Introduzione a KaleidaGraph™

II edizione

Andrea De Mauro
andrea.demauro@polito.it

12 marzo 2005

Indice

Sommario

Questa guida si propone d'introdurre il lettore all'utilizzo del software per l'analisi dei dati scientifici KaleidaGraph™ (Synergy Software, <www.synergy.com>). Per venire incontro a chi ha la necessità di utilizzare da subito il programma, i paragrafi che trattano gli argomenti indispensabili sono evidenziati da una nota marginale come quella riportata qui sulla destra. La guida è corredata di brevi e superficiali richiami teorici di statistica e si chiude con un esempio che ricapitola le più importanti tecniche descritte, l'analisi quantitativa del moto di un pendolo semplice[1].

Questa è una nota marginale

[1]La presente guida nasce come supporto didattico per gli allievi ingegneri che seguono il corso di "Laboratorio di fisica generale" presso il Politecnico di Torino e che vogliono utilizzare KaleidaGraph™ come supporto per la compilazione delle relazioni sull'esperienze svolte. L'esempio riportato a pag. 21 si basa, infatti, su un'esperienza normalmente affrontata durante il predetto corso.

1 Introduzione

KaleidaGraph™ è un software che permette l'elaborazione e l'analisi grafica di qualsiasi tipo di dato numerico. Essendo stato sviluppato appositamente per applicazioni scientifiche, esso è sicuramente più adatto dei normali fogli elettronici ad accogliere dati sperimentali con un alto numero di cifre significative. Inoltre, i suoi strumenti di elaborazione soddisfano, in larga parte, le esigenze di fisici sperimentali e ingegneri. Per questi motivi è assolutamente consigliabile utilizzare KaleidaGraph™ e non altri software (come, ad esempio, Microsoft® Excel, più adatto per calcoli finanziari) per applicazioni di tipo scientifico.

Nelle pagine che seguono verranno elencate, in maniera piuttosto veloce, le più importanti funzioni di KaleidaGraph™ cercando di evidenziare il significato pratico di alcuni risultati matematici. Il lettore interessato potrà approfondire gli argomenti servendosi della guida, sicuramente più completa, fornita con l'applicazione. È possibile effettuare un download della versione shareware di questo software dal sito <www.kaleidagraph.com>.

2 L'introduzione dei dati

La schermata iniziale di KaleidaGraph™ prevede la visualizzazione di due finestre: la prima, avente il generico nome di `Data 1`, costituisce il supporto per l'introduzione dei dati; la seconda, chiamata `Formula Entry`, permette l'inserimento delle formule matematiche che agiscono sui dati inseriti.

La prima fase di qualsiasi tipo di lavoro sarà, evidentemente, l'introduzione dei dati sperimentali. Occorre specificare che, nel caso essi siano stati precedentemente immagazzinati in file di tipo diverso da quello predefinito da KaleidaGraph™ (`.qda`), questi possono essere ugualmente "importati" tramite la funzione `Import` del menù `File`. Grazie a questa funzione possiamo importare i dati memorizzati antecedentemente tramite l'utilizzo dell'applicazione Microsoft® Excel (`.xls`) o, semplicemente, su un file di testo (`.txt`).

Numerazione relativa e assoluta

La tabella `Data 1` appare suddivisa in righe[2], numerate a partire da 0, e in colonne[3], identificate da una o più lettere ordinate, all'inizio, alfabeticamente. Cliccando sull'intestazione di una colonna essa viene selezionata: per motivi che si chiariranno di seguito, è importante considerare che KaleidaGraph™ permette di identificare le colonne facendo riferimento alla

[2] Per inserire una riga si prema `Ctrl + J`.

[3] Per inserire una colonna si prema `Ctrl + H`.

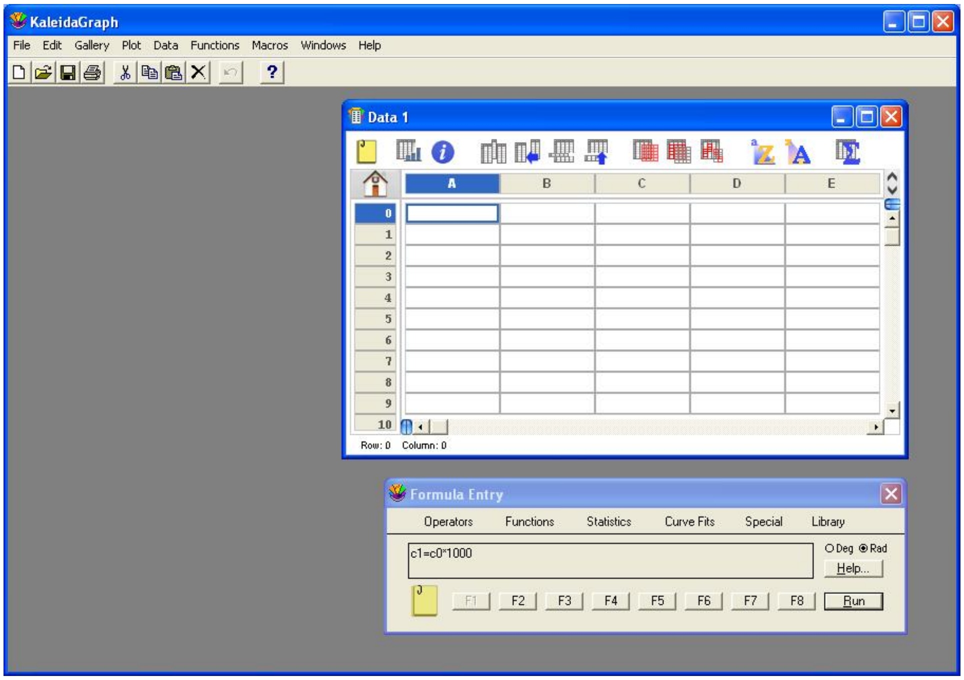

Figura 1: La schermata iniziale di KaleidaGraph™ .

loro posizione relativa rispetto a quella selezionata. In pratica, la colonna selezionata è identificata come `c0`, quella subito alla sua sinistra come `c1`, e così via. Questo sistema di numerazione verrà indicato come *numerazione relativa*. Per motivi analoghi, la denominazione alfabetica sarà di tipo *assoluto*[4]. L'utilizzo della numerazione relativa è indispensabile quando si vorranno applicare delle formule matematiche sui dati inseriti, come vedremo a pag. 4.

Inserire i dati

L'inserimento dei dati è facile e veloce: basta, infatti, selezionare la casella iniziale, digitare il numero (magari utilizzando il più comodo tastierino numerico) e scorrere a destra (con il tasto `TAB` o con la freccia destra) o in basso (con il tasto `Invio` o con l'apposita freccia).

