ≤ 70	50	68	3400	231200
70 < x ≤ 74	54	72	3888	279936
74 < x ≤ 78	36	76	2736	207936
78 < x ≤ 82	14	80	1120	89600
	200		14028	989648

Intervallo di confidenza per la media

media	70,14
varianza	28,764
deviazione standard	5,363
grado fiducia	0,95
ampiezza metà intervallo	0,74
estremo sinistro	69,40
estremo destro	70,88

Intervallo di confidenza per la varianza

grado libertà	199
grado fiducia	0,95
alfa	0,05
alfa/2	0,025
1-alfa/2	0,975
valori critici	239,96
	161,83
estremo sinistro	23,854
estremo destro	35,372

Torna su

Soluzione Esercizio 66
Intervalli di confidenza per la varianza (grandi campioni)

Indice

Ritorna Esercizio 66

Per calcolare l'intervallo di confidenza per la varianza di una popolazione occorre che la popolazione da cui viene estratto il campione abbia distribuzione normale
Il metodo descritto nell'esercizio precedente, basato sulla distribuzione chi quadro può essere usato sia per piccoli che per grandi campioni
Nel caso di un **grande campione** si può trovare un **intervallo di confidenza per la deviazione standard** anche con la formula seguente

$$\frac{s}{1+\dfrac{z_{\frac{\alpha}{2}}}{\sqrt{2n}}} < \sigma < \frac{s}{1-\dfrac{z_{\frac{\alpha}{2}}}{\sqrt{2n}}}$$

La formula è basata sulla distribuzione normale e l'intervallo è approssimato.
L'intervallo di confidenza per la varianza si ottiene elevando al quadrato gli estremi dell'intervallo di confidenza per la deviazione standard

Esercizio 66.1
I dati della tabella 1 sono il risultato di 80 determinazioni, in una data unità di misura, dell'emissione giornaliera di un gas inquinante da un impianto industriale.
Trovare l'intervallo di confidenza per la varianza con grado di fiducia 95%, usando la distribuzione chi quadro.
Trovare l'intervallo di confidenza per la deviazione standard con grado di fiducia 95%, usando la distribuzione normale e ricavare l'intervallo di confidenza per la varianza.

SUGGERIMENTI
Dopo aver trovato l'intervallo di confidenza per la deviazione standard, si può trovare il corrispondente intervallo per la varianza semplicemente elevando al quadrato gli estremi dell'intervallo per la deviazione standard.
L'intervallo basato sulla distribuzione chi quadro è esatto, mentre quello basato sulla normale è approssimato.

Tabella 1

15,8	24,8	17,3	13,9	23,9
22,7	26,1	6,2	28,6	17,5
26,8	20,9	18	19,4	11
19,1	21,4	22,9	21,6	20,4
18,5	18	24,6	13,5	16,2
14,4	24,3	19,4	24,6	20,8
8,3	11,8	12,3	20	13,3
25,9	17,9	15,9	24,1	18,1
26,4	18,7	11,2	9	13,2
9,8	12,8	14,7	17,6	23,7
22,7	15,5	20,5	16,7	10,7
15,2	19,2	26,6	16,9	19
23	7,7	20,1	23,5	14,5
29,6	22,5	17	18,4	18,1
21,9	19,3	22,3	25,7	31,8
10,5	9,4	27,5	20,1	28,5

numero dati	80
media campionaria	18,90
varianza campionaria	32,00

grado di libertà	79
grado di fiducia	0,95
alfa	0,05
alfa/2	0,025
1-alfa/2	0,975
valori critici (chi quadro)	105,47 56,31

Intervallo di confidenza per la varianza (esatto) con la distribuzione chi quadro

estremo sinistro	23,965
estremo destro	44,889

Intervallo di confidenza per la deviazione standard (approssimato) con la distribuzione normale

valore critico (normale)	1,96
estremo sinistro	4,898
estremo destro	6,694

Intervallo di confidenza per la varianza (approssimato) con la distribuzione normale

estremo sinistro	23,986
estremo destro	44,805

Torna su

7. TEST DI IPOTESI

Soluzione Esercizio 67
Test di ipotesi. Introduzione e definizioni
Test di ipotesi sulla media
(varianza della popolazione nota - grandi campioni)

Indice

Ritorna Esercizio 67

Un test di ipotesi è un procedimento inferenziale che mette a confronto due ipotesi riguardanti una popolazione, una delle quali è la negazione dell'altra, dette ipotesi nulla H0 e ipotesi alternativa H1. Sulla base dei valori di un campione di dati, si decide se rifiutare o non rifiutare l'ipotesi nulla H0, con un determinato margine di errore.

I **punti fondamentali di un test di ipotesi** sono i seguenti:

1 formulare l'**ipotesi nulla H0** e l'**ipotesi alternativa H1**: le due ipotesi si escludono a vicenda
In base alla scelta delle due ipotesi, il test sarà a una coda o a due code.
Il **test a due code** viene usato per decidere se il parametro che si sottopone a test è diverso da un valore assegnato;
il **test a una coda** viene usato per decidere se il parametro che si sottopone a test è maggiore (coda di destra) o minore (coda di sinistra) di un valore assegnato;

2 scegliere il **livello di significatività alfa** a cui si vuole eseguire il test
Il livello di significatività alfa è uguale alla probabilità di rifiutare l'ipotesi nulla quando essa è vera (**errore del primo tipo**)

3 determinare i **valori critici** e la **regione di rifiuto**
La regione di rifiuto è l'insieme dei valori che conducono al rifiuto dell'ipotesi nulla
I valori critici sono i valori che separano la regione di rifiuto da quella di accettazione

4 calcolare, sulla base dei dati del campione, il valore della **statistica test**
Il valore della statistica test è un numero che riassume le informazioni contenute nei dati del campione.
La formula per calcolare la statistica test dipende dal test che si effettua.

5 decidere se **rifiutare o non rifiutare l'ipotesi nulla** al livello di significatività scelto.
Il **primo metodo** per decidere consiste nel confrontare il valore della statistica test con la regione di rifiuto
Se il valore della statistica test cade nella regione di rifiuto si rifiuta l'ipotesi nulla,
se invece cade al di fuori non si rifiuta.
Il **secondo metodo** per decidere se rifiutare l'ipotesi nulla è basato sul calcolo del **p-value**
e sul confronto tra il p-value e il livello di significatività (Vedere Esercizio 68)

Test di ipotesi sulla media - varianza nota, grandi campioni

La statistica test per questo tipo di test si calcola con la formula seguente

$$\boxed{Z = \frac{\overline{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}}$$

n	numero di elementi del campione
$\overline{X}$	valor medio calcolato dai dati del campione
μ_0	valore della media assunto nell'ipotesi nulla
σ	scarto quadratico medio della popolazione

I valori critici si calcolano con la distribuzione normale (funzione INV.NORM.ST)
Se la popolazione da cui proviene il campione è normale, questo test per grandi campioni può essere usato anche nel caso di piccoli campioni con varianza nota (Esercizio 67.7)

Esempio 67.1 – Test a una coda (coda destra)

Un campione casuale di 150 maschi adulti residenti nell'Italia Settentrionale ha una statura media di 173 cm; lo scarto quadratico medio della popolazione si suppone noto e uguale a 30 cm

Sottoporre a test l'ipotesi che la statura media della popolazione maschile sia maggiore di 170 cm

SUGGERIMENTI

Si effettua un **test a una coda**

La regione di rifiuto è costituita dai valori a destra del valore critico (coda destra)

Ricordare che la funzione INV.NORM.ST opera sulla coda sinistra: per usare la coda destra (simmetrica) prendere il valore assoluto (vedere il commento alla cella F94)

Per poter prendere la decisione e concludere il test si può utilizzare la funzione SE, che esegue un test basato su una condizione da verificare su valori o formule

In questo caso si deve verificare se il valore della statistica test cade nella regione di rifiuto o no.

La funzione SE restituisce un valore se la condizione specificata ha valore VERO e un altro valore se essa ha valore FALSO

Sintassi

SE(test; se_vero; se_falso)

Test — è un valore o un'espressione qualsiasi che può dare come risultato VERO o FALSO. Ad esempio, A1>10 è un'espressione logica; se il valore contenuto nella cella A1 è maggiore di 10, l'espressione darà come risultato VERO. In caso contrario, l'espressione darà come risultato FALSO.

Se_vero — è il valore che viene restituito se test è VERO. Ad esempio, se questo argomento è la stringa di testo (fra apici) "Rifiuto" e l'argomento test dà come risultato VERO, allora la funzione SE visualizzerà il testo "Rifiuto".

Se_falso — è il valore che viene restituito se test è FALSO. Ad esempio, se questo argomento è la stringa di testo (fra apici) "Non rifiuto" e l'argomento test dà come risultato FALSO, allora la funzione SE visualizzerà il testo "Non rifiuto".

Test a una coda (coda destra)

ipotesi nulla H0	media popolazione <=	170
ipotesi alternativa H1	media popolazione >	170

n° dati campione	150
media campione	173
media ipotesi H0	170
scarto quadr. medio popolazione	30
livello significatività alfa	0,05
statistica test	1,2247
valore critico	1,6449

=ASS(INV.NORM.ST(F92))

Decisione	Non rifiuto H0

=SE(F93>F94;"Rifiuto H0";"Non rifiuto H0")

Conclusione: la statura media è minore o uguale a 170 cm

Nota importante: la frase corretta per la conclusione dovrebbe essere: "Sulla base dei dati del campione non possiamo rifiutare l'ipotesi che la statura media sia minore o uguale a 170 cm, con livello di significatività del 5%", frase che sintetizziamo per brevità, in questo come negli esempi ed esercizi successivi.

Esempio 67.2 – Test a una coda (coda sinistra)

Un campione casuale di 150 femmine adulte residenti nell'Italia Settentrionale ha una statura media di 160 cm; lo scarto quadratico medio della popolazione si suppone noto e uguale a 30 cm
Sottoporre a test l'ipotesi che la statura media della popolazione femminile sia minore di 165 cm

Test a una coda (coda sinistra)

ipotesi nulla H0	media popolazione >=	165
ipotesi alternativa H1	media popolazione <	165

n° dati campione	150
media campione	160
media ipotesi H0	165
scarto quadr. medio popolazione	30
livello significatività alfa	0,05
statistica test	-2,0412
valore critico	-1,6449

=INV.NORM.ST(F123)

Decisione	Rifiuto H0

=SE(F124<F125;"Rifiuto H0";"Non rifiuto H0")

Conclusione: la statura media è minore di 165 cm

Esempio 67.3 – Test a due code

Un campione casuale di 150 maschi adulti residenti nell'Italia Settentrionale ha una statura media di 175 cm; lo scarto quadratico medio della popolazione si suppone noto e uguale a 30 cm
Sottoporre a test l'ipotesi che la statura media della popolazione maschile sia diversa da 170 cm

SUGGERIMENTI

Nel test a due code i valori critici sono due, simmetrici fra loro, e la regione di rifiuto è costituita dai valori esterni all'intervallo di estremi i valori critici.
Nel calcolo dei valori critici per il test a due code, ricordare di dividere a metà il valore di alfa
Per scrivere la condizione test nella funzione SE si usa la funzione O (che corrisponde a *oppure,* connettivo logico OR nella logica)
La funzione O restituisce VERO se uno o più argomenti hanno valore VERO e restituisce FALSO se tutti gli argomenti hanno valore FALSO.

Sintassi

O(logico1;logico2;...)

Logico1;logico2;... sono le condizioni logiche da verificare, che possono avere valore VERO o FALSO

Test a due code

ipotesi nulla H0	media popolazione =	170
ipotesi alternativa H1	media popolazione ≠	170

n° dati campione	150
media campione	175
media ipotesi H0	170
scarto quadr. medio popolazione	30
livello significatività alfa	0,05
statistica test	2,0412
valori critici	-1,9600 1,9600

=INV.NORM.ST(F157/2)

Decisione	Rifiuto H0

=SE(O(F158<F159;F158>F160);"Rifiuto H0";"Non rifiuto H0")

Conclusione: La statura media è diversa da 170 cm

Esercizio 67.4

Si vuole studiare il problema dei tempi di attesa al telefono per collegarsi al servizio clienti di una società telefonica e parlare con l'operatore

I dati di un campione di 50 osservazioni dei tempi di attesa in minuti sono raccolti nella tabella 1

Sottoporre a test l'ipotesi che il tempo medio di attesa sia di 5 minuti; si suppone che lo scarto quadratico medio della popolazione sia noto e uguale a 2 minuti

Tabella 1

0,6	3,6	4,8	5,8	6,9
0,9	4	4,9	5,9	7,2
1,4	4	4,9	6	7,2
1,8	4,1	5	6,1	7,4
2,5	4,1	5,2	6,2	7,4
2,6	4,3	5,5	6,3	7,6
2,7	4,3	5,5	6,4	8,1
3,2	4,6	5,6	6,4	8,2
3,5	4,6	5,6	6,6	8,7
3,5	4,7	5,8	6,6	9,1

Si effettua un **test a due code**

ipotesi nulla H0	media popolazione =	5
ipotesi alternativa H1	media popolazione ≠	5

n° dati campione	50
media campione	5,158
media ipotesi H0	5
scarto quadr. medio popolazione	2
livello significatività alfa	0,05
statistica test	0,5586
valori critici	-1,96 1,96

Decisione	Non rifiuto H0

Conclusione: il tempo medio di attesa è di 5 minuti

Esercizio 67.5

I carichi di rottura dei cavi prodotti da un’azienda hanno una media pari a 1800kg e uno scarto quadratico medio uguale a 100kg.

Si vuole stabilire se con una nuova tecnica di costruzione il carico può essere reso maggiore.

Per effettuare il test si effettua una prova su 50 cavi e si trova che il carico di rottura medio è 1850kg.

E’ possibile affermare che il carico di rottura è aumentato ad un livello di significatività dell’1%?

Si effettua un **test a una coda (coda destra)**

ipotesi nulla H0	media popolazione <=	1800
ipotesi alternativa H1	media popolazione >	1800

n° dati campione	50
media campione	1850
media ipotesi H0	1800
scarto quadr. medio popolazione	100
livello significatività alfa	0,01
statistica test	3,5355
valore critico	2,3263

Decisione	Rifiuto H0

Conclusione: La media della popolazione è maggiore di 1800kg, quindi
la nuova tecnica di costruzione permette di aumentare il carico di rottura

Esercizio 67.6

Un campione di 36 osservazioni avente media uguale a 86,2 viene estratto da una popolazione la cui distribuzione ha varianza uguale a 100.
In passato la media della distribuzione era uguale a 83, ma si ha motivo di ipotizzare che recentemente la media possa essere cambiata.
Usando il livello di significatività del 5%, sottoporre a test l'ipotesi nulla opportuna nei seguenti casi:
caso 1 - si supponga di non sapere, nel caso la media sia cambiata, se è aumentata o diminuita;
caso 2 - si supponga di sapere che, nel caso la media sia cambiata, può solo essere aumentata.

