
Misure ed analisi dei dati:

introduzione al Laboratorio di Fisica

Liana Martinelli

Luca Baldini

Quinta edizione, Ottobre 2014

Compilato il 15 Ottobre 2020 alle ore 14:02

Copyright (C) 2014 Liana Martinelli, Luca Baldini.



This work is licensed under a
Creative Commons Attribution-ShareAlike 4.0 International License.

PREFAZIONE

Il materiale raccolto in queste dispense proviene sostanzialmente da due nuclei fondamentali: le dispense originali tratte dalle lezioni del corso di Esperimentazioni di Fisica I per Fisici, tenute da L. M. presso l'Università degli Studi di Pisa, ed una breve introduzione all'uso del calcolatore nei laboratori didattici, originariamente distribuita direttamente tra gli studenti del corso. Benché molte altre cose si siano via via aggiunte con il passare del tempo, l'organizzazione concettuale del materiale riflette la presenza di questi due nuclei originali.

Nella prima parte si introducono alcuni tra i fatti fondamentali relativi alla propagazione delle incertezze nelle misure fisiche, al calcolo delle probabilità (con particolare attenzione alle distribuzioni più spesso usate in fisica) ed al trattamento statistico dei dati (incluse le più comuni tecniche di *fit*). Questa parte contiene inoltre una trattazione, sia pure sommaria, di alcuni argomenti appena più avanzati (correlazione, distribuzione *t*, cambiamento di variabile nelle funzioni di distribuzione) che, pur non essendo così pesantemente usati nelle esperienze del primo anno, devono a buon diritto far parte del bagaglio culturale del fisico medio. Notiamo, per inciso, che, come i numerosi esempi dimostrano, la discussione è tendenzialmente focalizzata sul significato fisico delle nozioni introdotte, più che sul formalismo, e che, come è ovvio, il tono dell'esposizione non è sempre completamente rigoroso.

Lo scopo fondamentale della seconda parte è quello di rendere il meno traumatico possibile l'incontro con il calcolatore degli studenti del primo anno e, anche là dove questo incontro si rivela tutto sommato indolore, promuovere un uso *consapevole* del calcolatore stesso (cosa spesso tutt'altro che scontata). Questo include (ma non è limitato a) il conseguimento della capacità di utilizzare il *computer* per rappresentare ed analizzare dati sperimentali e (perché no?) presentarli in una forma esteticamente appetibile.



Il simbolo di curva pericolosa che ricorre occasionalmente, nel seguito, ha lo scopo di attrarre l'attenzione su alcuni passaggi particolarmente delicati che meritano di essere letti almeno due volte.

Gli autori desiderano ringraziare sinceramente Alessandro Strumia, che per primo ha trascritto in \LaTeX il contenuto originale delle lezioni di Esperimentazioni di Fisica I e tutti i docenti (in particolare Stefano Sanguinetti) e gli studenti (in particolare Francesca Menozzi) che, segnalando errori ed imprecisioni, hanno contribuito a rendere queste dispense più comprensibili ed accurate.

Pisa, 13 Agosto 2007

L. Martinelli

L. Baldini

INDICE

I	STATISTICA ED ANALISI DEI DATI	1
1	PROPAGAZIONE DEGLI ERRORI. PARTE I	5
1.1	Terminologia	5
1.2	Propagazione dell'incertezza massima (errore massimo)	7
1.2.1	Somma	8
1.2.2	Differenza	9
1.2.3	Prodotto	10
1.2.4	Quoziente	12
1.2.5	Caso generale	13
1.3	Errore relativo	16
1.3.1	Propagazione dell'errore relativo	16
1.4	Cifre significative e convenzioni di scrittura	17
1.5	Esercizi	22
2	PROBABILITÀ E STATISTICA	25
2.1	Definizione di probabilità	26
2.1.1	Prima definizione (Laplace, 1812)	26
2.1.2	Seconda definizione (Venn, 1866)	27
2.1.3	Terza definizione (Kolmogorov, 1933)	27
2.2	Proprietà e leggi della probabilità	28
2.2.1	Addizione	28
2.2.2	Moltiplicazione	29
2.3	Variabili casuali e funzioni di distribuzione	32
2.3.1	Variabili casuali discrete	33
2.3.2	Variabili casuali continue	33
2.4	Rappresentazione di fenomeni casuali	35
2.5	Caratteristiche comuni alle distribuzioni	40
2.5.1	Premessa	40
2.5.2	Notazioni	41
2.5.3	Definizioni di base: normalizzazione e valor medio	41
2.5.4	Valore di aspettazione o di previsione	46
2.5.5	Dispersione intorno alla media: varianza e deviazione standard	46
2.5.6	La semilarghezza a metà altezza	49
2.5.7	Momenti di una distribuzione	51
2.5.8	Discussione e cenno al Teorema di Tschebyscheff	52
2.5.9	Cenno alla teoria dei campioni	53
2.6	Esercizi	55
3	DISTRIBUZIONI	59
3.1	Distribuzione binomiale	60
3.1.1	Premessa: i coefficienti binomiali	60