Quando la mole di dati da inserire è molto grossa (ad esempio nel caso

[4]Per cambiare l'intestazione alfabetica delle colonne basta farci doppio clic sopra (nel caso di misure sperimentali è opportuno, dopo il nome, specificare l'unità di misura tra parentesi). Facendo clic col tasto destro del mouse e selezionando la voce `Column Formatting...` si potrà modificare il carattere dell'etichetta della colonna, il tipo di dati utilizzato (virgola mobile, intero, data, intervallo di tempo, ecc.), il formato e il numero di cifre decimali.

in cui si vogliano analizzare un elevato numero di misure sperimentali) si consiglia di ricorrere ad alcuni semplici elementi i quali, adottati per un numero abbastanza alto di volte, fanno risparmiare del tempo prezioso. Ad esempio:

- per introdurre il numero 0.552 si digiti solo .552;
- per introdurre dati con "molti 0" si può utilizzare la notazione esponenziale: per introdurre 0.00032 basterà digitare 32e-5 in quanto $0.00032 = 3.2\ 10^{-5}$; allo stesso modo, per introdurre 38000 si digiterà solo 38e3;
- nel caso, molto frequente, in cui la maggior parte dei dati da introdurre sia decimale e abbia uno stesso numero di 0 dopo la virgola o quando i numeri sono tutti molto grandi e aventi uno stesso numero di 0, si consiglia di digitare i numeri nella maniera più comoda (senza molti 0) e, solo in un secondo momento, riadattare i dati moltiplicandoli o dividendoli tutti per una potenza di 10.

Per effettuare delle operazioni, come quest'ultima, sui dati di una o più colonne bisognerà utilizzare la finestra `Formula Entry` già precedentemente introdotta, nel modo indicato di seguito.

3 L'utilizzo delle formule

Tramite l'utilizzo della finestra `Formula Entry` si possono elaborare in maniera rapida i dati precedentemente inseriti. KaleidaGraph™ permette di memorizzare otto funzioni ad accesso rapido e di salvare ed eseguire file contenti i listati per l'esecuzione di funzioni complesse. Le otto funzioni ad accesso rapido sono selezionabili cliccando sui tasti `F1`, `F2`, ..., `F8`. Per la memorizzazione di ogni funzione e, nel caso, per la loro sovrascrittura, basterà digitare la formula nella casella di testo centrale. Per l'esecuzione della formula, dopo averla selezionata, si dovrà cliccare sul tasto `Run` posizionato in basso a destra.

Scrivere semplici formule

Le formule editabili in KaleidaGraph™ sono, sostanzialmente, di due tipi:

- formule *ad assegnazione*;
- formule *a restituzione*.

Le prime hanno una struttura del tipo $y = f(x)$: il valore restituito da $f(x)$ viene assegnato a y dove x e y possono essere una colonna intera o il valore di una singola cella. Per indicare il singolo valore si usa la scrittura `cell(`x`,`y`)` dove x è il numero della riga mentre y è quello della colonna nella numerazione relativa. Appare evidente, quindi, come la scrittura `cell(0,0)` non si riferisca necessariamente ad un unico valore: essa fa riferimento, infatti, alla prima cella della colonna che, al momento della pressione del tasto `Run`, risulta essere selezionata. Un esempio di formula di questo tipo potrebbe essere:

```
cell(3,2)=2*log(cell(2,1))+2
```

Alcune funzioni possono operare anche su intere colonne: può, ad esempio, essere utile, per certe applicazioni, sommare tutte le celle di una colonna con quelle che si trovano subito sulla loro destra, cioé quelle della colonna successiva, memorizzando il risultato sulla colonna successiva ancora. Per indicare una colonna intera si utilizza la scrittura `c`x dove x è il numero relativo della colonna da considerare. Per effettuare l'operazione introdotta di sopra basta, allora, digitare `c2=c0+c1`, avendo l'accortezza di selezionare col mouse la prima colonna da sommare per riadattare la numerazione relativa.

Con questo metodo riusciamo, finalmente, a risolvere il problema che ci si è posti a pag. 4. Se i dati da inserire fossero stati:

58000
60000
56000
62000
...

e noi, per velocizzare il processo d'introduzione dei dati, avessimo inserito:

58
60
56
62
...

occorrerà moltiplicare tutti i valori della colonna inseriti per 1000, tramite la formula `c0=c0*1000`: cliccando sul tasto `Run` tutte tutti i valori delle celle appartenenti alla colonna selezionata verranno automaticamente "corredati" di tre zeri sulla loro destra. In altri casi, evidentemente, può essere opportuno dividere per potenze di 10 in modo tale da ottenere cifre decimali.

Non è detto che le formule debbano contenere necessariamente valori che si riferiscano o tutti a singole celle o tutti ad intere colonne. Se noi volessimo, ad esempio, memorizzare nella cella (0,1) il valore della media dei dati inseriti nella colonna selezionata, basterà scrivere `cell(0,1)=mean(c0)` dove la funzione `mean()` è, evidentemente, quella che restituisce la media di un set di valori[5]. Le funzioni inseribili nel `Formula Entry` sono selezionabili cliccando sui nomi che compaiono in alto nella finestra, ovvero `Operators, Functions, ..., Library` (nel caso di utilizzo di funzioni trigonometriche, occorre specificare se gli angoli sui quali agiscono sono espressi in gradi (`Deg`) o in radianti (`Rad`) tramite il bottone situato sulla destra della finestra, sopra il tasto di aiuto). Le funzioni matematiche più comuni utilizzabili in KaleidaGraph™ sono riportate in tab. 1.

Le formule *a restituzione* hanno una struttura del tipo $f(x)$: il valore di $f(x)$ non viene assegnato ad alcuna cella né colonna ma viene visualizzato in una finestra di dialogo denominata `Macro Results`. Cliccando sul tasto `To Clipboard`, il risultato viene memorizzato nel buffer del calcolatore dimodoché possa, successivamente, essere "incollato" in una qualsiasi cella[8].

4 Le funzioni statistiche

Ottenere le statistiche sui dati

Una volta inseriti i dati e dopo averli elaborati opportunamente, occorrerà sicuramente calcolarne le statistiche come media, deviazione standard e via dicendo. Per fare ciò basta selezionare una o più colonne di dati e cliccare sulla voce `Statistics...` dal menù `Functions`. La finestra visualizzata (fig. 2) è costituita da due colonne, nelle quali si possono alternare le colonne selezionate precedentemente agendo sulle barre di scorrimento orizzontali in basso, e diverse righe riportanti sulla sinistra i nomi delle funzioni

[5] Un set di valori non è necessariamente costituito dagli elementi di una colonna. Per riferirsi a valori disposti diversamente si può utilizzare l'operatore matriciale `[ ]`. La sintassi di questo operatore è:

`[` *riga iniziale* : *riga finale* , *colonna iniziale* : *colonna finale*]

Se i valori delle righe o quelli delle colonne sono omessi, la scrittura si riferirà a tutte le righe o, rispettivamente, a tutte le colonne presenti nella tabella. Come esempio riportiamo:

`[,]`	Opera sull'intera tabella dei dati
`[,2:90]`	Opera su tutte le righe, dalla colonna 2 alla colonna 90
`[0:30,]`	Opera su tutte le colonne, dalla riga 0 alla riga 30
`[0:19,0:19]`	Opera sui dati contenuti nelle prime 20 righe e nelle prime 20 colonne

[6] È l'integrale della distribuzione gaussiana (vedi la (5) a pag. 10) calcolata tra 0 e x.