Caso 1 - Si effettua un **test a due code**

ipotesi nulla H0	media popolazione =	83
ipotesi alternativa H1	media popolazione ≠	83

n° dati campione	36
media campione	86,2
media ipotesi H0	83
scarto quadr. medio popolazione	10
livello significatività alfa	0,05
statistica test	1,9200
valori critici	-1,96
	1,96

Decisione	Non rifiuto H0

Conclusione: la media non è cambiata

Caso 2 - Si effettua un **test a una coda (coda destra)**

ipotesi nulla H0	media popolazione <=	83
ipotesi alternativa H1	media popolazione >	83

n° dati campione	36
media campione	86,2
media ipotesi H0	83
scarto quadr. medio popolazione	10
livello significatività alfa	0,05
statistica test	1,9200
valore critico	1,6449

Decisione	Rifiuto H0

Conclusione: la media è aumentata

Esercizio 67.7

Si supponga che i punteggi di un test sul quoziente di intelligenza di una certa popolazione di adulti si distribuiscano normalmente con uno scarto quadratico medio uguale a 15.
Un campione di 25 adulti estratti da questa popolazione ha un punteggio medio di 105.
Sottoporre a test l’ipotesi che il punteggio medio sia 100, con un livello di significatività del 5%

La popolazione da cui proviene il campione ha distribuzione normale, quindi il test per grandi campioni può essere effettuato anche se si ha un piccolo campione.

Si effettua un **test a due code**

ipotesi nulla H0	media popolazione =	100
ipotesi alternativa H1	media popolazione ≠	100

n° dati campione	25
media campione	105
media ipotesi H0	100
scarto quadr. medio popolazione	15
livello significatività alfa	0,05
statistica test	1,6667
valori critici	-1,96
	1,96

Decisione	Non rifiuto H0

Conclusione: Il punteggio medio è uguale a 100

Esercizio 67.8

Da una popolazione normale avente scarto quadratico medio uguale a 2, si estrae un campione di ampiezza n = 10. Il valor medio del campione è 18,58
Sottoporre a test l’ipotesi che la media sia uguale a 20 ai livelli di significatività 1% e 5%

Poiché la popolazione da cui proviene il campione è normale, si può effettuare il test per grandi campioni anche se il campione è piccolo

Si effettua un **test a due code**

ipotesi nulla H0	media popolazione =	20
ipotesi alternativa H1	media popolazione ≠	20

n° dati campione	10	
media campione	18,58	
media ipotesi H0	20	
scarto quadr. medio popolazione	2	
livello significatività alfa	0,05	0,01
statistica test	-2,2452	-2,2452
valori critici	-1,9600	-2,5758
	1,9600	2,5758

Decisione	Rifiuto H0	Non rifiuto H0

Torna su

Conclusione:
La decisione che si prende dipende dal livello di significatività fissato: la differenza fra la media del campione e il valore 20 ipotizzato per il parametro della popolazione viene ritenuta statisticamente significativa al livello del 5%, ma non al livello dell’1%. Si tratta dunque di un **caso critico**.

Soluzione Esercizio 68

Test di ipotesi sulla media. Calcolo del p-value (varianza della popolazione nota - grandi campioni)

Indice

Ritorna Esercizio 68

Calcolo del p-value

Il **secondo metodo** per decidere se rifiutare l'ipotesi nulla è basato sul calcolo del **p-value** e sul confronto tra il p-value e i livelli di significatività
Il p-value è il più piccolo valore del livello di significatività alfa per cui i dati del campione consentono di rifiutare l'ipotesi nulla
Un p-value prossimo a 0 indica che la probabilità di sbagliare rifiutando l'ipotesi nulla è molto vicina a 0, ossia si è praticamente certi di non sbagliare rifiutando l'ipotesi nulla
Un p-value vicino ai classici livelli di significatività indica che la decisione è critica e dipende in modo cruciale dal livello di significatività
Un p-value maggiore indica che si è praticamente certi di non sbagliare non rifiutando l'ipotesi nulla

Per i test basati sulla distribuzione normale il p-value si calcola con le seguenti formule:

$$p-\text{value} = \begin{cases} 1 - P(Z < Z_0) & \text{per il test a una coda con} \quad H_0 : \mu \le \mu_0 \quad H_1 : \mu > \mu_0 \\ P(Z < Z_0) & \text{per il test a una coda con} \quad H_0 : \mu \ge \mu_0 \quad H_1 : \mu < \mu_0 \\ 2[1 - P(Z < |Z_0|)] & \text{per il test a due code con} \quad H_0 : \mu = \mu_0 \quad H_1 : \mu \ne \mu_0 \end{cases}$$

Z_0 è il valore della statistica test, calcolato in base ai dati campionari
Il p-value viene fornito da Excel quando si eseguono i test con gli Strumenti Analisi Dati

Esempio 68.1

I carichi di rottura dei cavi prodotti da un'azienda hanno una media pari a 1800kg e uno scarto quadratico medio uguale a 100kg.
Si vuole stabilire se mediante una nuova tecnica di costruzione il carico di rottura può essere reso maggiore. Per effettuare il test si effettua una prova su 50 cavi e si trova che il carico di rottura medio è di 1850kg. E' possibile affermare che il carico di rottura è aumentato?

Si effettua un **test a una coda (coda destra)**

ipotesi nulla H0	media popolazione <=	1800
ipotesi alternativa H1	media popolazione >	1800

n° dati campione	50
media campione	1850
media ipotesi H0	1800
scarto quadratico medio popolazione	100
statistica test	3,5355

p-value	0,0002
Decisione	Rifiuto H0

=1-DISTRIB.NORM.ST(E46)

=SE(D48<0,01;"Rifiuto H0";SE(D48>0,05;"Non rifiuto H0";"Caso critico"))

Il p-value è prossimo a 0, perciò si rifiuta l'ipotesi nulla
Conclusione: la nuova tecnica di costruzione permette di aumentare il carico di rottura

Esercizio 68.2

Una ditta produttrice di pneumatici afferma che la durata media di un certo tipo di pneumatici per auto è di almeno 50000km.
Per sottoporre a test questa affermazione un campione di 40 pneumatici viene sottoposto a prove su strada e si misura una durata media di 48900km. Lo scarto quadratico medio della popolazione, noto da precedenti studi, è uguale a 2000km.
E' possibile accettare l'affermazione del produttore?

Si effettua un **test a una coda (coda sinistra)**

ipotesi nulla H0	media popolazione >=	50000
ipotesi alternativa H1	media popolazione <	50000

n° dati campione	40
media campione	48900
media ipotesi H0	50000
scarto quadratico medio popolazione	2000
statistica test	-3,4785

p-value	0,0003
Decisione	Rifiuto H0

Il p-value è prossimo a 0, perciò si rifiuta l'ipotesi nulla
Conclusione: l'affermazione del produttore non può essere accettata

Esercizio 68.3

Da una popolazione normale avente scarto quadratico medio uguale a 2, si estrae un campione di ampiezza n = 10. Il valor medio del campione è 18,58
Sottoporre a test l'ipotesi che la media sia uguale a 20

Poiché la popolazione da cui proviene il campione è normale, si può effettuare il test per grandi campioni anche se il campione è piccolo

Si effettua un **test a due code**

ipotesi nulla H0	media popolazione =	20
ipotesi alternativa H1	media popolazione ≠	20

n° dati campione	10
media campione	18,58
media ipotesi H0	20
scarto quadratico medio popolazione	2
statistica test	-2,2452

p-value	0,02475
Decisione	Caso critico

Conclusione: In questo caso il p-value è compreso fra i classici livelli di significatività e ciò indica che la decisione è critica. Per poter prendere una decisione occorre disporre di un campione di maggior ampiezza.

Esercizio 68.4

Si supponga che i punteggi di un test sul quoziente di intelligenza di una certa popolazione di adulti si distribuiscano normalmente con uno scarto quadratico medio uguale a 15.
Un campione di 25 adulti estratti da questa popolazione ha un punteggio medio di 105.
Sottoporre a test l'ipotesi che il punteggio medio sia 100, con un livello di significatività del 5%

La popolazione da cui proviene il campione ha distribuzione normale, quindi il test per grandi campioni può essere effettuato anche se si ha un piccolo campione.

Si effettua un **test a due code**

ipotesi nulla H0	media popolazione =	100
ipotesi alternativa H1	media popolazione ≠	100

n° dati campione	25
media campione	105
media ipotesi H0	100
scarto quadratico medio popolazione	15
statistica test	1,6667

p-value	0,09558
Decisione	Non rifiuto H0

Il p-value è maggiore di 0,5 perciò non si rifiuta l'ipotesi nulla
Conclusione: Il punteggio medio è uguale a 100

Torna su

Soluzione Esercizio 69

Test di ipotesi sulla media
(varianza della popolazione incognita - grandi campioni)

Indice

Ritorna Esercizio 69

Il test di ipotesi illustrato nell'esercizio 67 richiede che sia noto il valore dello scarto quadratico medio della popolazione da cui è estratto il campione; se lo scarto quadratico medio non è noto, ma il campione è grande, si può utilizzare il valore s dello scarto quadratico medio del campione, al posto dello scarto della popolazione, commettendo un errore di approssimazione.

In alternativa potrà anche essere usato il test basato sull'uso della distribuzione t di Student che sarà descritto nel prossimo Esercizio 70 (in questo modo si evita l'errore di approssimazione)

Esercizio 69.1

Una ditta produttrice di pneumatici afferma che la durata media di un certo tipo di pneumatici per auto è di almeno 50000km.
Per sottoporre a test questa affermazione un campione di 40 pneumatici viene sottoposto a prove su strada e si misura una durata media di 48900 km, con uno scarto quadratico medio (calcolato dal campione) s = 2500km.
Sottoporre a test l'affermazione, con un livello di significatività 1%

Si effettua un **test a una coda (coda sinistra)**

ipotesi nulla H0	media popolazione >=	50000
ipotesi alternativa H1	media popolazione <	50000

n° dati campione	40
media campione	48900
media ipotesi H0	50000
scarto quadratico medio campione	2500
livello significatività alfa	0,01
statistica test	-2,7828
valore critico	-2,3263

Decisione	Rifiuto H0

Conclusione: La durata dei pneumatici è minore di 50000 km

Calcolo del p-value

p-value	0,00269
Decisione	Rifiuto H0

Il p-value è prossimo a 0, perciò si rifiuta l'ipotesi nulla

Esercizio 69.2

In un dato anno il voto medio all'esame di maturità classica è stato di 78/100.
In una commissione che ha esaminato 70 candidati, si è registrato un voto medio di 81,2/100 con uno scarto quadratico medio s = 14.
Verificare l'ipotesi che non ci sia differenza significativa tra la media generale e la media del campione, al livello di significatività del 5%.

Si effettua un **test a due code**

ipotesi nulla H0	media popolazione =	78
ipotesi alternativa H1	media popolazione ≠	78

n° dati campione	70
media campione	81,2
media ipotesi H0	78
scarto quadratico medio campione	14
livello significatività alfa	0,05
statistica test	1,9124
valori critici	-1,9600
	1,9600

Decisione	Non rifiuto H0

Conclusione: Non c'è differenza significativa fra le medie

Calcolo del p-value

p-value	0,05583
Decisione	Non rifiuto H0

Il p-value è maggiore di 0,5 perciò non si rifiuta l'ipotesi nulla

Esercizio 69.3

Negli Stati Uniti i pazienti che necessitano di un trapianto di cuore rimangono in lista d'attesa in media 60 giorni.
In un determinato ospedale, per un campione di 40 pazienti la media è di 65 giorni con scarto quadratico medio di 7 giorni.
La media in questo ospedale è maggiore di quella complessiva?

Si effettua un **test a una coda (coda destra)**

ipotesi nulla H0	media popolazione <=	60
ipotesi alternativa H1	media popolazione >	60

n° dati campione	40
media campione	65
media ipotesi H0	60
scarto quadratico medio campione	7
livello significatività alfa	0,01
statistica test	4,5175
valore critico	2,3263

Decisione	Rifiuto H0

Conclusione: La media in questo ospedale è maggiore della media di 60 giorni

Calcolo del p-value

p-value	0,0000031
Decisione	Rifiuto H0

Torna su

Il p-value è prossimo a 0, perciò si rifiuta l'ipotesi nulla

Soluzione Esercizio 70

Test di ipotesi sulla media (varianza della popolazione incognita)

Indice

Ritorna Esercizio 70

Test di ipotesi sulla media - varianza incognita

La statistica test per questo tipo di test si calcola con la formula seguente:

$$T = \frac{\overline{X} - \mu_0}{\frac{s}{\sqrt{n}}}$$

n	numero di elementi del campione
$\overline{X}$	valor medio calcolato dai dati del campione
μ_0	valore della media assunto nell'ipotesi nulla
s	scarto quadratico medio del campione

La **popolazione** da cui proviene il campione deve avere **distribuzione normale**

I valori critici si calcolano con la distribuzione t di Student (funzione INV.T), con grado di libertà n–1

Esempio 70.1

Le bottiglie di vino poste in vendita contengono usualmente 750ml di vino.

Si effettua un controllo su un campione di 6 bottiglie e si misurano i valori in ml della tabella 1

Tabella 1

747,5	747	749
747	751,5	752

Stabilire se questi dati confermano con un livello di significatività del 5% l'affermazione che le bottiglie hanno un contenuto medio uguale a quello dichiarato

SUGGERIMENTI

Si effettua un **test a due code**

Per il calcolo dei valori critici si usa la distribuzione t di Student (funzione INV.T)

La funzione INV.T opera su due code; per ottenere il valore critico relativo a una coda moltiplicare per 2 il valore del livello di significatività alfa

La funzione INV.T restituisce sempre un valore positivo: fare attenzione al segno nel calcolo dei valori critici, dove necessario

ipotesi nulla H0	media popolazione =	750
ipotesi alternativa H1	media popolazione ≠	750

n° dati campione	6
grado libertà	5
media campione	749
media ipotesi H0	750
scarto quadratico medio campione	2,2583
livello significatività alfa	0,05
statistica test	-1,0847
valori critici	-2,5706 2,5706

Attenzione al segno =-INV.T(D48;D44)

Decisione	Non rifiuto H0

=SE(O(D49<D50;D49>D51);"Rifiuto H0";"Non rifiuto H0")

Conclusione: Le bottiglie contengono in media la quantità di vino dichiarata

Se il test è effettuato per tutelare l'interesse del consumatore, l'ipotesi nulla e l'ipotesi alternativa sono invece le seguenti e si effettua un **test a una coda (coda sinistra)**
Fare attenzione al segno nel calcolo del valore critico e ricordare di moltiplicare il livello di significatività alfa per 2

ipotesi nulla H0	media popolazione >=	750
ipotesi alternativa H1	media popolazione <	750

n° dati campione	6
grado libertà	5
media campione	749
media ipotesi H0	750
scarto quadratico medio campione	2,2583
livello significatività alfa	0,05
statistica test	-1,0847
valore critico	-2,0150

Attenzione al segno. Moltiplicare alfa per 2 =-INV.T(2*D69;D65)

Decisione	Non rifiuto H0

=SE(D70<D71;"Rifiuto H0";"Non rifiuto H0")

Conclusione: Le bottiglie contengono in media la quantità di vino dichiarata

Esercizio 70.2
Si estrae un campione di 8 confezioni di detersivo in polvere da una grossa produzione.
La tabella 2 riporta il peso in g delle 8 confezioni

Tabella 2

1998	2002
1999	2005
2011	2007
2002	2005

Assumendo che popolazione da cui proviene il campione abbia distribuzione normale, verificare se al livello di significatività del 5%, si può affermare che il peso medio delle confezioni di questa produzione è maggiore di 2000g

Si effettua un **test a una coda (coda destra)**

ipotesi nulla H0	media popolazione <=	2000
ipotesi alternativa H1	media popolazione >	2000

n° dati campione	8
grado libertà	7
media campione	2003,625
media ipotesi H0	2000
scarto quadratico medio campione	4,2741
livello significatività alfa	0,05
statistica test	2,3989
valore critico	1,8946

Decisione	Rifiuto H0

Conclusione: Il peso medio delle confezioni è maggiore di 2000g

Esercizio 70.3

Un problema comune per molte aziende è quello di controllare il processo automatico di riempimento di confezioni alimentari.
Se la quantità di prodotto inserito è inferiore al dichiarato, si avranno reclami da parte dei consumatori, se è maggiore si avrà un costo per l'azienda.
Per controllare le confezioni di caffè allo scopo di accertare che contengano 250 g di prodotto si prende un campione di 24 confezioni e si registra il peso, riportato nella tabella 3

Tabella 3

245	245	252	250
244	248	253	250
250	247	251	245
251	250	252	251
254	253	250	253
247	250	247	252

Si assume che il peso delle confezioni sia distribuito normalmente.
Il peso medio è diverso da 250 g?
Scegliere il livello di significatività del 5%

Dato che si vuole individuare un'eventuale differenza dal valore di 250 g in entrambe le direzioni, si effettua un **test a due code**

ipotesi nulla H0	media popolazione =	250
ipotesi alternativa H1	media popolazione ≠	250

n° dati campione	24
grado libertà	23
media campione	249,58
media ipotesi H0	250
scarto quadratico medio campione	2,9180
livello significatività alfa	0,05
statistica test	-0,6995
valori critici	-2,0687 2,0687

Decisione	Non rifiuto H0

Conclusione: Le confezioni di caffè sono riempite correttamente

Esercizio 70.4

Si effettuano 8 misure sperimentali per stimare il valore di una lunghezza in cm
I risultati ottenuti sono riportati nella tabella 4

Tabella 4

1,8	1,9
2,3	2,2
1,9	2,6
2,1	2,8

Sulla base di questo campione di misure è possibile affermare che la media della popolazione da cui proviene il campione è uguale a 2, con un livello di significatività del 5%?

Si effettua un **test a due code**

ipotesi nulla H0	media popolazione =	2
ipotesi alternativa H1	media popolazione ≠	2

n° dati campione	8
grado libertà	7
media campione	2,20
media ipotesi H0	2
scarto quadratico medio campione	0,3546
livello significatività alfa	0,05
statistica test	1,5954
valori critici	-2,3646
	2,3646

Decisione	Non rifiuto H0

Conclusione: La media della popolazione è uguale a 2

Torna su

Soluzione Esercizio 71
Test di ipotesi sulla proporzione

Indice

Ritorna Esercizio 71

Test di ipotesi sulla proporzione - grandi campioni

Si sottopone a test l'ipotesi che la proporzione della popolazione abbia un valore p_0

La statistica test per questo tipo di test si calcola con la formula seguente

$$Z = \frac{X - np_0}{\sqrt{np_0(1-p_0)}}$$

n	numero di elementi del campione
X	numero di volte in cui la caratteristica osservata si presenta nel campione
p_0	valore della proporzione assunto nell'ipotesi nulla

I valori critici si calcolano con la distribuzione normale (funzione INV.NORM.ST)

Esempio 71.1 – Test a una coda (coda sinistra)

Una ditta farmaceutica asserisce che un suo farmaco è efficace nel 90% dei casi.
In un campione di 200 persone che lo hanno usato, il farmaco si è rivelato efficace in 160 casi.
Stabilire se l'affermazione della ditta farmaceutica è legittima con livello di significatività 1%

Test a una coda (coda sinistra)

ipotesi nulla H0	proporzione popolazione >=	0,9
ipotesi alternativa H1	proporzione popolazione <	0,9

n° dati campione	200
X	160
proporzione ipotesi H_0	0,9
livello significatività alfa	0,01
statistica test	-4,714
valore critico	-2,326

Decisione	Rifiuto H0

=SE(E36<E37;"Rifiuto H0";"Non rifiuto H0")

Conclusione: L'affermazione non è legittima

Esercizio 71.2

Una compagnia aerea afferma che non più del 6% dei bagagli smarriti viene definitivamente perso.
Sottoporre a test questa affermazione, sapendo che su un campione di 200 persone che hanno subito lo smarrimento del bagaglio, 17 non l'hanno più ritrovato; scegliere il livello di significatività 1%

Test a una coda (coda destra)

ipotesi nulla H0	proporzione popolazione <=	0,06
ipotesi alternativa H1	proporzione popolazione >	0,06

n° dati campione	200
X	17
proporzione ipotesi H_0	0,06
livello significatività alfa	0,01
statistica test	1,4887
valore critico	2,3263

Decisione	Non rifiuto H0

Conclusione: L'affermazione della compagnia aerea non può essere contestata

Esercizio 71.3

Si effettuano 500 lanci di una moneta e si ottiene 267 volte testa.

1 Decidere se la moneta è truccata oppure no, con un livello di significatività 5%.

2 Ripetere il calcolo nel caso che il numero di volte in cui si ottiene testa sia 280

Per una moneta non truccata la probabilità che esca testa è 0,5

Test a due code

ipotesi nulla H0	proporzione popolazione =	0,5
ipotesi alternativa H1	proporzione popolazione ≠	0,5

1

n° dati campione	500
X	267
proporzione ipotesi H_0	0,5
livello significatività alfa	0,05
statistica test	1,5205
valori critici	-1,9600
	1,9600

Decisione	Non rifiuto H0

Conclusione: La moneta non può ritenersi truccata

2 Servirsi dell'aggiornamento automatico: nella cella E77 inserire il dato 280 (al posto di 267) e osservare il cambiamento nella decisione

Conclusione In questo caso la moneta può ritenersi truccata

Esercizio 71.4

Un fabbricante dichiara che almeno il 95% della merce fornita a una ditta è conforme alle esigenze del cliente

L'esame di un campione di 200 esemplari della merce rivela che 18 esemplari sono difettosi

Sottoporre a test la dichiarazione del fabbricante al livello di significatività 1% e 5%

Test a una coda (coda sinistra)

ipotesi nulla H0	proporzione popolazione >=	0,95
ipotesi alternativa H1	proporzione popolazione <	0,95

n° dati campione	200
X	182
proporzione ipotesi H_0	0,95
livello significatività alfa	0,05
statistica test	-2,596
valore critico	-1,645

Decisione	Rifiuto H0

Cambiare il livello di significatività e verificare che la decisione non cambia
Conclusione: L'affermazione del fabbricante è falsa a entrambi i livelli di significatività

Esercizio 71.5

Un'università afferma che il 60% dei candidati che sostengono il test di ammissione per l'iscrizione al corso di laurea in Biotecnologie vengono ammessi al corso.
Da un campione di 300 diplomati che hanno sostenuto il test, ne vengono ammessi 148.
L'affermazione sostenuta dall'università è vera al livello di significatività 1%?

Test a una coda (coda sinistra)

ipotesi nulla H0	proporzione popolazione >=	0,6
ipotesi alternativa H1	proporzione popolazione <	0,6

n° dati campione	300
X	148
proporzione ipotesi H_0	0,6
livello significatività alfa	0,01
statistica test	-3,771
valore critico	-2,326

Decisione	Rifiuto H0

Conclusione: L'affermazione dell'università viene rifiutata, la percentuale degli ammessi è significativamente inferiore

Torna su

Soluzione Esercizio 72
Test di ipotesi sulla varianza

Indice

Ritorna Esercizio 72

Test di ipotesi sulla varianza

La statistica test per questo tipo di test si calcola con la formula seguente

$$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2}$$

n numero di elementi del campione

σ_0^2 valore della varianza assunto nell'ipotesi nulla

S^2 varianza del campione

I valori critici si calcolano con la distribuzione chi quadro (funzione INV.CHI)

Esempio 72.1

In un ospedale i chirurghi stanno sperimentando una nuova tecnica di intervento su una data patologia e si studia la variabilità della lunghezza della degenza

Si vuole verificare se con la nuova procedura la varianza della degenza sia inferiore rispetto a quella con la procedura tradizionale

Con la procedura tradizionale la lunghezza della degenza ha uno scarto quadratico medio di 5 giorni

Osservando un campione di 40 pazienti che sono stati sottoposti al nuovo tipo di intervento si osserva uno scarto quadratico medio s = 2,836 giorni.

Si può affermare che la varianza con la nuova procedura sia inferiore al livello di significatività 5%?

SUGGERIMENTI

Si effettua un **test a una coda (coda sinistra)**

La regione di rifiuto è costituita dai valori a sinistra del valore critico (coda sinistra)

Ricordare che la funzione INV.CHI opera sulla coda destra: per la coda sinistra usare per la probabilità il valore 1–alfa

Test a una coda (coda sinistra)

ipotesi nulla H0	varianza popolazione >=	25
ipotesi alternativa H1	varianza popolazione <	25

n° dati campione	40
scarto quadratico medio campione	2,836
scarto quadratico medio ipotesi H0	5
grado libertà	39
livello significatività alfa	0,05
statistica test	12,547
valore critico	25,695

=INV.CHI(1-D46;D45)

Decisione	Rifiuto H0

=SE(D47<D48;"Rifiuto H0";"Non rifiuto H0")

Conclusione: La varianza del tempo di degenza con la nuova procedura è inferiore
La lunghezza della degenza è meno variabile e può essere prevista con un minor margine di errore, il che è utile nella gestione della durata dei ricoveri e dei costi di degenza

Esercizio 72.2

Il peso di certi pacchetti confezionati automaticamente è distribuito secondo una distribuzione normale con scarto quadratico medio uguale a 0,25g.
L'esame di un campione di 20 confezioni ha permesso di calcolare uno scarto quadratico campionario s = 0,32g
Si può affermare che lo scarto quadratico medio è aumentato al livello di significatività 5%?
E al livello 1%?

Test a una coda (coda destra)

ipotesi nulla H0	varianza popolazione <=	0,0625
ipotesi alternativa H1	varianza popolazione >	0,0625

n° dati campione	20
scarto quadratico medio campione	0,32
scarto quadratico medio ipotesi H0	0,25
grado libertà	19
livello significatività alfa	0,05
statistica test	31,130
valore critico	30,144

Decisione	Rifiuto H0

Cambiare il livello di significatività e osservare il cambiamento nella decisione

Conclusione Si tratta di una decisione critica, che dipende dal livello di significatività
Per poter prendere una decisione univoca occorre un campione più grande

Esercizio 72.3

Lo scarto quadratico medio delle temperature annuali di una città in un periodo di 100 anni è stato di 8°C.
Misurando la temperatura media del quindicesimo giorno di ogni mese durante gli ultimi 15 anni si è riscontrato che lo scarto quadratico medio delle temperature annuali è stato di 5°C
Sottoporre a test l'ipotesi che la temperatura della città sia diventata meno variabile che in passato con livello di significatività 1%

Test a una coda (coda sinistra)

ipotesi nulla H0	varianza popolazione >=	64
ipotesi alternativa H1	varianza popolazione <	64

n° dati campione	15
scarto quadratico medio campione	5
scarto quadratico medio ipotesi H0	8
grado libertà	14
livello significatività alfa	0,05
statistica test	5,469
valore critico	6,571

Decisione	Rifiuto H0

Conclusione: La variabilità della temperatura è diminuita

Esercizio 72.4

E' noto che una certa popolazione avente distribuzione normale ha varianza $\sigma^2 = 22{,}5$
Da un'altra popolazione viene estratto il campione di dati della tabella 1

Tabella 1

16	12
10	8
0	12
10	6
10	8
4	2

Si può concludere al livello di significatività del 5% che la seconda popolazione abbia la stessa varianza della prima?

SUGGERIMENTI

Si effettua un test a due code
Ricordare che la funzione INV.CHI opera sulla coda destra: per operare su due code e trovare i due valori critici usare per la probabilità i valori alfa/2 e 1–alfa/2

Test a due code

ipotesi nulla H0	varianza popolazione =	22,5
ipotesi alternativa H1	varianza popolazione ≠	22,5

n° dati campione	12
varianza campione	20,697
varianza ipotesi H0	22,5
grado libertà	11
livello significatività alfa	0,05
statistica test	10,119
valori critici	3,816
	21,920

Decisione	Non rifiuto H0

Conclusione: Le due popolazioni hanno la stessa varianza

Esercizio 72.5

(Vedere anche Esercizio 70.3)
L'azienda che produce confezioni di caffè da 250 g vuole verificare se la varianza della popolazione del peso delle confezioni è uguale a 9 g^2
Per effettuare il test si preleva un campione di 24 confezioni e si registra il peso, riportato nella Tabella 3

Tabella 3

245	245	252	250
244	248	253	250
250	247	251	245
251	250	252	251
254	253	250	253
247	250	247	252

Test a due code

ipotesi nulla H0	varianza popolazione =	9
ipotesi alternativa H1	varianza popolazione ≠	9

n° dati campione	24
varianza campione	8,514
varianza ipotesi H0	9
grado libertà	23
livello significatività alfa	0,05
statistica test	21,759
valori critici	11,689
	38,076

Decisione	Non rifiuto H0

Conclusione: La varianza è uguale a 9 g^2

Torna su

Soluzione Esercizio 73

Test di ipotesi sulla differenza fra due medie (varianze delle popolazioni note)

Strumenti di Analisi: Test Z, due campioni per medie

Indice

Ritorna Esercizio 73

Per effettuare i test di ipotesi sul confronto fra medie si possono utilizzare gli **Strumenti Analisi Dati**; gli strumenti di analisi disponibili richiedono come input due campioni di **dati disposti in colonna**, la differenza ipotizzata fra le medie e il livello di significatività

Test di ipotesi sulla differenza fra due medie - varianze note

La statistica test per questo tipo di test si calcola con la formula seguente

$$\boxed{Z = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - d}{\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}}}$$

n_1 , n_2	ampiezze dei due campioni
$\overline{X}_1$, $\overline{X}_2$	medie dei due campioni
σ_1^2 , σ_2^2	varianze note delle due popolazioni
d	differenza ipotizzata fra le medie dei due campioni

Nel caso in cui le varianze delle due popolazioni siano note si usa lo strumento
Test Z: due campioni per medie
Se i due campioni sono piccoli, le popolazioni da cui provengono i campioni devono essere normali; se invece i campioni sono grandi il test può essere usato anche nel caso di popolazioni qualsiasi.