[7] Per ottenere un numero casuale intero compreso tra 0 ed $n-1$ basta scrivere `int(ran()*`n`)`.

[8] Selezionando l'opzione `Data only` viene copiato solo il valore numerico e non la sua etichetta.

Tabella 1: Funzioni matematiche

Funzione	Descrizione
abs(x)	Valore assoluto
erf(x)	Funzione degli errori[6]
erfc(x)	Complementare della funzione degli errori
exp(x)	Esponenziale (e^x)
log(x)	Logaritmo base 10
ln(x)	Logaritmo naturale
sqrt(x)	Radice quadrata
fract(x)	Mantissa (parte decimale di x)
int(x)	Parte intera di x
pi	Restituisce la costante π
sin(x)	Seno
cos(x)	Coseno
tan(x)	Tangente
invsin(x)	Arcoseno
invcos(x)	Arcocoseno
invtan(x)	Arcotangente
ran()	Restituisce un numero casuale compreso tra zero e uno[7]

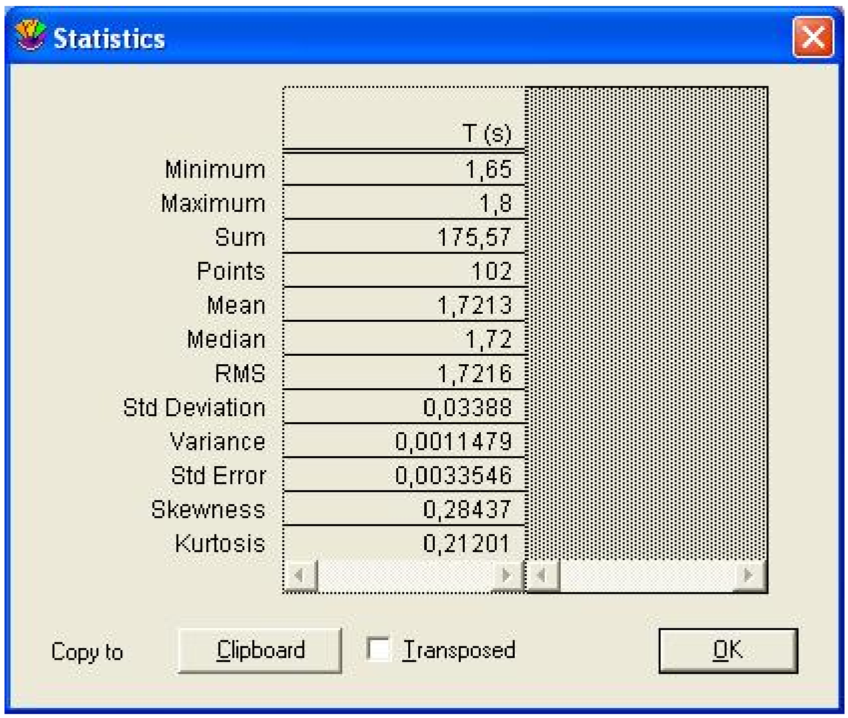

Figura 2: La finestra `Statistics`

statistiche visualizzate. Prima di entrare nel merito di quest'ultime occorre specificare che, cliccando sul tasto `Clipboard`, KaleidaGraph™ copia le statistiche[9] nel buffer dando la possibilità all'utente di "incollarle" nella relazione o nel rapporto che sta compilando. Ecco di seguito elencate le funzioni statistiche visualizzate:

`Minimum` e `Maximum` costituiscono, rispettivamente, il valore minimo e quello massimo del set di dati.

`Sum` è la somma algebrica di tutti i valori selezionati.

`Points` è il numero dei valori selezionati (senza contare le caselle vuote).

`Mean` è la media aritmetica dei valori selezionati calcolata secondo la nota formula:

$$\bar{x} = \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{1}$$

dove x_i è il valore i-esimo del set di dati ed n è il numero dei punti (`Points`).

`Median` costituisce la mediana[10] del set di dati.

`RMS` (root mean square) è la media quadratica semplice, ovvero:

$$Q = \frac{\sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}}{n} = \frac{1}{n}\sqrt{\sum_{i=1}^{n} x_i^2} \tag{2}$$

`Std Deviation` è la deviazione standard σ ovvero la radice quadrata positiva della varianza, cioè:

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2} = \ldots^{11} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i^2) - n\,\bar{x}^2} \tag{3}$$

[9] Selezionando l'opzione `Transposed` vengono copiate righe e colonne in ordine inverso (quello di una matrice trasposta).

[10] Ordinando i valori in senso crescente, si dice *mediana* il valore che bipartisce la successione ossia quello non inferiore alla metà dei valori e non superiore all'altra metà.

[11] In quanto:

$$\begin{aligned}\sum_{i=1}^{n}(x_i - \bar{x})^2 &= (x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2 = \\ &= x_1^2 + x_2^2 + \ldots + x_n^2 + n\,\bar{x}^2 - 2\,x_1\,\bar{x} - 2\,x_2\,\bar{x} - \ldots - 2\,x_n\,\bar{x} =\end{aligned}$$

`Variance` è la varianza σ^2, cioè la (3) senza il simbolo della radice.

`Std Error` (standard error) è la deviazione standard dalla media, pari a:

$$\begin{aligned}\sigma_{\bar{x}} &= \frac{\sigma}{\sqrt{n}} = \sqrt{\frac{1}{n(n-1)}\sum_{i=1}^{n}(x_i - \bar{x})^2} = \\ &= \sqrt{\frac{1}{n(n-1)}\sum_{i=1}^{n}(x_i^2) - n\,\bar{x}^2} \qquad (4)\end{aligned}$$

Prima di andare oltre occorre specificare a cosa servono questi ultimi tre dati nell'analisi di un set di misure sperimentali. Senza entrare troppo nel dettaglio[12], si ricorda che, nel caso in cui i dati sperimentali siano le misure di una data grandezza fisica:

- se le misure sono soggette a molte piccole sorgenti di errori casuali e a trascurabili errori sistematici, allora i valori misurati saranno distribuiti secondo una *distribuzione normale* (o *di Gauss*) centrata sul "valor vero" X e avente "larghezza" σ, dove la funzione di densità è:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp \frac{-(x-\bar{x})^2}{2\sigma^2} \qquad (5)$$

- La larghezza σ è il limite di confidenza del 68 % (ovvero, in parole povere, lo scarto tra il valore misurato x e quello vero X è più piccolo di σ con probabilità del 68 % circa);

- Si dimostra che la miglior stima del "valor vero" X è la media aritmetica $\bar{x}$ (`mean`) e la miglior stima di σ è la deviazione standard (`Std Error`);

- L'incertezza di $\bar{x}$ (ovvero della miglior stima del valor vero) è pari alla deviazione standard dalla media $\sigma_{\bar{x}}$ (`Std Error`).

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

[12] Per approfondire l'argomento si consiglia il testo: Vicario, Levi, *Calcolo delle probabilità e statistica per ingegneri*, Bologna, Progetto Leonardo, 1997.

In conseguenza di ciò, possiamo concludere che la misura può essere rappresentata con la scrittura:

$$X = \bar{x} \pm \sigma_{\bar{x}}$$

cioè, guardando la finestra `Statistics`, possiamo scrivere che la misurazione effettuata ha avuto come "risultato" `Mean` $\pm$ `StdError` (senza dimenticare mai di posticipare l'unità di misura)[13]. Quindi, grazie alla funzione `Statistics`, dopo aver inserito un set di misure sperimentali, KaleidaGraph™ ci permette di ottenere direttamente la migliore stima del valor vero e la sua attendibilità.

Il significato delle statistiche

Ogni grandezza contenuta nella (5) ha un suo preciso significato grafico quindi le varie voci della finestra `Statistics` ci informano sull'aspetto che avrà il grafico della distribuzione di frequenza (di cui si parlerà avanti). Per esempio, $\bar{x}$ (`Mean`) ci informa del valore su cui è centrata la distribuzione ($x = \bar{x}$ è la retta che meglio approssima un ipotetico asse di simmetria); σ (`Std Deviation`) ci dà un idea di quanto è "stretta" la distribuzione (più grande è σ, più la "campana" è allargata).

Le ultime due funzioni statistiche ci indicano come e di quanto la distribuzione delle frequenze risulta essere "sproporzionata" rispetto alla curva di Gauss.

`Skewness` è il *coefficiente di asimmetria* S ovvero il rapporto tra il momento terzo centrato sul valor medio e il cubo della deviazione standard:

$$S = \frac{\mu_3}{\sigma^3} = \frac{\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(x_i - \bar{x})^3}{\left(\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^{3/2}} \tag{6}$$

Questo parametro ci indica come e di quanto è "squilibrata" orizzontalmente la distribuzione del set di dati, precisamente:

- `Skewness` > 0: distribuzione deviata a sinistra;
- `Skewness` $= 0$: distribuzione simmetrica;
- `Skewness` < 0: distribuzione deviata a destra.

[13] La misura e l'incertezza si scrivono fra parentesi, fuori dalle quali si specifica l'unità di misura. Un esempio di scrittura corretta è

$$m = (\texttt{59,3} \pm \texttt{0,1})\,\texttt{kg}$$

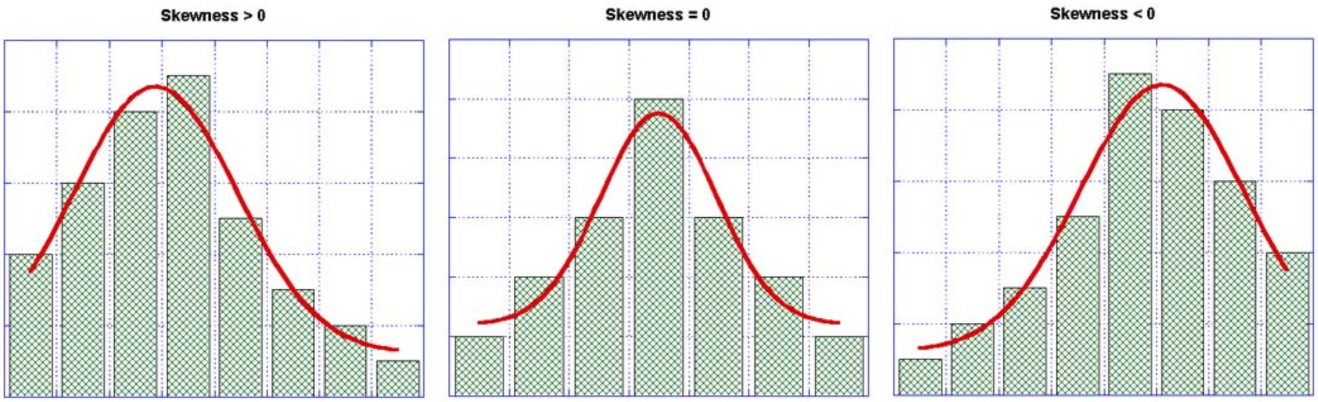

Figura 3: Significato grafico del coefficiente di asimmetria

`Kurtosis.` Il momento del quarto ordine centrato nella media fornisce l'*eccesso* o *curtosi* e misura, in pratica, l'appiattimento di una funzione di densità (nel nostro caso, quella di Gauss) rispetto al suo centro. In verità, per dare un significato grafico, si utilizza il *coefficiente di curtosi* E o *di appiattimento*, il nostro `Kurtosis`, pari al rapporto tra il momento quarto centrato nella media e la deviazione standard alla quarta, diminuito di 3:

$$E = \frac{\mu_4}{\sigma^4} - 3 = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^4}{\left(\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2\right)^2} - 3 \tag{7}$$

Il coefficiente di curtosi da' un'indicazione sulla posizione relativa del picco centrale della distribuzione delle frequenze rispetto a quello della funzione di densità normale, precisamente:

- `Kurtosis` > 0: distribuzione più allungata al centro di quanto non lo sia la campana di Gauss;
- `Kurtosis` $= 0$: distribuzione proporzionata, in altezza, in maniera simile alla campana di Gauss;
- `Kurtosis` < 0: distribuzione più schiacciata al centro di quanto non lo sia la campana di Gauss.