Esempio 73.1

Per verificare l'efficacia di un nuovo farmaco per il controllo dell'ipertensione vengono esaminati due gruppi di pazienti: al primo gruppo viene somministrato il nuovo farmaco sperimentale, al secondo gruppo viene somministrato un farmaco già comunemente usato
Il primo gruppo è formato da 40 persone, il secondo da 35 persone.
Si suppone che le popolazioni da cui sono estratti i campioni abbiano varianza
$\sigma_1^2 = 170, \quad \sigma_2^2 = 140$
Si vuole verificare se il nuovo farmaco è più efficace del vecchio, ossia se la pressione media del gruppo trattato con il nuovo farmaco è inferiore a quella del gruppo trattato con il vecchio, con livello di significatività del 5% .
I dati della pressione sanguigna massima dei due campioni sono riportati nella tabella 1

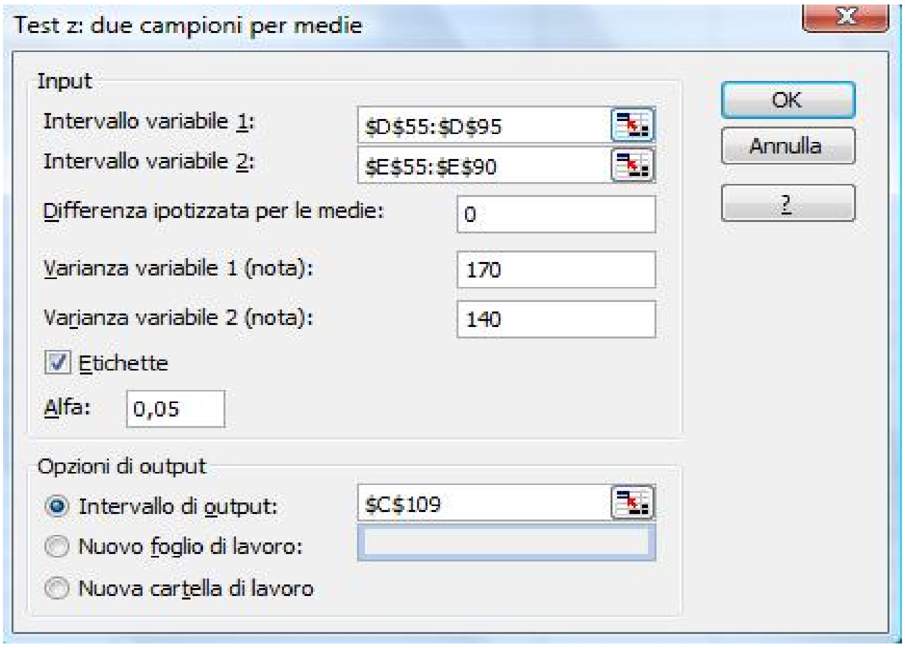
Tabella 1

nuovo farmaco	vecchio farmaco
143	158
154	150
168	172
160	158
140	155
151	168
135	171
145	153
165	181
147	174
146	184
157	154
120	180
146	185
148	175
152	169
132	162
146	173
162	156
134	186
149	174
139	160
160	158
169	158
153	176
129	164
161	179
153	173
159	171
155	174
152	171
149	144
157	173
165	155
163	169
141	
155	
113	
127	
134	

SUGGERIMENTI

Ipotesi nulla H0	media1 >= media2
Ipotesi alternativa H1	media1 < media2

Finestra di dialogo dello strumento Test z: due campioni per medie



Test z: due campioni per medie

	nuovo farmaco	*vecchio farmaco*
Media	148,35	167,51
Varianza nota	170	140
Osservazioni	40	35
Differenza ipotizzata per le medie	0	
z	-6,6721	
P(Z<=z) una coda	1,260E-11	
z critico una coda	1,645	
P(Z<=z) due code	2,521E-11	
z critico due code	1,960	

Lettura dell'output

Z	valore della statistica test Il valore di Z viene calcolato con la formula riportata all'inizio del foglio
P(Z<=z) una coda	valore del p-value per il test a una coda La decisione con il p-value segue le solite regole
z critico una coda	valore critico per il test a una coda La decisione può anche essere presa confrontando il valore della statistica test con il valore critico Il valore critico riportato nell'output è quello positivo, usato per il test a una coda (coda destra) Se si effettua un test a una coda (coda sinistra) il valore critico è il simmetrico (con segno negativo) di quello riportato nell'output
P(Z<=z) due code	valore del p-value per il test a due code
z critico due code	valori critici per il test a due code; i valori critici sono quello riportato nell'output e il suo simmetrico (negativo)

Conclusione: In questo esempio il p-value (una coda) è molto piccolo, perciò non si sbaglia rifiutando l'ipotesi nulla

Decisione	Rifiuto H0

La decisione può anche essere presa confrontando il valore della statistica test Z = -6,6721 con il valore critico z = -1,645; in questo caso il valore della statistica test appartiene alla regione di rifiuto, costituita dai valori minori del valore critico

Esercizio 73.2

Nella tabella 2 sono riportate le misure del peso in g di due campioni di 10 oggetti dello stesso tipo prodotti da due macchine diverse; gli oggetti sono scelti a caso da due popolazioni aventi entrambe la distribuzione normale, con varianze

$$\sigma_1^2 = 1{,}8 \qquad\qquad \sigma_2^2 = 1{,}3$$

Sottoporre a test l'ipotesi che le due popolazioni abbiano la stessa media con livello di significatività del 5%

Tabella 2

Campione 1	Campione 2
37,2	35,6
39,7	35
37,2	34,9
38,8	36
37,7	36,6
36,6	36,1
37,5	35,8
40,5	34,9
38,2	38,6
36,6	36,5

Ipotesi nulla H0	media1 = media2
Ipotesi alternativa H1	media1 ≠ media2

I due campioni sono piccoli, ma provengono da due popolazioni aventi la distribuzione normale, con varianze note, perciò il test Z può ancora essere utilizzato

Test z: due campioni per medie

	Campione 1	*Campione 2*
Media	38	36
Varianza nota	1,8	1,3
Osservazioni	10	10
Differenza ipotizzata per le medie	0	
z	3,5921	
P(Z<=z) una coda	0,00016	
z critico una coda	1,6449	
P(Z<=z) due code	0,00033	
z critico due code	1,9600	

Conclusione: Il p-value a due code è prossimo a 0, perciò non si sbaglia rifiutando l'ipotesi nulla. Le medie delle popolazioni sono diverse

Decisione	Rifiuto H0

Esercizio 73.3

Un docente universitario è interessato a confrontare quanto gli studenti imparano attraverso un corso tradizionale in aula rispetto a un corso on-line. Gli studenti di entrambi i corsi sostengono

un esame a fine corso e il docente raccoglie e analizza i dati della tabella 3, che riporta i voti conseguiti all'esame
Sottoporre a test l'ipotesi che i risultati conseguiti dagli studenti del corso tradizionale siano migliori. Sceglliere il livello di significatività del 5%
Si suppone che le varianze delle due popolazioni siano note

varianza corso in aula = 8
varianza corso on line = 10

Poiché i campioni sono grandi non è necessario ipotizzare che le popolazioni abbiano distribuzione normale

Tabella 3

corso in aula	corso on line
81	80
88	84
87	87
84	83
83	81
88	85
86	80
91	78
91	81
85	81
83	86
84	83
84	91
84	85
87	82
86	85
84	81
84	86
83	84
85	81
85	86
85	80
83	88
93	85
83	81
85	83
85	80
86	86
86	81
85	83
84	80
85	84
79	83
89	80
83	83
86	90
83	85
86	86
86	86
82	84

Ipotesi nulla H0	media1 <= media2
Ipotesi alternativa H1	media1 > media2

Test z: due campioni per medie

	corso in aula	*corso on line*
Media	85,18	83,45
Varianza nota	8	10
Osservazioni	40	40
Differenza ipotizzata per le medie	0	
z	2,5715	
P(Z<=z) una coda	0,00506	
z critico una coda	1,6449	
P(Z<=z) due code	0,01013	
z critico due code	1,9600	

Conclusione: Il p-value a una coda è prossimo a 0, perciò non si sbaglia rifiutando l'ipotesi nulla
I risultati degli studenti che seguono il corso in aula sono migliori

Decisione	Rifiuto H0

Torna su

Soluzione Esercizio 74

Test di ipotesi sulla differenza fra due medie
(varianze delle popolazioni incognite, varianze uguali)
Strumenti di Analisi: Test t, due campioni assumendo uguale varianza

Indice

Ritorna Esercizio 74

Test di ipotesi sulla differenza fra due medie - varianze incognite uguali
La statistica test per questo tipo di test si calcola con la formula seguente

$$T = \frac{\left(\overline{X}_1 - \overline{X}_2\right) - d}{\sqrt{S^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

S^2 stima congiunta della varianza, ottenuta con la formula

$$S^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

$S_1^2,\ S_2^2$ varianze dei due campioni

n_1 , n_2 ampiezze dei due campioni
d differenza ipotizzata fra le medie

Nel caso in cui le varianze delle due popolazioni non siano note, ma siano uguali,
si usa lo Strumento di Analisi
Test t: due campioni assumendo uguale varianza
Le popolazioni da cui provengono i due campioni devono essere normali

Esempio 74.1
Nella tabella 1 sono riportate le lunghezze in cm di due campioni di oggetti dello stesso tipo prodotti da due macchine diverse
Sottoporre a test l'ipotesi che gli oggetti prodotti abbiano lunghezza media significativamente diversa al livello di significatività 5%, supponendo che le popolazioni da cui provengono i campioni abbiano distribuzione normale con la stessa varianza

Tabella 1

Campione 1	Campione 2
8,26	7,95
8,13	7,89
8,35	7,9
8,07	8,14
8,34	7,92
	7,84
	7,94

SUGGERIMENTI

Si effettua un test a due code

Ipotesi nulla H0	media1 = media2
Ipotesi alternativa H1	media1 ≠ media2

Test t: due campioni assumendo uguale varianza

	Campione 1	*Campione 2*
Media	8,23	7,94
Varianza	0,01575	0,0091
Osservazioni	5	7
Varianza complessiva	0,01176	
Differenza ipotizzata per le medie	0	
gdl	10	
Stat t	4,5671	
P(T<=t) una coda	0,0005	
t critico una coda	1,8125	
P(T<=t) due code	0,0010	
t critico due code	2,2281	

Lettura dell'output

Varianza complessiva	stima congiunta della varianza ottenuta con la formula sopra riportata
Stat t	valore della statistica test t Il valore di t viene calcolato con la formula riportata all'inizio del foglio
P(T<=t) una coda	valore del p-value per il test a una coda In questo esempio il p-value è prossimo a 0, quindi non si sbaglia rifiutando l'ipotesi nulla
t critico una coda	valore critico per il test a una coda La decisione può essere presa anche confrontando il valore della statistica test con il valore critico Il valore critico riportato nell'output è quello positivo, usato per il test a una coda (coda destra) Se si effettua un test a una coda (coda sinistra) il valore critico è il simmetrico (segno negativo) di quello riportato nell'output
P(T<=t) due code	valore del p-value per il test a due code La decisione con il p-value segue le solite regole
t critico due code	valori critici per il test a due code I valori critici sono quello riportato nell'output e il suo simmetrico (negativo)

Conclusione: il p-value a due code è prossimo a 0, quindi non si sbaglia rifiutando l'ipotesi nulla. Le lunghezze medie sono uguali

Decisione	Rifiuto H0

Esercizio 74.2

Una banca vuole migliorare il servizio alla clientela tra le 12 e le 13 nelle sue filiali
Per una settimana viene rilevato il tempo di attesa in minuti di ciascun cliente in due filiali, una in un quartiere commerciale e l'altra in un quartiere residenziale
Per due campioni di 15 rilevazioni si ottengono i tempi della tabella 2

Tabella 2

Quartiere commerciale	Quartiere residenziale
4,21	9,66
5,55	5,9
3,02	8,02
5,13	5,79
4,77	8,73
2,34	3,82
3,54	8,01
3,2	8,35
4,5	10,49
6,1	6,68
0,38	5,64
5,12	4,08
6,46	6,17
6,19	9,91
3,79	5,47

Si assume che le varianze dei tempi di attesa siano uguali per entrambe le filiali e che i tempi di attesa si distribuiscano secondo la distribuzione normale
Verificare se c'è una differenza significativa fra i tempi medi di attesa nelle due filiali, con livello di significatività 5%

Si effettua un test a due code

Ipotesi nulla H0	media1 = media2
Ipotesi alternativa H1	media1 ≠ media2

Test t: due campioni assumendo uguale varianza

	Quartiere commerciale	*Quartiere residenziale*
Media	4,2867	7,1147
Varianza	2,6830	4,3355
Osservazioni	15	15
Varianza complessiva	3,5093	
Differenza ipotizzata per le medie	0	
gdl	28	
Stat t	-4,1343	
P(T<=t) una coda	0,0001	
t critico una coda	1,7011	
P(T<=t) due code	0,0003	
t critico due code	2,0484	

Conclusioni:
Il valore del p-value 0,0003 è prossimo a 0, perciò non si sbaglia rifiutando l'ipotesi nulla
La regione di rifiuto per il test a due code è formata dai valori maggiori del valore t critico due code (2,0484) e minori del valore t critico simmetrico del precedente (-2,0484);
il valore della statistica test t è minore del valore critico negativo, perciò al livello di significatività 5% si rifiuta l'ipotesi nulla e si conclude che i tempi medi d'attesa non sono uguali

Decisione	Rifiuto H0

Esercizio 74.3

Due campioni rispettivamente di 10 automobilisti di Torino e di 8 automobilisti di Asti consumano in un mese le quantità di benzina (in litri) riportate nella tabella 3

Tabella 3

Torino	Asti
55	42
55	39
46	36
54	41
57	38
50	42
52	46
47	44
53	
51	

Sottoporre a test l'ipotesi che la differenza fra i consumi medi di torinesi e astigiani sia diversa dal valore d=10 litri di benzina al mese
Supporre che le popolazioni da cui sono tratti i campioni abbiano distribuzione normale con uguale varianza. Scegliere il livello di significatività 5%

Si effettua un test a due code

Ipotesi nulla H0	media1 – media2 = 10
Ipotesi alternativa H1	media1 – media2 ≠ 10

Test t: due campioni assumendo uguale varianza

	Torino	*Asti*
Media	52	41
Varianza	12,6667	10,5714
Osservazioni	10	8
Varianza complessiva	11,75	
Differenza ipotizzata per le medie	10	
gdl	16	
Stat t	0,615021	
P(T<=t) una coda	0,273595	
t critico una coda	1,745884	
P(T<=t) due code	0,547189	
t critico due code	2,119905	

Conclusioni:

Il valore del p-value per il test a due code indica che non si sbaglia non rifiutando l'ipotesi nulla
La stessa conclusione si trae confrontando il valore della statistica test (0,6150)
con i valori critici (–2,1199 e 2,1199)

Decisione	Non rifiuto H0

Esercizio 74.4

Si pesano due campioni di 8 pompelmi gialli e di 10 pompelmi rosa: i pesi sono riportati nella tabella 4

Stabilire con un test al livello di significatività 5% se c'è differenza fra i pesi medi dei due tipi di frutti

Supporre che le varianze siano uguali e che le popolazioni siano normali

Tabella 4

Pompelmi gialli	Pompelmi rosa
241	220
204	185
224	203
214	213
209	215
215	202
247	207
219	205
	215
	211

Si effettua un test a due code

Ipotesi nulla H0	media1 = media2
Ipotesi alternativa H1	media1 ≠ media2

Test t: due campioni assumendo uguale varianza

	Pompelmi gialli	*Pompelmi rosa*
Media	221,625	207,6
Varianza	229,125	97,156
Osservazioni	8	10
Varianza complessiva	154,892	
Differenza ipotizzata per le medie	0	
gdl	16	
Stat t	2,3757	
P(T<=t) una coda	0,0152	
t critico una coda	1,7459	
P(T<=t) due code	0,0303	
t critico due code	2,1199	

Conclusioni:

Il valore del p-value 0,0303 è compreso fra i valori classici del livello di significatività, perciò si tratta di un caso critico

Decisione	Caso critico

Le conclusioni si possono anche trarre senza servirsi del p-value, ma in questo modo occorre effettuare il test ai due livelli di significatività consueti del 5% e dell'1% e poi confrontare i risultati ottenuti (questo secondo modo è meno veloce).