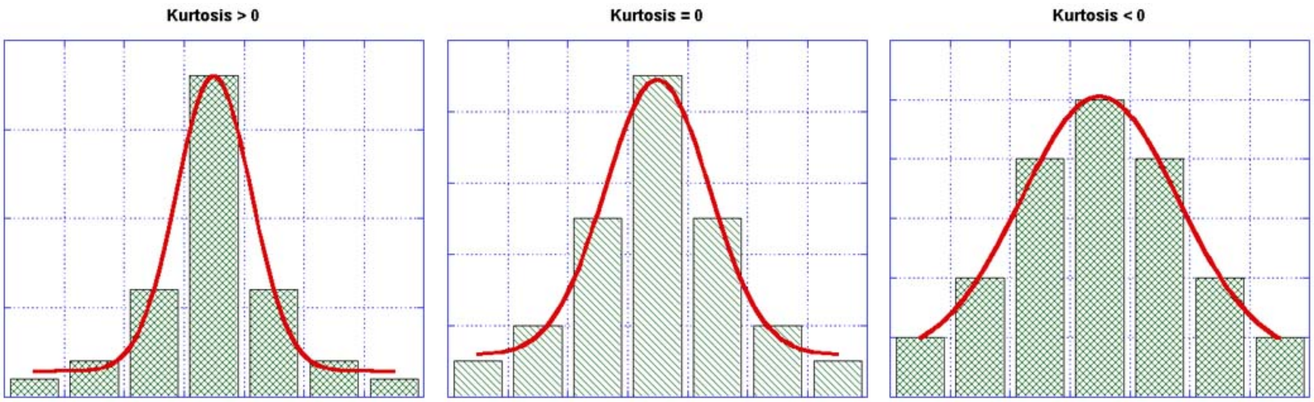

Figura 4: Significato grafico del coefficiente di curtosi

5 I grafici

KaleidaGraph™ permette di creare grafici a partire dai dati sperimentali inseriti o di visualizzare quelli di funzioni matematiche. Quest'ultima operazione è effettuabile tramite l'utilizzo della finestra `Function Plot` richiamabile cliccando sulla voce `Function` dal menù `Gallery`. Nella casella di testo superiore (quella dopo '`y =`') va inserita la funzione da graficare utilizzando x come variabile indipendente[14] cioè, ad esempio, scrivendo:

```
5*x^2+cos(x)
```

Le funzioni matematiche utilizzabili in questo contesto sono le stesse elencate nella tab. 1 di pag. 7. Nelle caselle `X min` e `X max` si devono specificare i punti estremi, rispettivamente quello sinistro e quello destro, dell'intervallo sul quale disegnare la funzione. Visto che KaleidaGraph™ grafica le funzioni matematiche "per punti", occorre specificargli su quanti punti calcolarne il valore tramite la casella `Points`: inserendo un valore piccolo si guadagnerà in velocità di calcolo ma si perderà in risoluzione, e viceversa. Premendo il tasto `New Plot` (o `Replot`, in fase di aggiornamento) si otterrà un grafico simile a quello riportato in fig. 5.

Personalizzare un grafico

I vari elementi di un grafico possono essere eliminati o personalizzati opportunamente tramite la casella degli strumenti apposita (fig. 5, in alto a sulla sinistra) che appare automaticamente insieme al grafico. Selezionando

[14] In verità si possono graficare funzioni anche utilizzando come dominio l'insieme dei valori di una colonna. Per far ciò basta utilizzare come variabile indipendente la colonna considerata, specificandola nella solita notazione relativa (es: `c2`).

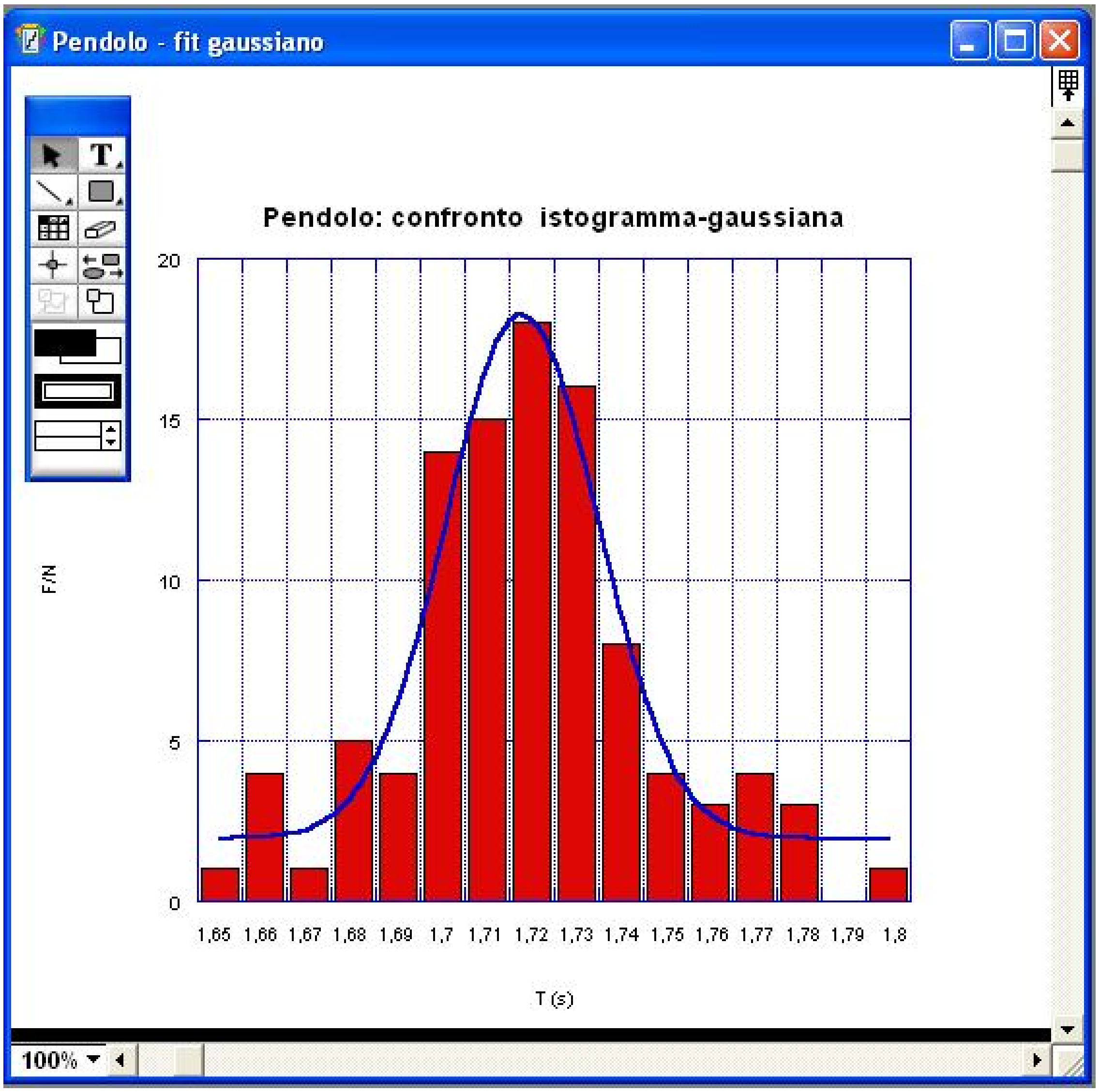

Figura 5: Un esempio di grafico

l'icona rappresentante la freccia è possibile evidenziare i vari elementi del grafico (etichette, legende, ecc.) per spostarli, ridimensionarli o eliminarli (con la pressione del tasto `Canc \ Del`). Selezionando la `T` si possono inserire le caselle di testo: tramite la finestra `Edit String` si possono modificare formato, tipo di carattere, dimensione, stile e colore. Gli altri strumenti servono per inserire segmenti, rettangoli, tabelle a doppia entrata oppure per zoomare una zona interessante del grafico e così via. Cliccando col tasto destro del mouse in un punto qualsiasi della finestra contenente il grafico, appare un menù pop-up dal quale sono richiamabili alcune finestre importanti per la personalizzazione del grafico: tramite la finestra `Axis Options` si può selezionare, ad esempio, la scala degli assi orientati (se lineare, `Linear`, o logaritmica, `Log`); con la finestra `Variable Settings` si possono modificare: la forma dei punti dei singoli punti del grafico, il riempimento delle eventuali superfici, lo stile e la dimensione del tratto, e il colore della linea. Utilizzando la lista visualizzata a sinistra (`Plot Variable`) si seleziona la linea alla quale si fà riferimento (ovvero quella da modificare).

Graficare dei dati sperimentali

Tutte queste possibilità di personalizzazione dei grafici sono del tutto valide anche quando si vogliono graficare dei dati sperimentali. Per fare ciò occorre, semplicemente, selezionare il tipo di grafico dal menù `Gallery` e specificare le colonne della tabella dei dati da utilizzare. Dopo avere selezionato il tipo di grafico, infatti, apparirà una finestra, denominata `Plot`, tramite la quale, grazie all'uso di alcuni bottoni radio, si dovranno specificare le colonne da utilizzare come coordinate X e quelle da utilizzare come coordinate Y dei vari punti del grafico. Evidentemente, per un grafico a coordinare polari si dovranno specificare quali colonne utilizzare per i valori di Θ (angolo) e di R (raggio).

Aggiungere le barre d'incertezza

Una volta ottenuto il grafico dei dati sperimentali può essere opportuno aggiungere le *barre d'incertezza*: esse servono a dare un'idea, a chi esamina il grafico, della distribuzione delle incertezze di cui sono affette le varie misure sperimentali. Per aggiungere le barre degli errori si deve far apparire il solito menù pop-up (clic col tasto destro del mouse sul grafico) e cliccare sulla voce `Error Bars...`: apparirà una finestra denominata `Error Bar Variables` grazie alla quale selezionare i set di dati ai quali aggiungere le barre d'incertezza. Cliccando su una casella qualsiasi si aprirà una nuova finestra, `Error Bar Settings`, grazie alla quale si potranno modificare le impostazioni delle barre: i due menù a tendina centrali servono per informare KaleidaGraph™ sulle modalità di assegnazione dell'errore in eccesso (in alto) e in difetto (in basso). Le modalità disponibili sono:

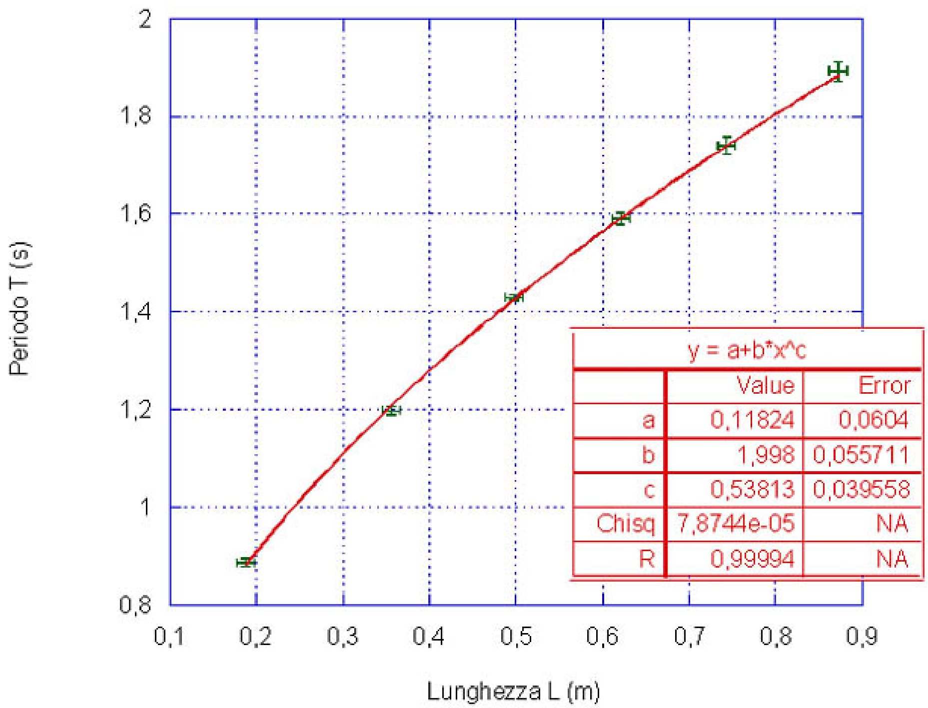

Figura 6: Un grafico (fit) con le barre d'incertezza

- `% of Value`. La lunghezza delle barre sono proporzionali ai valori ai quali fanno riferimento: la costante di proporzionalità (*errore relativo percentuale*) viene inserita nell'apposita casella di testo sulla destra.

- `Fixed Value`. La lunghezza è fissa per tutti i valori e specificata nella casella `Fixed Error`.

- `Standard Dev.` La lunghezza è pari ad un numero n preciso di deviazioni standard ($n\cdot$ `Std Deviation`) specificato nella casella `# Standard Dev`.

- `Standard Error`. La lunghezza delle barre è pari alla deviazione standard dalla media della colonna considerata.

- `Data Column`. Viene data la possibilità di specificare i singoli valori delle lunghezze delle barre degli errori: basta selezionare la colonna della tabella dei dati che contiene questi valori.