Al livello di significatività alfa=0,05 si utilizzano i risultati appena ottenuti

La regione di rifiuto per il test a due code è formata dai valori maggiori del valore t critico due code (2,1199) e minori del valore t critico simmetrico del precedente (-2,1199);

il valore della statistica test t è maggiore del valore critico a destra (2,1199), perciò al livello di significatività 5% si rifiuta l'ipotesi nulla
Ripetiamo ora il test per il livello di significatività 1%

Test t: due campioni assumendo uguale varianza

	Pompelmi gialli	*Pompelmi rosa*
Media	221,625	207,6
Varianza	229,125	97,156
Osservazioni	8	10
Varianza complessiva	154,892	
Differenza ipotizzata per le medie	0	
gdl	16	
Stat t	2,3757	
P(T<=t) una coda	0,0152	
t critico una coda	2,5835	
P(T<=t) due code	0,0303	
t critico due code	2,9208	

Conclusioni:
Al livello di significatività dell'1% a regione di rifiuto per il test a due code è formata dai valori maggiori del valore t critico due code (2,9208) e minori del valore t critico simmetrico del precedente (-2,9208); il valore della statistica test t è compreso fra i valori critici, perciò al livello di significatività 1% non si rifiuta l'ipotesi nulla
Confrontando le conclusioni ai due livelli di significatività si ha conferma di essere in presenza di un caso critico, che richiede ulteriori indagini e campioni di maggior ampiezza

Torna su

Soluzione Esercizio 75
Test di ipotesi sul rapporto fra due varianze
Strumenti di Analisi: Test F a due campioni per varianze

Indice

Ritorna Esercizio 75

Test di ipotesi sul rapporto fra due varianze
Si utilizza questo test per la verifica dell'ipotesi che due popolazioni abbiano la stessa varianza
Questo test è utile per stabilire se il test t per l'uguaglianza fra le medie di due popolazioni
con la stessa varianza possa essere applicato
Il test presuppone che i campioni siano estratti da popolazioni con distribuzione normale
e viene eseguito calcolando il rapporto fra le varianze: se le popolazioni hanno la stessa
varianza, ci si attende che il rapporto fra le varianze sia uguale a 1
Il test è di solito a due code, l'ipotesi alternativa è che le varianze siano diverse
Per eseguire il test viene utilizzata la distribuzione F
La statistica test per questo tipo di test si calcola con la formula seguente

$$\boxed{F = \frac{s_1^2}{s_2^2}}$$

s_1^2 , s_2^2 varianze dei due campioni

Il test può essere eseguito con lo strumento di Analisi **Test F a due campioni per varianze**

Esempio 75.1
Nella tabella 1 sono riportate le lunghezze in cm di due campioni A e B di oggetti dello stesso
tipo prodotti da due macchine diverse (Vedi Esercizio 74, Esempio 74.1)
Sottoporre a test l'ipotesi che le due popolazioni da cui provengono i campioni abbiano la stessa
varianza

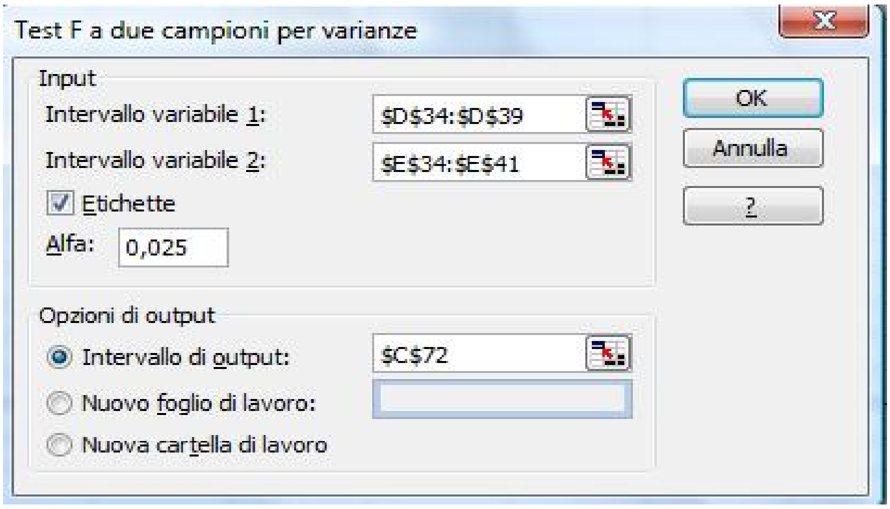
Tabella 1

Campione 1	Campione 2
8,26	7,95
8,13	7,89
8,35	7,9
8,07	8,14
8,34	7,92
	7,84
	7,94

SUGGERIMENTI
Si effettua un test a due code

Ipotesi nulla H0	varianze uguali
Ipotesi alternativa H1	varianze diverse

Per effettuare il test a due code con livello di significatività 5%, nella finestra di dialogo
dello strumento di analisi **Test F a due campioni per varianze** introdurre nella casella Alfa
il valore 0,025



Test F a due campioni per varianze

	Campione 1	*Campione 2*
Media	8,23	7,94
Varianza	0,01575	0,0091
Osservazioni	5	7
gdl	4	6
F	1,7308	
P(F<=f) una coda	0,2609	
F crtitico una coda	6,2272	

Lettura dell'output

Media	media dei due campioni
Varianza	varianza dei due campioni
Osservazioni	numero di dati dei due campioni
gdl	gradi di libertà del numeratore e del denominatore
F	valore della statistica test
	Il valore di F viene calcolato con la formula riportata all'inizio del foglio
P(F<=f) una coda	valore del p-value per il test a una coda (coda destra)
F critico una coda	valore critico per il test a una coda

Il valore del p-value per il test a due code è uguale al doppio di quello a una coda
In questo esempio si ha quindi

p-value (due code)	0,5219

Il p-value ha un valore elevato, perciò non si sbaglia non rifiutando l'ipotesi nulla

Decisione	Non rifiuto H0

Conclusione Le due varianze sono uguali

Esercizio 75.2
(Vedere Esercizio 74.2)
Una banca vuole migliorare il servizio alla clientela tra le 12 e le 13 nelle sue filiali
Per una settimana viene rilevato il tempo di attesa in minuti di ciascun cliente in due filiali, una in un quartiere commerciale e l'altra in un quartiere residenziale
Per due campioni di 15 rilevazioni si ottengono i tempi della tabella 2

Tabella 2

Quartiere commerciale	Quartiere residenziale
4,21	9,66
5,55	5,9
3,02	8,02
5,13	5,79
4,77	8,73
2,34	3,82
3,54	8,01
3,2	8,35
4,5	10,49
6,1	6,68
0,38	5,64
5,12	4,08
6,46	6,17
6,19	9,91
3,79	5,47

Per poter effettuare il test sulla differenza fra i tempi medi di attesa nelle due filiali (Vedere Esercizio 74.2) occorre assumere che le varianze dei tempi di attesa siano uguali per entrambe le filiali: verificare con il test F se l'ipotesi è soddisfatta, con livello di significatività del 5%

Si effettua un test a due code
Per il livello di significatività del 5%, nella finestra di dialogo per Alfa inserire 0,025

Ipotesi nulla H0	varianze uguali
Ipotesi alternativa H1	varianze diverse

Test F a due campioni per varianze

	Quartiere commerciale	*Quartiere residenziale*
Media	4,2867	7,1147
Varianza	2,6830	4,3355
Osservazioni	15	15
gdl	14	14
F	0,6188	
P(F<=f) una coda	0,1900	
F crtitico una coda	0,3357	

p-value (due code)	0,3800

Conclusione: Il p-value ha un valore elevato, perciò non si sbaglia non rifiutando l'ipotesi nulla. Le due varianze sono uguali

Decisione	Non rifiuto H0

Esercizio 75.3

(Vedere Esercizio 74.3)
Due campioni rispettivamente di 10 automobilisti di Torino e di 8 automobilisti di Asti consumano in un mese le quantità di benzina (in litri) riportate nella tabella 3

Tabella 3

Torino	Asti
55	42
55	39
46	36
54	41
57	38
50	42
52	46
47	44
53	
51	

Per poter stabilire con un se la differenza fra i consumi medi di torinesi e astigiani sia diversa dal valore d=10 litri di benzina al mese, verificare che le popolazioni da cui sono tratti i campioni abbiano la stessa varianza
Scegliere il livello di significatività 5% (nella finestra di dialogo per Alfa inserire 0,025)

Si effettua un test a due code

Ipotesi nulla H0	varianze uguali
Ipotesi alternativa H1	varianze diverse

Test F a due campioni per varianze

	Torino	*Asti*
Media	52	41
Varianza	12,6667	10,5714
Osservazioni	10	8
gdl	9	7
F	1,1982	
P(F<=f) una coda	0,4150	
F crtitico una coda	4,8232	

p-value (due code)	0,8300

Conclusione: Il p-value ha un valore elevato, perciò non si sbaglia non rifiutando l'ipotesi nulla. Le due varianze sono uguali

Decisione	Non rifiuto H0

Esercizio 75.4
(Vedere Esercizio 74.4)
Si pesano due campioni di 8 pompelmi gialli e di 10 pompelmi rosa: i pesi sono riportati nella tabella 4
Per poter stabilire con un test se c'è differenza fra i pesi medi dei due tipi di frutti verificare che le varianze sono uguali

Tabella 4

Pompelmi gialli	Pompelmi rosa
241	220
204	185
224	203

214	213
209	215
215	202
247	207
219	205
	215
	211

Si effettua un test a due code

Ipotesi nulla H0	varianze uguali
Ipotesi alternativa H1	varianze diverse

Test F a due campioni per varianze

	Pompelmi gialli	Pompelmi rosa
Media	221,625	207,6
Varianza	229,125	97,1556
Osservazioni	8	10
gdl	7	9
F	2,3583	
P(F<=f) una coda	0,1150	
F crtitico una coda	4,1970	

p-value (due code)	0,2301

Conclusione: Il p-value ha un valore elevato, perciò non si sbaglia non rifiutando l'ipotesi nulla. Le due varianze sono uguali

Decisione	Non rifiuto H0

Torna su

8. TEST CHI-QUADRO

Soluzione Esercizio 76
Test chi quadro di adattamento

Indice

Ritorna Esercizio 76

Con il **test chi quadro di adattamento** si stabilisce se i dati osservati di un campione provengono da una assegnata distribuzione di una popolazione.
Per un test chi quadro di adattamento le ipotesi sono così formulate

Ipotesi nulla H0 i dati provengono da una popolazione con una specificata distribuzione di probabilità

Ipotesi alternativa H1 i dati non provengono da una popolazione con la distribuzione di probabilità specificata

Il procedimento consiste nel confrontare la distribuzione delle frequenze osservate con la distribuzione delle frequenze attese
Le **frequenze osservate** si definiscono come il numero reale di osservazioni relative a ogni classe in una distribuzione di frequenza
Le **frequenze attese** si definiscono come il numero di osservazioni che si avrebbero per ogni classe se i dati del campione si distribuissero secondo la distribuzione ipotizzata
Per valutare l'adattamento delle frequenze osservate alle frequenze attese si utilizza la **statistica test chi quadro** calcolata con la formula

$$\boxed{\chi^2 = \sum_{i=1}^{k} \frac{(O_i - A_i)^2}{A_i}}$$

Oi frequenze osservate
Ai frequenze attese
k numero delle classi nella distribuzione di frequenza

Per la correttezza del test le frequenze attese devono essere maggiori di 5
La statistica test ha approssimativamente la distribuzione chi quadro con grado di libertà

$$\boxed{\nu = k - 1 - m}$$

dove m è il numero di parametri della distribuzione teorica, stimati servendosi dei dati del campione

La decisione è basata sul confronto fra la statistica test e il valore critico χ^2_α
dove alfa è il **livello di significatività**, di solito uguale a 0,05 oppure 0,01
La **regione di rifiuto** è data dai valori tali che

$$\boxed{\chi^2 > \chi^2_\alpha}$$

Esempio 76.1
Alle ultime elezioni amministrative in un comune si sono presentate quattro liste che hanno ottenuto le seguenti percentuali

Lista	Percentuale
1	26%
2	32%
3	15%
4	27%
Totale	100%

Nella Sezione elettorale A del comune i 350 voti validi sono risultati così suddivisi

Lista	Voti
1	80
2	120
3	60
4	90
Totale	350

Nella Sezione elettorale B invece i 320 voti validi sono risultati così suddivis

Lista	Voti
1	65
2	120
3	40
4	95
Totale	320

I risultati delle due sezioni si adattano bene ai risultati complessivi oppure le differenze sono statisticamente rilevanti?

SUGGERIMENTI

Le frequenze attese si calcolano moltiplicando la frequenza percentuale di ciascuna lista per il numero corrispondente di voti della lista

Sezione A

Ipotesi nulla H0	c'è adattamento
Ipotesi alternativa H1	non c'è adattamento

Calcolo delle frequenze attese e della statistica test chi quadro

Lista	Frequenze osservate O	Frequenze attese A	(O-A)^2/A
1	80	91	1,330
2	120	112	0,571
3	60	52,5	1,071
4	90	94,5	0,214
	350	350	**3,187**

chi quadro

In questo caso il numero di classi è k=4, e nessun parametro della popolazione è stimato dai dati del campione, quindi m=0 e il grado di libertà è uguale a 3

Grado di libertà	3	
Livello di significatività	0,05	
Regione di rifiuto:	**chi quadro alfa**	7,8147

=INV.CHI(D97;D96)

Conclusione:

Il valore della statistica test chi quadro calcolato non appartiene alla regione di rifiuto, perciò non si può rifiutare l'ipotesi nulla.

I dati del campione permettono di concludere che non c'è una differenza statisticamente rilevante: la sezione A si adatta ai risultati generali.

Sezione B

Ipotesi nulla H0	c'è adattamento
Ipotesi alternativa H1	non c'è adattamento

Calcolo delle frequenze attese e della statistica test chi quadro

Lista	Frequenze osservate O	Frequenze attese A	(O-A)^2/A
1	65	83,2	3,981
2	120	102,4	3,025
3	40	48	1,333
4	95	86,4	0,856
	320	320	**9,196**

chi quadro

Gradi di libertà	3	
Livello di significatività	0,05	
Regione di rifiuto.	**chi quadro alfa**	7,8147

=INV.CHI(D120;D119)

Conclusione:
Il valore della statistica test chi quadro calcolato appartiene alla regione di rifiuto, perciò si rifiuta l'ipotesi nulla.
I dati del campione permettono di concludere che c'è una differenza statisticamente rilevante: la sezione B non si adatta ai risultati generali.

Esercizio 76.2
Nelle classi prime dei licei scientifici di una città i risultati degli scrutini finali sono i seguenti

Promossi con media superiore a 7/10	18%
Promossi con media inferiore a 7/10	38%
Promossi con debito	37%
Non Promossi	7%

Nelle classi prime del liceo scientifico Newton si hanno i seguenti risultati:

Promossi con media superiore a 7/10	24
Promossi con media inferiore a 7/10	35
Promossi con debito	38
Non Promossi	18

Stabilire se i risultati delle prime del liceo Newton si adattano a quelli generali con un livello di significatività del 5%

Ipotesi nulla H0	c'è adattamento
Ipotesi alternativa H1	non c'è adattamento

Calcolo delle frequenze attese e della statistica test chi quadro

Lista	Frequenze osservate O	Frequenze attese A	(O-A)^2/A
Promossi media > 7/10	24	20,7	0,526
Promossi media < 7/10	35	43,7	1,732
Promossi con debito	38	42,55	0,487
Non Promossi	18	8,05	12,298
	115	115	**15,043**

chi quadro

Gradi di libertà	3	
Livello di significatività	0,05	
Regione di rifiuto:	**chi quadro alfa**	7,8147

Conclusione:
Il valore della statistica test chi quadro calcolato appartiene alla regione di rifiuto, perciò si rifiuta l'ipotesi nulla.
I dati del campione permettono di concludere che c'è una differenza statisticamente rilevante: i risultati delle classi prime del liceo Newton non si adattano ai risultati generali.

Torna su

Soluzione Esercizio 77

Test chi quadro di adattamento. Calcolo del p-value

Indice

Ritorna Esercizio 77

Esercizio 77.1

I tipi di specializzazione post-laurea in medicina sono suddivisi in quattro categorie.
La tabella 1 mostra la distribuzione di frequenza percentuale dei medici di ciascuna categoria nel 1990

Tabella 1

Specializzazione	Frequenza percentuale
A	14,1%
B	30,6%
C	26,4%
D	28,9%
	100,0%

Si vuole stabilire se nel 2006 la distribuzione è cambiata.
Un campione di 500 medici specialisti mostra la seguente distribuzione di frequenza (Tabella 2)

Tabella 2

Specializzazione	Frequenza assoluta
A	62
B	172
C	130
D	136
	500

La distribuzione è rimasta invariata o è cambiata?

Ipotesi nulla H0	c'è adattamento
Ipotesi alternativa H1	non c'è adattamento

Calcolo delle frequenze attese e della statistica test chi quadro

Tabella 3

Specializzazione	Frequenze osservate O	Frequenze attese A	(O-A)^2/A
A	62	70,5	1,0248
B	172	153	2,3595
C	130	132	0,0303
D	136	144,5	0,5000
	500	500	**3,9146**

chi quadro

In questo caso il numero di classi è k=4, e nessun parametro della popolazione è stimato dai dati del campione, quindi m=0 e il grado di libertà è uguale a 3

Gradi di libertà	3	
Livello di significatività	0,05	
Regione di rifiuto:	chi quadro alfa	7,815

Conclusione:
Il valore della statistica test chi quadro calcolato non appartiene alla regione di rifiuto, perciò non si può rifiutare l'ipotesi nulla.
I dati del campione non permettono di concludere che la distribuzione sia cambiata

SUGGERIMENTI

Con Excel si può fare il test di adattamento con la funzione TEST.CHI che restituisce in uscita il p-value, ossia il più piccolo livello di significatività (calcolato a partire dalle tabelle di frequenze osservate e frequenze attese) per il quale si può rifiutare l'ipotesi nulla.
Un p-value molto prossimo a 0 indica che la probabilità di sbagliare rifiutando l'ipotesi nulla è molto vicina a 0, ossia si è praticamente certi di non sbagliare rifiutando l'ipotesi nulla
Un p-value vicino ai classici livelli di significatività indica che la decisione è critica.
Un p-value maggiore indica che non si sbaglia non rifiutando l'ipotesi nulla.

La funzione TEST.CHI utilizza i valori delle frequenze osservate e delle frequenze attese, già calcolati nella tabella 3
Sintassi
TEST.CHI(**int_effettivo**;**int_previsto**)
int_effettivo: intervallo delle frequenze osservate
int_previsto: intervallo dele frequenze teoriche attese

Riferendoci all'esercizio precedente, riportiamo la tabella delle frequenze osservate e attese e calcoliamo il p-value

Specializzazione	Frequenze osservate O	Frequenze attese A
A	62	70,5
B	172	153
C	130	132
D	136	144,5
	500	500

p-value	**0,2708**

Conclusione:
In questo caso il valore del p-value è grande, rispetto ai solito livelli di significatività, perciò non si sbaglia non rifiutando l'ipotesi nulla.

Esercizio 77.2

E' noto per esperienza che il livello di qualità di ciascun oggetto prodotto da una certa macchina è descritto dalla tabella 4

Tabella 4

Livello di qualità	Percentuale
Di altissima qualità	38%
Di alta qualità	32%
Di media qualità	26%
Di bassa qualità	4%
	100%

Una nuova macchina, progettata per la stessa produzione, ha prodotto 500 oggetti con i risultati riportati nella tabella 5

Tabella 5

Livello di qualità	Percentuale
Di altissima qualità	222
Di alta qualità	171
Di media qualità	98
Di bassa qualità	9
	500

La differenza è dovuta solo al caso? Utilizzare il p-value per la decisione

Ipotesi nulla H0	c'è adattamento
Ipotesi alternativa H1	non c'è adattamento

Calcolo delle frequenze attese

Livello di qualità	Frequenze osservate O	Frequenze attese A
Di altissima qualità	222	190
Di alta qualità	171	160
Di media qualità	98	130
Di bassa qualità	9	20
	500	500

Calcolo del p-value con la funzione TEST.CHI

p-value	0,0001640

Conclusione: Il p-value è prossimo a 0, perciò si rifiuta l'ipotesi nulla: la differenza non è dovuta al caso.

Torna su

Soluzione Esercizio 78
Test chi quadro di adattamento alla distribuzione uniforme discreta

Indice

Ritorna Esercizio 78

Esercizio 78.1

Per provare l'ipotesi che un dado non sia truccato si effettuano 600 lanci e si osservano le seguenti uscite (Tabella 1)

Tabella 1

numero uscito	frequenza
1	95
2	115
3	105
4	90
5	115
6	80
Totale	600

Stabilire con il test di adattamento alla distribuzione uniforme discreta, se il dado è truccato
Provare a cambiare i valori delle uscite osservate e vedere come cambia la conclusione del test.

Ipotesi nulla H0	Il dado non è truccato
Ipotesi alternativa H1	Il dado è truccato

Calcolo delle frequenze attese e della statistica test chi quadro

numero uscito	frequenze osservate O	frequenze attese A	(O-A)^2/A
1	95	100	0,25
2	115	100	2,25
3	105	100	0,25
4	90	100	1
5	115	100	2,25
6	80	100	4
totale	600	600	**10**

chi quadro

In questo caso il numero di classi è k=6, e nessun parametro della popolazione è stimato dai dati del campione, quindi m=0 e il grado di libertà è uguale a 5

Gradi di libertà	5
Livello di significatività	0,05
Regione di rifiuto	11,070

Conclusione:

Il valore della statistica test chi quadro calcolato non appartiene alla regione di rifiuto, perciò non si può rifiutare l'ipotesi nulla; si conclude che il dado non è truccato.

Soluzione con la funzione TEST.CHI

p-value	**0,0752**

Conclusione:
Il valore del p-value è sufficientemente grande, perciò non si può rifiutare l'ipotesi nulla.
Il dado non è truccato

Esercizio 78.2

Un dado apparentemente non truccato viene lanciato 120 volte, ottenendo le frequenze della tabella 2
Si osserva che in questo campione di risultati i numeri 2 e 4 escono con maggior frequenza.
Stabilire con un test di adattamento se il dado è truccato
Effettuare il test per i livelli di significatività alfa=0,05 e alfa=0,01.

Tabella 2

numero uscito	frequenza
1	10
2	32
3	16
4	34
5	16
6	12
Totale	120

Ipotesi nulla H0	Il dado non è truccato
Ipotesi alternativa H1	Il dado è truccato

Calcolo delle frequenze attese e della statistica test chi quadro

numero uscito	frequenze osservate O	frequenze attese A	(O-A)^2/A
1	10	20	5
2	32	20	7,2
3	16	20	0,8
4	34	20	9,8
5	16	20	0,8
6	12	20	3,2
totale	120	120	**26,8**

chi quadro

In questo caso il numero di classi è k=6, e nessun parametro della popolazione è stimato dai dati del campione, quindi m=0 e il grado di libertà è uguale a 5

Gradi di libertà	5	
Livello di significatività	0,05	0,01
Regione di rifiuto	11,070	15,086

Conclusione:
Il valore della statistica test chi quadro calcolato appartiene alla regione di rifiuto per entrambi i livelli di significatività, perciò si può rifiutare l'ipotesi nulla; si conclude che il dado è truccato.

Soluzione con la funzione TEST.CHI

p-value	**0,0001**

Conclusione:
Il valore del p-value è prossimo a zero, perciò non si sbaglia rifiutando l'ipotesi nulla.
Il dado è truccato

Torna su

Soluzione Esercizio 79
Test chi quadro di adattamento alla distribuzione binomiale

Indice

Ritorna Esercizio 79

Esercizio 79.1
Sono state lanciate 2000 volte cinque monete; il numero di teste in ciascuno dei lanci ha avuto le frequenze riportate nella tabella 1

Tabella 1

numero di teste	frequenze osservate
0	59
1	316
2	596
3	633
4	320
5	76
Totale	2000

Se le monete sono tutte eque, le probabilità teoriche che in un singolo lancio si ottenga un dato numero di teste sono date da una distribuzione binomiale di parametri

$n = 5 \qquad p = 0,5$

Stabilire con il test chi quadro se le monete sono eque.

SUGGERIMENTI
Le probabilità teoriche possono essere calcolate con la distribuzione binomiale usando la funzione DISTRIB.BINOM

Per calcolare le frequenze attese, ossia le frequenze che si avrebbero in 2000 lanci, se le monete fossero eque, si moltiplica ogni probabilità teorica per il numero totale dei lanci

Ipotesi nulla H0	le monete sono eque
Ipotesi alternativa H1	le monete non sono eque

n	5
p	0,5

Calcolo delle frequenze attese e della statistica test chi quadro

numero di teste	frequenze osservate O	Probabilità (distribuzione binomiale)	frequenze attese A	(O-A)^2/A
0	59	0,03125	63	0,196
1	316	0,15625	313	0,039
2	596	0,3125	625	1,346
3	633	0,3125	625	0,102
4	320	0,15625	313	0,180
5	76	0,03125	63	2,916
totale	2000	1	2000	**4,779**

In questo caso il numero di classi è k=6, e nessun parametro della popolazione è stimato dai dati del campione, quindi m=0 e il grado di libertà è uguale a 5

Grado di libertà	5
Livello di significatività	0,05
Regione di rifiuto	11,0705

Conclusione:
Il valore della statistica test chi quadro calcolato non appartiene alla regione di rifiuto, perciò non si può rifiutare l'ipotesi nulla; si conclude che le monete sono eque.

Soluzione con la funzione TEST.CHI

p-value	**0,4434**

Conclusione:
Il valore del p-value è sufficientemente grande, perciò non si sbaglia non rifiutando l'ipotesi nulla.

Esercizio 79.2
Una compagnia di assicurazioni ritiene che il numero di automobilisti che indossano la cintura di sicurezza sia una variabile binomiale con p = 0,70.
Per verificare questa assunzione si istituiscono dei posti di controllo e si verificano a campione 10 automobilisti ogni due ore. Raccogliendo i dati di 1000 campioni di 10 guidatori si ottiene la distribuzione di frequenza della Tabella 2

N° guidatori con cintura	frequenze osservate
0	0
1	0
2	1
3	1
4	7
5	28
6	95
7	215
8	300
9	258
10	95
	1000

Stabilire se i dati osservati si distribuiscono secondo una distribuzione binomiale con parametri n = 10 e p = 0,80. utilizzare il p-value per la decisione.

Ipotesi nulla H0	c'è adattamento
Ipotesi alternativa H1	non c'è adattamento

Calcolo delle frequenze attese e della statistica test chi quadro
Parametri distribuzione binomiale

n	10
p	0,8

N° guidatori con cintura	**frequenze osservate O**	**Probabilità (distribuzione binomiale)**	**frequenze attese A**
0	0	0,00000	0
1	0	0,00000	0
2	1	0,00007	0
3	1	0,00079	1
4	7	0,00551	6
5	28	0,02642	26
6	95	0,08808	88
7	215	0,20133	201
8	300	0,30199	302
9	258	0,26844	268
10	95	0,10737	107
totale	1000	1,00000	1000

Soluzione con la funzione TEST.CHI

p-value	0,11434

Conclusione:
Il valore del p-value è sufficientemente grande, perciò non si sbaglia non rifiutando l'ipotesi nulla. i dati si distribuiscono secondo la distribuzione binomiale

Torna su

Soluzione Esercizio 80
Test chi quadro di adattamento alla distribuzione normale

Indice

Ritorna Esercizio 80

Esempio 80.1

Si vuole sottoporre a test il fatto che le età dei residenti di una certa regione abbiano una distribuzione approssimativamente normale.
Si esamina un campione casuale di 500 residenti nella regione e si ottengono i seguenti valori di media e scarto quadratico medio, calcolati dal campione (tabella 1)

Tabella 1

età media (in anni)	34,96
scarto quadratico medio	21,82

e la seguente distribuzione di frequenza (tabella 2)

Tabella 2

Classi	Freq. assolute osservate
x<=15	91
15<x<=25	105
25<x<=35	84
35<x<=45	65
45<x<=55	55
55<x<=65	46
x>65	54
	500

Ipotesi nulla H0	Le età sono distribuite in modo approx. normale
Ipotesi alternativa H1	Le età non sono distribuite normalmente

Calcolo delle frequenze attese e del valore della statistica test chi quadro
Usare come valor medio e come scarto quadratico medio della distribuzione normale
i valori delle statistiche ottenuti dal campione (quindi, nel calcolo del grado di libertà, m=2)
Ricordare che per la correttezza del test le frequenze attese devono essere maggiori di 5;
in caso contrario occorre accorpare delle classi adiacenti (vedi esempio 80.4)

Classi	Estremi destri classi	Frequenze assolute osservate O	Probabilità (distribuzione normale)	Frequenze attese A	(O-A)^2/A
x<=15	15	91	0,1802	90,08	0,0094
15<x<=25	25	105	0,1439	71,93	15,1987
25<x<=35	35	84	0,1767	88,35	0,2143
35<x<=45	45	65	0,1766	88,28	6,1383
45<x<=55	55	55	0,1435	71,76	3,9131
55<x<=65	65	46	0,0949	47,45	0,0443
x>65		54	0,0843	42,15	3,3316
	totali	500	1,0000	500	**28,8497**

chi quadro

Grado di libertà	4
Livello di significatività	0,05
Regione di rifiuto	9,488

Conclusione:
Il valore delle statistica test chi quadro appartiene alla regione di rifiuto, perciò si può rifiutare l'ipotesi nulla: i dati del campione permettono di concludere che la distribuzione non è normale

Soluzione con la funzione TEST.CHI

p-value	**6,49535E-05**

Conclusione: Il valore del p-value è molto piccolo perciò non si sbaglia rifiutando l'ipotesi nulla.

Esercizio 80.2

Nella tabella di distribuzione di frequenza sono raccolti i dati di un campione di misure della lunghezza in cm di 100 sbarrette di metallo (Tabella 3)

Tabella 3

classi	**frequenza assoluta**
20<x<=30	7
30<x<=40	16
40<x<=50	25
50<x<=60	27
60<x<=70	17
70<x<=80	8

Il valor medio e lo scarto quadratico medio del campione sono riportati nella tabella 4

Tabella 4

valor medio (in cm)	51,1
scarto quadratico medio	12,85

Verificare l'adattamento dei dati a una distribuzione normale

Ipotesi nulla H0	I dati sono distribuiti in modo approx. normale
Ipotesi alternativa H1	I dati non sono distribuiti normalmente

Calcolo delle frequenze attese e del valore della statistica test chi quadro
Usare come valor medio e come scarto quadratico medio della distribuzione normale i valori delle statistiche ottenuti dal campione (quindi m=2)

classi	**estremi destri classi**	**Frequenze assolute osservate O**	**Probabilità (distribuzione normale)**	**Frequenze attese A**	**(O-A)^2/A**
20<x<=30	30	7	0,0503	5,03	0,7722
30<x<=40	40	16	0,1436	14,36	0,1884
40<x<=50	50	25	0,2720	27,20	0,1787
50<x<=60	60	27	0,2898	28,98	0,1357
60<x<=70	70	17	0,1736	17,36	0,0075
70<x<=80		8	0,0707	7,07	0,1232
		100	1,0000	100	**1,4057**

chi quadro

Grado di libertà	3
Livello di significatività	0,05
Regione di rifiuto	7,815

Conclusione:
Il valore delle statistica test chi quadro non appartiene alla regione di rifiuto, perciò non si può rifiutare l'ipotesi nulla: i dati del campione permettono di concludere che la distribuzione è approssimativamente normale

Soluzione con la funzione TEST.CHI

p-value	**0,924**

Conclusione: Il valore del p-value è grande perciò non si sbaglia non rifiutando l'ipotesi nulla.