Graficare la distribuzione delle frequenze

Nel caso in cui si voglia graficare la distribuzione delle frequenze di alcuni dati sperimentali (che, nelle ipotesi di pag. 10, dovrebbe "assomigliare" alla campana di Gauss), occorre agire ancora sulla tabella dei dati arricchendola di due importanti colonne: una riportante i valori iniziali delle varie classi di dati (`Bins`), l'altra le rispettive frequenze (ovvero il numero dei valori che cadono nello specifico intervallo). Queste due colonne vengono calcolate automaticamente dal programma, tramite la funzione `Bin Data...` richiamabile dal menù `Functions` (dopo aver selezionato una o più colonne di dati sperimentali). Nella finestra `Bin Data` sono visualizzate: un'anteprima delle due colonne di cui sopra (una coppia per ogni colonna di dati sperimentali selezionata all'inizio); un'anteprima dell'istogramma delle frequenze; una serie di caselle di testo e bottoni radio grazia ai quali impostare i calcoli. Nelle caselle `Min` e `Max` si settano, rispettivamente, il valore minimo e quello massimo dell'intervallo da suddividere in classi. Nella casella `# of Bins` si possono dichiarare il numero di classi nelle quali suddividere l'intervallo considerato. I bottoni radio che seguono servono per predisporre le modalità di output del risultato. Cliccando su `Recalculate` vengono aggiornate le anteprime delle colonne e dell'istogramma. All'apertura della finestra, le condizioni iniziali di `Min`, `Max` e `# of Bins` sono ottimali per una corretta suddivisione in classi: si consiglia, pertanto, di non modificare tali valori. Cliccando sul tasto `Clipboard` si copia il contenuto delle colonne dell'anteprima nel buffer: una volta chiusa la finestra, selezionando due colonne vuote e premendo `Ctrl + V` si incolla il contenuto del buffer nella tabella dei dati.

A questo punto, per ottenere il grafico della distribuzione dei dati sperimentali, basta scegliere il tipo di grafico, ad esempio cliccando su `Gallery >> Bar >> Column`, e selezionare la colonna riportante i minimi delle classi (di solito, denominata 'Histogram X') come X e l'altra, o le altre, come Y. Per completezza, occorre affermare che esiste un altro metodo, molto più veloce di questo, per ottenere il grafico della distribuzione dei dati: basta cliccare su `Gallery >> Stat >> Histogram`, selezionare la colonna o le colonne dei dati sperimentali (come Y), cliccare sul tasto `New Plot` e poi su `Ok` nella finestra che segue. Questo metodo, però, è meno generale di quello esposto precedentemente e non permette la visualizzazione delle colonne riportanti la suddivisione in classi di frequenza né l'aggiunta di fit.

Copiare un grafico nella relazione

Per concludere con l'argomento "grafici", s'illustra il metodo per copiare un qualsiasi grafico nel documento che si sta compilando: basta selezionare la finestra del grafico (cliccandoci sopra in un punto qualsiasi col tasto sinistro) e premere `Ctrl + C` o cliccare su `Edit >> Copy Graph...`. Una volta copiato il grafico nel buffer lo si può "incollare" nel documento che si sta preparando, tramite le apposite funzioni (per la maggior parte delle applicazioni, basta selezionare il punto opportuno del documento nel quale va inserito il grafico e premere il solito `Ctrl + V`).

6 I fit

Ammettiamo di possedere un set di dati sperimentali riguardanti due grandezze fisiche, x e y, e di sapere, grazie a delle considerazioni sulla natura del fenomeno, che tra esse intercorre una relazione di tipo lineare (ovvero che, all'aumentare dell'una, aumenti o diminuisca l'altra con una certa "velocità" costante). Ci si potrebbe chiedere quale sia l'equazione della retta che meglio possa approssimare la distribuzione dei dati sperimentali sul piano xy, ovvero quel grafico ottenuto disegnando un punto per ogni coppia di misure, utilizzando come coordinate, rispettivamente, i valori di x e di y. La risposta a questa domanda sarebbe costituita, in pratica, da due numeri[15] affetti, evidentemente, da una certa incertezza dovuta all'imprecisione delle misure. In parole povere, la situazione è questa: se le misure fossero perfette

[15]Essendoci una relazione lineare, cioè di primo grado, l'equazione della "curva" in questione (in questo caso, di una retta) sarebbe del tipo:

$$y = mx + q$$

dove il coefficiente del termine di primo grado, m, è detto *coefficiente angolare* e misura la "pendenza" della retta mentre il termine di grado zero, q, è detto *intercetta* e indica il valore dell'ordinata con la quale la retta attraversa l'asse y.

e, quindi, le loro incertezze nulle, i punti del grafico di cui sopra dovrebbero giacere tutti su una stessa retta la cui pendenza e “altezza” sarebbero descritte dalla sua equazione, la quale, in termini fisici, rappresenterebbe la “formula” associata a quel dato fenomeno fisico preso in esame. In verità, essendo le misure affette da incertezza, i punti si trovano “sparpagliati” attorno ad una fantomatica retta della quale si può solo stimare l’equazione. Il procedimento statistico per effettuare questa stima, si chiama *fit lineare* (detto anche *regressione lineare* o *metodo dei minimi quadrati*). Nel caso in cui la relazione tra le due grandezze non sia lineare, questo procedimento[16] va sotto il nome più generale di *fit dei dati sperimentali*.

Effettuare un fit personalizzato

KaleidaGraph™ effettua i più frequenti e utili tipi di fit grazie a metodi iterativi di calcolo numerico. Per ottenere un fit occorre, anzitutto, selezionare il grafico su cui lavorare. Dopodiché occorre selezionare, dal menù `Curve Fit`, il tipo di fit da effettuare e, successivamente, la colonna di dati a cui riferirsi tramite la finestra `Curve Fit Selections`: nella scelta del tipo di fit, cliccando su `General` si ha la possibilità di creare nuovi tipi di fit personalizzati o di utilizzare quelli creati precedentemente. Selezionando `Edit General` si può gestire il menù dei fit personalizzati: cliccando su `Add` si crea un nuovo fit, al quale può essere assegnato un nuovo nome (sostituendo quello di default `'New Fit'`) e modificata l’equazione. Per effettuare quest’ultima operazione bisogna, dopo aver selezionato il fit dalla lista a sinistra, cliccare sul tasto `Edit`: nella casella di testo centrale va inserita l’equazione utilizzando, come coefficienti del fit, le lettere latine minuscole[17] e ricordandosi che le funzioni riconosciute sono quelle elencate nella tabella a pag. 7. Ad esempio, se ci si vuol chiedere quale sia la miglior parabola che approssimi i dati sperimentali, occorre scrivere nell’apposita casella di testo:

```
a*x^2+b*x+c
```

o, se si vuole trovare la miglior curva esponenziale, basta digitare:

```
a*exp(b*x)
```

[16] Per spiegare matematicamente quali siano gli specifici calcoli da effettuare, si dovrebbe semplicemente risolvere un sistema nel quale le derivate parziali degli scarti (quadratici) sono poste a zero (*minimizzazione degli scarti*).

[17] Il programma usa di default etichette diverse, costituite dalla lettera `M` seguita da un numero progressivo.

Per effettuare un *fit gaussiano*, ovvero per riprodurre nel grafico della distribuzione delle frequenze la curva di Gauss che meglio l'approssima[18] (vedi la fig. 5 a pag. 14), bisogna creare un nuovo fit avente equazione:

`gaussfit(a,b,c,d)`

KaleidaGraph™ prevede diversi tipi di fit predefiniti che elenchiamo di seguito, ricordando che il metodo del `General Fit` rimane quello più efficiente grazie alla sua versatilità:

`Linear` : corrisponde al fit lineare di equazione `a+b*x`.

`Polynomial` : permette di effettuare un fit polinomiale di equazione `a+b*x+c*x^2+...+i*x^9`. Dopo averlo selezionato, occorre informare il programma del grado del polinomio da trovare, tramite la finestra `Polynomial Order`.

`Exponential` : equivale al fit generale di equazione `a*exp(b*x)` e non è applicabile sui dati minori o uguali a zero.

`Logarithmic` : ha equazione `a+b*log(x)` e non può essere applicato a dati non positivi.

`Power` : corrisponde a `a*x^b` ed è sottoposto alle stesse limitazioni di dominio dei due casi precedenti.

`Smooth` : utilizzando una *funzione di Stineman*, KaleidaGraph™ disegna la curva più regolare possibile che approssimi i dati al $\pm$ 10%.

Effettuando uno qualsiasi dei fit sopraelencati, all'infuori di `smooth`, si ottiene, oltre che all'aggiunta della curva del fit sul grafico, la visualizzazione di una tabella (vedi la fig. 6 di pag. 16) o di una casella di testo, avente i caratteri di colore rosso, nella quale sono riportate le stime numeriche dei vari parametri (o l'equazione della curva del fit) e, in alcuni casi, dei parametri aggiuntivi che ci informano sull'accuratezza dell'approssimazione come `R` (*coefficiente di correlazione lineare*: un valore vicino a 1 corrisponde ad una buona approssimazione raggiunta) e `Chisq` (*chi quadro*[19]: in questo caso, più piccolo è, meglio il fit approssima i dati). Per fare in modo che queste informazioni vengano o meno visualizzate occorre selezionare la voce `Display Equation` dal menù `Plot` o dal menù pop-up che appare cliccando col tasto destro del mouse sul grafico.

[18] I parametri di questo fit sono, evidentemente, la media e la deviazione standard dalla media.

[19] Il valore del chi quadro è pari alla somma degli scarti quadratici tra i dati sperimentali e quelli calcolati con il fit.

7 Un esempio: il pendolo semplice

Per ricapitolare le modalità di utilizzo delle più importanti e utili funzionalità di KaleidaGraph^{TM}, immaginiamo di dover elaborare alcune misure sperimentali riguardanti il fenomeno fisico del moto di un pendolo semplice. Per minimizzare l'errore sistematico dovuto ai riflessi del "cronometrista", conviene misurare il periodo di cinque oscillazioni complete: in questo caso, però, bisogna dividere tutti i dati sperimentali per cinque, in modo da ottenere una media aritmetica dei periodi di un'oscillazione singola. Occorre, quindi, utilizzare la semplice formula `c0=c0/5` immettendola nella casella di testo della finestra `Formula Entry`, selezionare la colonna delle misure dei periodi e premere il tasto `Run`.

A questo punto, per ottenere il periodo medio e la sua incertezza (considerando trascurabili gli errori sistematici) occorre utilizzare la funzione `Statistics` del menù `Functions`, magari ricopiando la tabella ottenuta nella relazione che si sta compilando, utilizzando il tasto `Clipboard`, ed incollandola nell'editor di testi. Se, ad esempio, avessimo ottenuto la finestra `Statistics` riportata nella fig. 2 di pag. 8, potremmo affermare che il periodo vale:

$$T = (\texttt{1,721} \pm \texttt{0,003})\ \texttt{s}$$

Adesso si potrebbe disegnare la distribuzione delle frequenze arricchendola con la curva di Gauss che meglio l'approssima. Per fare ciò, bisogna cliccare su `Functions >> Bin Data`, poi sul tasto `Clipboard` e, successivamente, dopo aver selezionato due colonne vuote della tabella dei dati, premere la combinazione di tasti `Ctrl + V`. A questo punto, si clicca su `Curve Fit >> General >> Edit General`, poi su `Add` e, dopo aver rinominato il `'New Fit'`, su `Edit`. Ora basta scrivere nella casella di testo `gaussfit(a,b,c,d)` e cliccare su `Ok` e poi, nuovamente, su `Ok`, per chiudere la finestra `General Curve Fits`. Infine occorre, semplicemente, selezionare il fit appena creato da `Curve Fit >> General` e, dopo aver selezionata la colonna di dati interessata (l'unica visualizzata), cliccare su `Ok`.

Supponiamo, ora, di eseguire sei diversi gruppi di misurazioni di periodi (ogni gruppo formato da un congruo numero di rilievi) corrispondenti a sei diverse lunghezze di filo e di ottenere una tabella avente sei righe e tre colonne: lunghezza del filo, periodo medio e deviazione standard dalla media (calcolata tramite la solita funzione `Statistics` per ogni gruppo di misurazioni). Con questi elementi si può tentare di ottenere un grafico, corredato di barre d'incertezza, della distribuzione delle due grandezze fisiche (lunghezza e periodo) e di un fit che metta in evidenza le loro modalità di correlazione. Per creare il grafico basta cliccare su `Gallery >> Linear >> Scatter`, selezionare la colonna delle lunghezze come X e quella dei periodi come Y, e cliccare su `New Plot`. Per aggiungere le barre d'incertezza occorre: cliccare su `Error Bars` dal menu pop-up (tasto destro del mouse sul grafico), cliccare su X e immettere l'incertezza della lunghezza[20] nella casella `Fixed Error`, cliccare su `Ok` e poi su Y, selezionare la voce `Data Column` dal menù a tendina, scegliere la colonna delle deviazioni standard dalla media, cliccare su `Ok` e, poi, di nuovo sull'`Ok` della finestra `Error Bar Variables`. Per ottenere il fit bisognerà crearne uno nuovo e poi selezionarlo come si è già fatto per quello gaussiano (visto in precedenza) usando, questa volta, come equazione: `a+b*x^c` (ottenibile mediante considerazioni di carattere dimensionale). Si dovrebbe ottenere un risultato simile a quello riportato nella fig. 6 di pag. 16.

I grafici ottenuti andranno, infine, copiati nella relazione che si sta compilando, premendo la solita combinazione `Ctrl + C` dopo averli selezionati, e incollati tramite la specifica funzione dell'editor di testi utilizzato.

[20]Di solito, per le misurazioni di lunghezza, l'incertezza è pari alla metà del minimo intervallo apprezzabile.