Esercizio 80.3

Nella tabella di distribuzione di frequenza sono raccolti i dati di un campione di misure del peso in g di 120 sferette di acciaio (Tabella 5)

Tabella 5

classi	frequenza assoluta
90<x<=110	7
110<x<130	21
130<x<=150	38
150<x<=170	33
170<x<=190	13
190<x<=210	8

Verificare l'adattamento dei dati a una distribuzione normale avente valor medio 150 e scarto quadratico medio 25 (valori teorici assegnati)

Ipotesi nulla H0	I dati sono distribuiti in modo approx. normale
Ipotesi alternativa H1	I dati non sono distribuiti normalmente

Calcolo delle frequenze attese e del valore della statistica test chi quadro

valor medio teorico	150
scarto quadratico medio teorico	25

classi	estremi destri classi	Frequenze assolute osservate O	Probabilità (distribuzione normale)	Frequenze attese A	(O-A)^2/A
90<x<=110	110	7	0,0548	6,58	0,0273
110<x<130	130	21	0,1571	18,85	0,2460
130<x<=150	150	38	0,2881	34,58	0,3388
150<x<=170	170	33	0,2881	34,58	0,0720
170<x<=190	190	13	0,1571	18,85	1,8138
190<x<=210		8	0,0548	6,58	0,3084
		120	1,0000	120	**2,8063**

chi quadro

Attenzione: nel calcolo del grado di libertà ricordare che i due parametri della distribuzione sono assegnati, non vengono calcolati dai dati del campione (m=0)

Grado di libertà	5
Livello di significatività	0,05
Regione di rifiuto	11,070

Conclusione:
Il valore delle statistica test chi quadro non appartiene alla regione di rifiuto, perciò non si può rifiutare l'ipotesi nulla: i dati del campione permettono di concludere che la distribuzione è approssimativamente normale

Soluzione con la funzione TEST.CHI

p-value	**0,730**

Conclusione: Il valore del p-value è grande perciò non si sbaglia non rifiutando l'ipotesi nulla.

Esempio 80.4
Nella tabella di distribuzione di frequenza sono raccolti i dati riguardanti le stature di 60 studenti (Tabella 6)

Tabella 6

classi	freq. assoluta
162<x<=165	2
165<x<=168	13
168<x<=171	24
171<x<=174	15
174<x<=177	6

Verificare l'adattamento a una distribuzione normale con media 170 e scarto quadratico medio 3 (valori teorici assegnati)

Ipotesi nulla H0	I dati si adattano alla distribuzione normale
Ipotesi alternativa H1	I dati non sono distribuiti normalmente

Calcolo delle frequenze attese e della statistica test chi quadro

valor medio teorico	170
scarto quadratico medio teorico	3

classi	estremi destri classi	frequenze osservate O	Probabilità (distrib. normale)	frequenze attese A
x<=165	165	2	0,0478	2,87
165<x<=168	168	13	0,2047	12,28
168<x<=171	171	24	0,3781	22,68
171<x<=174	174	15	0,2782	16,69
x>174		6	0,0912	5,47
Totale		60	1,0000	60

Poiché c'è una classe (la prima) con frequenza attesa minore di 5, occorre procedere all'accorpamento di due classi adiacenti (le prime due)

classi	estremi delle classi	frequenze osservate O	frequenze attese A	(O-A)^2/A
x<=168	168	15	15,15	0,0015
168<x<=171	171	24	22,68	0,0764
171<x<=174	174	15	16,69	0,1719
x>174		6	5,47	0,0508
			60,00	**0,3005**

chi quadro

Attenzione: nel calcolo del grado di libertà ricordare che i due parametri della distribuzione sono assegnati, non vengono calcolati dai dati del campione (m=0)

Grado di libertà	3
Livello di significatività	0,05
Regione di rifiuto	7,815

Conclusione:
Il valore di chi quadro calcolato non appartiene alla regione di rifiuto, perciò non si può rifiutare l'ipotesi nulla. I dati del campione permettono di concludere che si ha un buon adattamento alla distribuzione normale

Soluzione con la funzione TEST.CHI

p-value	**0,960**

Conclusione: il valore del p-value è elevato perciò non si rifiuta l'ipotesi nulla.

Esercizio 80.5
Nella tabella di distribuzione di frequenza sono raccolti i dati di un campione di 150 misure di peso in g (Tabella 7)

Tabella 7

classi	freq. assoluta
80<x<=110	3
110<x<=140	10
140<x<=170	24
170<x<=200	29
200<x<=230	36
230<x<=260	22
260<x<=290	18
290<x<=320	8

Verificare l'adattamento a una distribuzione normale con media 210 e scarto quadratico medio 50 (valori teorici assegnati)

Ipotesi nulla H0	I dati si adattano alla distribuzione normale
Ipotesi alternativa H1	I dati non sono distribuiti normalmente

Calcolo delle frequenze attese e della statistica test chi quadro

valor medio teorico	210
scarto quadratico medio teorico	50

classi	estremi destri classi	frequenze osservate O	Probabilità (distrib. normale)	frequenze attese A
x<=110	110	3	0,0228	3,41
110<x<=140	140	10	0,0580	8,70
140<x<=170	170	24	0,1311	19,66
170<x<=200	200	29	0,2089	31,33
200<x<=230	230	36	0,2347	35,20
230<x<=260	260	22	0,1859	27,89
260<x<=290	290	18	0,1039	15,58
x>290		8	0,0548	8,22
Totale		150	1,0000	150

Poiché c'è una classe (la prima) con frequenza attesa minore di 5, occorre procedere all'accorpamento di due classi adiacenti (le prime due)

classi	estremi delle classi	frequenze osservate O	frequenze attese A	(O-A)^2/A
x<=140	140	13	12,11	0,0649
140<x<=170	170	24	19,66	0,9557
170<x<=200	200	29	31,33	0,1737
200<x<=230	230	36	35,20	0,0181
230<x<=260	260	22	27,89	1,2433
260<x<=290	290	18	15,58	0,3764
x>290		8	8,22	0,0059
		150	150	**2,8380**

chi quadro

Attenzione: nel calcolo del grado di libertà ricordare che i due parametri della distribuzione sono assegnati, non vengono calcolati dai dati del campione (m=0)

Grado di libertà	6
Livello di significatività	0,05
Regione di rifiuto	12,592

Conclusione:
Il valore di chi quadro calcolato non appartiene alla regione di rifiuto, perciò non si può rifiutare l'ipotesi nulla. I dati del campione permettono di concludere che si ha un buon adattamento alla distribuzione normale

Soluzione con la funzione TEST.CHI

p-value	**0,829**

Conclusione: il valore del p-value è elevato perciò non si rifiuta l'ipotesi nulla.

Torna su

Soluzione Esercizio 81
Test chi quadro di adattamento alla distribuzione normale

Indice

Ritorna Esercizio 81

Esercizio 81.1
Sono assegnati i dati della tabella 1
Costruire una distribuzione di frequenza assoluta, raggruppando i dati in **6 classi chiuse a destra di uguale ampiezza**.
Disegnare l'istogramma della distribuzione di frequenza assoluta
Trovare l'intervallo di confidenza per la media con grado di fiducia del 95%.
Verificare l'adattamento dei dati a una distribuzione normale; usare come parametri della distribuzione normale i valori calcolati dai dati
Effettuare il test con livello di significatività del 5%

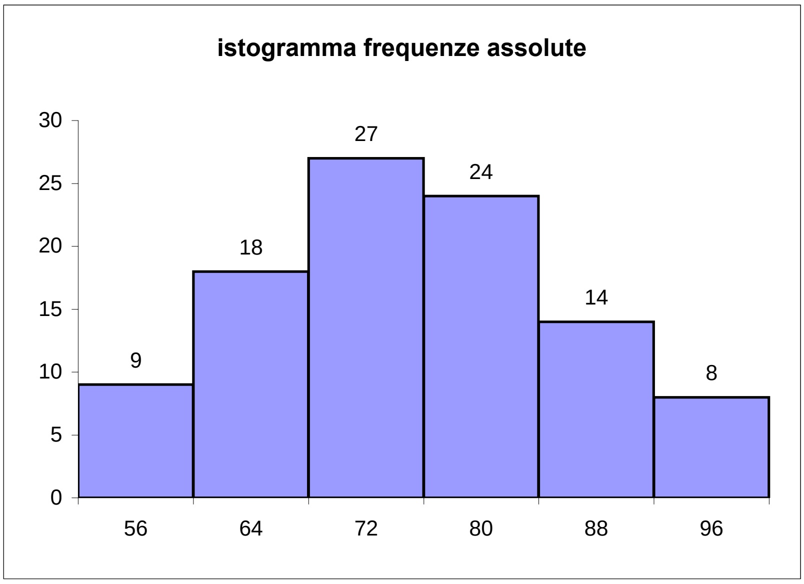
Tabella 1

68	73	61	66	80
84	79	65	78	78
75	88	75	82	89
82	73	87	75	61
86	60	94	94	75
90	93	62	77	95
59	71	95	69	60
88	59	78	74	79
76	85	63	68	83
81	75	78	60	71
79	87	86	61	66
62	80	67	65	78
59	80	73	75	82
97	57	81	87	75
78	88	72	74	82
85	78	63	62	77
76	62	76	95	69
87	76	75	78	74
71	53	85	63	68
75	74	96	72	60

numero dati	100
minimo	53
massimo	97
media	75,580
varianza	108,105
scarto quad. medio	10,397

range	44
num classi	6
ampiezza	8

classi	**estremo destro**	**frequenza assoluta**	**valori centrali**
52<x<=60	60	9	56
60<x<=68	68	18	64
68<x<=76	76	27	72
76<x<=84	84	24	80
84<x<=92	92	14	88
92<x<=100	100	8	96



Intervallo di confidenza per la media

grado di fiducia	0,950
metà intervallo	2,04
estremo sinistro	73,54
estremo destro	77,62

Test di adattamento alla normale

Ipotesi nulla H0	I dati si adattano alla distribuzione normale
Ipotesi alternativa H1	I dati non si adattano alla distribuzione normale

I parametri della distribuzione normale sono calcolati dai dati osservati (m=2)

valor medio	75,58
scarto quadratico medio	10,40

classi	estremo destro	frequenza assoluta	probabilità (normale)	frequenze attese	(O-A)^2/A
x<=60	60	9	0,067	6,701	0,789
60<x<=68	68	18	0,166	16,598	0,118
68<x<=76	76	27	0,283	28,312	0,061
76<x<=84	84	24	0,275	27,487	0,442
84<x<=92	92	14	0,152	15,188	0,093
x>92	92	8	0,057	5,714	0,915
		100	1,000	100,000	**2,418**

chi quadro

gradi libertà	3
livello significatività	0,05
regione rifiuto	7,815

Conclusione:
Il valore di chi quadro calcolato non appartiene alla regione di rifiuto, perciò non si può rifiutare l'ipotesi nulla. I dati del campione permettono di concludere che si ha un buon adattamento alla distribuzione normale

Soluzione con la funzione TEST.CHI

p-value	**0,7888**

Conclusione: Il valore del p-value è grande perciò non si rifiuta l'ipotesi nulla.

Esercizio 81.2

Sono assegnati i dati della tabella 2

Costruire una distribuzione di frequenza assoluta, raggruppando i dati in **6 classi chiuse a destra di uguale ampiezza**.

Disegnare l'istogramma della distribuzione di frequenza assoluta

Verificare l'adattamento dei dati a una distribuzione normale; usare come parametri della distribuzione normale i valori calcolati dai dati

Effettuare il test con livello di significatività del 5%

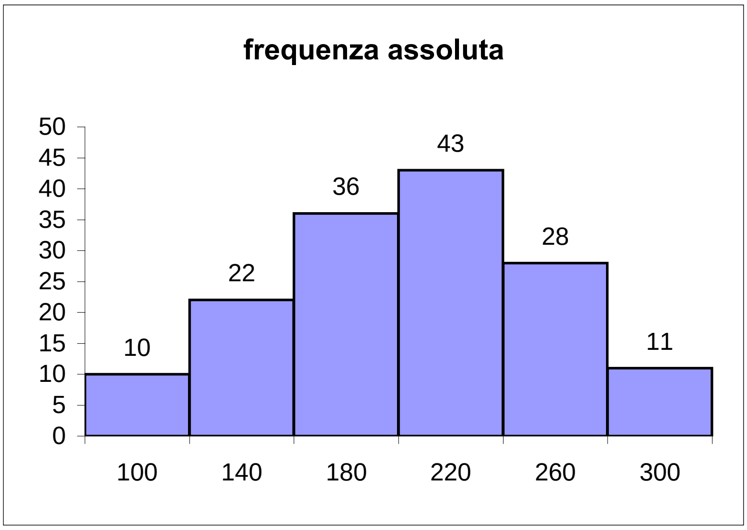
Tabella 2

148	192	214	206	274	212
149	272	179	209	317	139
220	217	272	166	209	300
250	173	164	142	106	177
208	204	203	183	215	250
188	194	159	280	221	269
281	277	190	299	194	103
165	238	259	126	209	270
244	220	207	221	98	265
306	166	199	198	117	163
221	153	148	206	245	299
152	236	118	244	187	126
233	280	219	144	110	221
153	155	226	144	182	198
215	129	217	215	156	206
245	294	96	215	237	198
267	185	237	186	197	178
196	155	198	187	217	169
235	131	268	249	227	251
102	268	206	164	143	127
203	284	209	190	205	194
158	297	166	259	193	201
107	254	142	216	267	166
260	204	183	285	303	104
280	217	280	197	217	266

numero dati	150
minimo	96
massimo	317
media	205,13
varianza	2720,38
scarto quadr. medio	52,16

numero classi	6
range	221
ampiezza classi	40

classi	**estremo destro**	**frequenza assoluta**	**valori centrali**
80<x<=120	120	10	100
120<x<=160	160	22	140
160<x<=200	200	36	180
200<x<=240	240	43	220
240<x<=280	280	28	260
280<x<=320	320	11	300
	totale	150	



Test di adattamento alla normale

Ipotesi nulla H0	I dati si adattano alla distribuzione normale

I parametri della distribuzione normale sono calcolati dai dati osservati (m=2)

media	205,13
scarto quadratico medio	52,16

classi	estremo destro	frequenza assoluta	probabilità (normale)	frequenze attese	(O-A)^2/A
x<=120	120	10	0,051	7,699	0,688
120<x<=160	160	22	0,142	21,320	0,022
160<x<=200	200	36	0,267	40,108	0,421
200<x<=240	240	43	0,287	43,092	0,000
240<x<=280	280	28	0,176	26,445	0,091
x>280	280	11	0,076	11,335	0,010
	totale	150	1,000	150	**1,232**

chi quadro

grado libertà	3
liv. significatività	0,05
regione rifiuto	7,815

Conclusione:
Il valore di chi quadro calcolato non appartiene alla regione di rifiuto, perciò non si può rifiutare l'ipotesi nulla. I dati del campione permettono di concludere che si ha un buon adattamento alla distribuzione normale

Soluzione con la funzione TEST.CHI

p-value	**0,9418**

Conclusione: Il valore del p-value è elevato perciò non si rifiuta l'ipotesi nulla.

Torna su

Soluzione Esercizio 82
Test chi quadro di indipendenza

Indice

Ritorna Esercizio 82

Con il **test chi quadro di indipendenza** si può stabilire la dipendenza o indipendenza fra due variabili qualitative misurate sullo stesso insieme di dati
Le ipotesi da verificare sono le seguenti

Ipotesi nulla H0	le due variabili sono indipendenti
Ipotesi alternativa H1	le due variabili sono dipendenti

I dati raccolti sulle due variabili qualitative sono riassunti in una tabella detta **tabella di contingenza**.
Il procedimento consiste nel confrontare la distribuzione delle frequenze osservate con la distribuzione delle frequenze attese
Il numero contenuto in ogni cella della tabella è la **frequenza osservata** associata a una delle catagorie rispetto alla prima variabile e a una delle categorie rispetto alla seconda variabile.
Le **frequenze attese**, ossia le frequenze che si avrebbero se l'ipotesi nulla fosse vera.
Per la correttezza del test le frequenze attese devono essere maggiori di 5; in caso contrario occorre procedere all'accorpamento di categorie, e ciò deve avvenire con criterio logico, sulla base del problema trattato
Per valutare l'adattamento delle frequenze osservate alle frequenze attese si utilizza la **statistica test chi quadro** calcolata con la formula

$$\boxed{\chi^2 = \sum_{i=1}^{r}\sum_{j=1}^{c}\frac{\left(O_{ij} - A_{ij}\right)^2}{A_{ij}}}$$

Oij	frequenze osservate
Aij	frequenze attese
r	numero righe della tabella di contengenza
c	numero colonne della tabella di contengenza

La statistica test ha approssimativamente la distribuzione chi quadro con grado di libertà

$$\boxed{\nu = (r-1)(c-1)}$$

La decisione è basata sul confronto fra la statistica test e il valore critico χ^2_α
dove alfa è il **livello di significatività**, di solitp uguale a 0,05 oppure 0,01
La **regione di rifiuto** è data dai valori tali che

$$\boxed{\chi^2 > \chi^2_\alpha}$$

Esempio 82.1
Per stabilire l'efficacia di un vaccino anti-influenzale è stata condotta una ricerca, somministrando il vaccino a 500 persone e controllando il loro stato di salute per sei mesi dopo la vaccinazione.
Lo stesso controllo è stato fatto su un altro gruppo di 500 persone non vaccinate.
I dati ottenuti sono riassunti nella seguente tabella 1

Tabella 1 - Frequenze osservate

	nessuna influenza	una influenza	più di una influenza	Totale
vaccinati	252	145	103	500
non vaccinati	224	136	140	500
Totale	476	281	243	1000

Calcolo delle frequenze attese

Oij frequenze osservate

Aij frequenze attese

Ogni frequenza attesa Aij si calcola moltiplicando il totale della riga i per il totale della colonna j e dividendo per il totale dei dati nella tabella delle frequenze osservate

Nota: nel calcolo delle frequenze attese fare attenzione all'**uso dei riferimenti misti** necessari per il corretto trascinamento delle formule

Frequenze attese

	nessuna influenza	una influenza	più di una influenza	Totale
vaccinati	238	140,5	121,5	500
non vaccinati	238	140,5	121,5	500
Totale	476	281	243	1000

Calcolo della statistica test chi quadro

Si applica la formula

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{\left(O_{ij} - A_{ij}\right)^2}{A_{ij}}$$

r numero righe

c numero colonne

Si possono disporre i calcoli nella seguente tabella, dove in ogni cella compare il corrispondente termine della sommatoria; nella cella corrispondente al valore di chi quadro si fa la somma di tutti gli elementi di questa tabella (utile solo per svolgere i calcoli)

	nessuna influenza	una influenza	più di una influenza
vaccinati	0,8235	0,1441	2,8169
non vaccinati	0,8235	0,1441	2,8169

chi quadro

7,569

Ipotesi nulla H0	variabili indipendenti (il vaccino non è efficace)
Ipotesi alternativa H1	variabili dipendenti (il vaccino è efficace)

Grado di libertà	2
Livello di significatività	0,05
Regione di rifiuto	5,991

Conclusioni: si rifiuta l'ipotesi nulla, il vaccino è efficace

Livello di significatività	0,01
Regione di rifiuto	9,210

Conclusioni: non si rifiuta l'ipotesi nulla, il vaccino non è efficace

Confrontando le conclusioni contrastanti ai due livelli di significatività, si deduce che si è in una situazione critica, per cui occorre un ulteriore studio del problema, con un campione più grande.

Soluzione con la funzione TEST.CHI
Con Excel si può fare il test di indipendenza con la funzione TEST.CHI che restituisce in uscita il p-value, ossia il più piccolo livello di significatività (calcolato a partire dalle tabelle di frequenze osservate e frequenze attese) per il quale si può rifiutare l'ipotesi nulla.
Un p-value molto prossimo a 0 indica che la probabilità di sbagliare rifiutando l'ipotesi nulla è molto vicina a 0 (ossia siamo praticamente certi di non sbagliare)
Un p-value vicino ai classici livelli di significatività indica che la decisione è critica.
Un p-value maggiore indica che non si può rifiutare l'ipotesi.

p-value	**0,0227**

Conclusione: il valore del p-value, vicino ai livelli di significatività comunemente usati, indica che ci si trova in una situazione critica.

Esercizio 82.2
C'è dipendenza tra tipo di lavoro e sesso del lavoratore?
Un campione di 250 lavoratori mostra i seguenti dati (**Tabella 1 - frequenze osservate**)

Tabella 1

	M	F	Totali
Manager Professionista	42	30	72
Tecnico Amministrativo	28	48	76
Servizi	12	18	30
Altro	60	12	72
Totali	142	108	250

Calcolo delle **frequenze attese (Tabella 2)**

Tabella 2

	M	F	Totali
Manager Professionista	40,90	31,10	72,00
Tecnico Amministrativo	43,17	32,83	76,00
Servizi	17,04	12,96	30,00
Altro	40,90	31,10	72,00
Totali	142,00	108,00	250,00

Calcolo del valore della statistica test chi quadro

	M	F
Manager Professionista	0,030	0,039
Tecnico Amministrativo	5,330	7,007
Servizi	1,491	1,960
Altro	8,924	11,734

chi quadro

36,51

Ipotesi nulla H0	non c'è dipendenza fra tipo di lavoro e sesso
Ipotesi alternativa H1	c'è dipendenza fra tipo di lavoro e sesso

Gradi di libertà	3
Livello di significatività	0,01
Regione di rifiuto	11,345

Conclusioni:
Il valore della statistica test chi quadro appartiene alla regione di rifiuto, perciò si rifiuta l'ipotesi nulla: c'è dipendenza fra tipo di lavoro e sesso del lavoratore

Soluzione con la funzione TEST.CHI

p-value	**5,829E-08**

Conclusione: Il valore del p-value è molto piccolo perciò si rifiuta l'ipotesi nulla.

Torna su

Soluzione Esercizio 83
Test chi quadro di indipendenza

Indice

Ritorna Esercizio 83

Esercizio 83.1

Dall'esame del colore dei capelli dei bambini di una certa regione si sono ricavati i seguenti dati (Tabella 1 - **frequenze osservate**); stabilire se c'è dipendenza fra colore dei capelli e sesso del bambino, confrontando i dule livelli di significatività

	Tabella 1 - Frequenze osservate					
	biondo	**rosso**	**castano**	**bruno**	**nero**	**Totali**
maschi	592	119	849	504	36	2100
femmine	544	97	677	451	14	1783
Totali	1136	216	1526	955	50	3883

Calcolo delle **frequenze attese (Tabella 2)**

	Tabella 2 - Frequenze attese					
	biondo	**rosso**	**castano**	**bruno**	**nero**	**Totali**
maschi	614,37	116,82	825,29	516,48	27,04	2100,00
femmine	521,63	99,18	700,71	438,52	22,96	1783,00
Totali	1136,00	216,00	1526,00	955,00	50,00	3883,00

Calcolo del valore della statistica test chi quadro

	biondo	**rosso**	**castano**	**bruno**	**nero**	
maschi	0,81	0,04	0,68	0,30	2,97	
femmine	0,96	0,05	0,80	0,36	3,50	
					chi-quadro	**10,47**

Ipotesi nulla H0	non c'è dipendenza fra colore dei capelli e sesso
Ipotesi alternativa H1	c'è dipendenza fra colore dei capelli e sesso

Grado di libertà	4
Livello di significatività	0,01
Regione di rifiuto	13,277

Conclusioni:
Il valore della statistica test chi quadro non appartiene alla regione di rifiuto, perciò non si rifiuta l'ipotesi nulla. C'è evidenza statistica di dipendenza fra colore dei capelli e sesso

Grado di libertà	4
Livello di significatività	0,05
Regione di rifiuto	9,488

Conclusioni:
Il valore della statistica test chi quadro appartiene alla regione di rifiuto, perciò si rifiuta l'ipotesi nulla
Non c'è evidenza statistica di dipendenza fra colore dei capelli e sesso

Confrontando le conclusioni ai due diversi livelli di significatività si evidenzia una situazione critica!

Soluzione con la funzione TEST.CHI

p-value	**0,03325**

Conclusione: il valore del p-value conferma la situazione critica

Esercizio 83.2

E' stata condotta un'indagine su un gruppo di 400 persone per accertare l'eventuale dipendenza fra il titolo di studio e il giudizio sulla linea politica di un importante quotidiano.
Il titolo di studio è stato distinto in 3 fasce: media inferiore, media superiore, laurea.
Sono previste due risposte al quesito: favorevole o contrario.
I risultati sono raccolti nella tabella 3 delle frequenze osservate
Stabilire con il test chi quadro se esiste o no dipendenza fra titolo di studio e giudizio

Tabella 3 - Frequenze osservate

	Favorevoli	**Contrari**	**Totale**
media inferiore	80	65	145
media superiore	58	32	90
laurea	110	55	165
Totale	248	152	400

Calcolo delle frequenze attese (Tabella 4)

Tabella 4 - Frequenze attese

	Favorevoli	**Contrari**	**Totale**
media inferiore	90	55	145
media superiore	56	34	90
laurea	102	63	165
Totale	248	152	400

Calcolo del valore della statistica test chi quadro

	Favorevoli	**Contrari**
media inferiore	1,0902	1,7788
media superiore	0,0867	0,1415
laurea	0,5796	0,9456

chi quadro

4,6224

Ipotesi nulla H0:	c'è indipendenza fra titolo di studio e giudizio
Ipotesi alternativa H1	c'è dipendenza fra titolo di studio e giudizio

livello significatività	0,05
grado libertà	2
chi quadro alfa	5,991

livello significatività	0,01
grado libertà	2
chi quadro alfa	9,210

Conclusione: non si rifiuta l'ipotesi nulla, c'è indipendenza fra titolo di studio e giudizio

Soluzione con la funzione TEST.CHI

p-value	0,099

Conclusione: il valore del p-value conferma che non si può rifiutare l'ipotesi nulla

Torna su

Soluzione Esercizio 84
Test chi quadro di indipendenza

Indice

Ritorna Esercizio 84

Esercizio 84.1

Una ditta le cui vendite dipendono esclusivamente dalla pubblicità vuole sapere se c'è dipendenza fra i diversi tipi di mezzi di comunicazione che usa per la pubblicità e l'età del cliente.
La ditta analizza gli ordini di 450 clienti, da cui risultano l'età del cliente e il mezzo di comunicazione attraverso il quale il cliente ha conosciuto l'azienda, e raccoglie le informazioni riportate nella tabella 1 (frequenze osservate)
Stabilire con il test chi quadro se c'è dipendenza fra età e mezzo di comunicazione.
Livello di significatività alfa = 0,05

	Tabella 1 - frequenze osservate				
	fascia di età				
	21—30	**31—40**	**41—50**	**>50**	**Totale**
internet	49	52	22	12	135
televisione	64	72	41	25	202
stampa	42	36	16	19	113
Totale	155	160	79	56	450

Calcolo delle **frequenze attese (Tabella 2)**

	Tabella 2 - frequenze attese				
	fascia di età				
	21—30	**31—40**	**41—50**	**>50**	**Totale**
internet	46,5	48	23,7	16,8	135
televisione	69,6	71,8	35,5	25,1	202
stampa	38,9	40,2	19,8	14,1	113
Totale	155	160	79	56	450

Calcolo del valore della statistica test chi quadro

	21—30	**31—40**	**41—50**	**>50**	
internet	0,13	0,33	0,12	1,37	
televisione	0,45	0,00	0,86	0,00	
stampa	0,24	0,43	0,74	1,73	
					6,43

chi quadro

Ipotesi nulla H0	non c'è dipendenza fra età e mezzo di comunicazione
Ipotesi alternativa H1	c'è dipendenza

Grado di libertà	6
Livello di significatività	0,05
Regione di rifiuto	12,592

Conclusioni:
Il valore della statistica test chi quadro non appartiene alla regione di rifiuto, perciò non si rifiuta l'ipotesi nulla. C'è evidenza statistica di indipendenza fra mezzo di comunicazione x rifiuta pubblicità e età del cliente.

Soluzione con la funzione TEST.CHI

p-value	**0,377**

Conclusione: il valore del p-value è grande e conferma l'indipendenza

Esercizio 84.2
Un'azienda è interessata a valutare se c'è una relazione tra il tempo impiegato dai propri impiegati per raggiungere il luogo di lavoro e il livello di stress sul lavoro.
Uno studio condotto su 116 impiegati con mansioni simili ha portato ai risultati della tabella 3 (frequenze osservate).
Stabilire con il test chi quadro se c'è dipendenza fra tempo impiegato per raggiungere il luogo di lavoro e livello di stress.
Confrontare i due livelli di significatività alfa = 5% e alfa =1% e trarre le conclusioni

	Tabella 3 - frequenze osservate			
	livello di stress			
tempo impiegato	**alto**	**medio**	**basso**	**Totale**
meno di 15 minuti	9	5	18	32
da 15 a 45 minuti	17	8	28	53
più di 45 minuti	18	6	7	31
Totale	44	19	53	116

Calcolo delle **frequenze attese (Tabella 4)**

	Tabella 4 - frequenze attese			
	livello di stress			
tempo impiegato	**alto**	**medio**	**basso**	**Totale**
meno di 15 minuti	12,14	5,24	14,62	32
da 15 a 45 minuti	20,10	8,68	24,22	53
più di 45 minuti	11,76	5,08	14,16	31
Totale	44	19	53	116

Calcolo del valore della statistica test chi quadro

	livello di stress			
tempo impiegato	**alto**	**medio**	**basso**	
meno di 15 minuti	0,81	0,01	0,78	
da 15 a 45 minuti	0,48	0,05	0,59	**chi quadro**
più di 45 minuti	3,31	0,17	3,62	
				9,83

Ipotesi nulla H0	non c'è dipendenza fra età e mezzo di comunicazione
Ipotesi alternativa H1	c'è dipendenza

Grado di libertà	4	
Livello di significatività	0,05	0,01
Regione di rifiuto	9,488	13,277

Conclusioni:
Per il livello di significatività del 5% la statistica test appartiene alla regione di rifiuto, mentre per il livello dell' 1% non appartiene alla regione di rifuto: questo fatto mostra che siamo in un caso critico

Soluzione con la funzione TEST.CHI

p-value	**0,043**

Conclusione: il valore del p-value è compreso fra i livelli classici di significatività e conferma che siamo in presenza di una decisione critica.

Torna su