3.1.2	La distribuzione binomiale	60
3.1.3	Normalizzazione, media e varianza	63
3.2	Distribuzione di Poisson	65
3.2.1	Premessa: il numero di Nepero	65
3.2.2	La distribuzione di Poisson	66
3.2.3	Normalizzazione, media e varianza	70
3.3	Distribuzione uniforme	72
3.3.1	Normalizzazione, media e varianza	72
3.4	Distribuzione esponenziale	75
3.4.1	Normalizzazione, media e varianza	77
3.5	Distribuzione di Gauss	79
3.5.1	Premessa: la funzione di Gauss	79
3.5.2	La distribuzione di Gauss	80
3.5.3	Normalizzazione, media e varianza	82
3.5.4	Integrazione della distribuzione di Gauss	83
3.5.5	Relazione con altre distribuzioni	84
3.6	Distribuzione del χ^2	88
3.6.1	Premessa: la funzione Γ di Eulero	88
3.6.2	La distribuzione del χ^2	88
3.6.3	Normalizzazione, media e varianza	89
3.7	Distribuzione di Cauchy	90
3.7.1	Normalizzazione, media e varianza	91
3.8	Sommario	95
3.9	Esercizi	96
4	PROPAGAZIONE DEGLI ERRORI. PARTE II	99
4.1	Media e varianza di funzioni di variabili casuali	99
4.1.1	Funzioni di una sola variabile	100
4.1.2	Funzioni di due variabili	102
4.1.3	La covarianza	104
4.1.4	Caso generale	105
4.2	Media e varianza campione	107
4.3	Il teorema del limite centrale	109
4.3.1	Premessa	109
4.3.2	Il teorema del limite centrale	110
4.4	Intervalli di confidenza	114
5	METODI DI FIT	117
5.1	Problemi di fit	117
5.2	Metodo dei minimi quadrati	118
5.2.1	Fit dei minimi quadrati nel caso di una funzione costante	119
5.2.2	Fit dei minimi quadrati nel caso di una funzione lineare	120
5.3	Metodo del minimo χ^2	123
5.3.1	Fit del minimo χ^2 per una funzione costante: la media pesata	123
5.3.2	Fit del minimo χ^2 nel caso lineare	125

5.4	Fit di tipo generale	128
5.5	Test del χ^2	130
5.5.1	Test del χ^2 per una serie di misure	131
5.5.2	Test del χ^2 per una distribuzione	132
5.6	Esercizi	135
6	FORMULE APPROSSIMATE	137
6.1	Derivata di una funzione data per punti	137
6.2	Integrale di una funzione data per punti	139
6.3	Interpolazione	140
7	COMPLEMENTI	141
7.1	Correlazione	141
7.1.1	Coefficiente di correlazione	143
7.2	La distribuzione di Student	145
7.2.1	Variabile t di Student a n gradi di libertà	145
7.2.2	Variabile t di Student a n-1 gradi di libertà	146
7.2.3	t-test	147
7.3	Funzione di distribuzione per funzioni di variabile casuale	149
7.3.1	Variabili casuali discrete	149
7.3.2	Variabili casuali continue	150
7.4	Esercizi	157
II	INTRODUZIONE AL CALCOLATORE	159
8	IL SISTEMA OPERATIVO LINUX	163
8.1	Terminologia	164
8.2	Una breve storia di Linux	164
8.3	Concetti di base	165
8.3.1	Il <i>login</i>	166
8.3.2	La <i>shell</i> ed i comandi	166
8.3.3	Il <i>logout</i>	167
8.4	Il <i>filesystem</i> Linux	167
8.5	Navigare il <i>filesystem</i>	168
8.5.1	<i>pwd</i>	168
8.5.2	<i>ls</i>	168
8.5.3	<i>cd</i>	170
8.6	Modificare il <i>filesystem</i>	170
8.6.1	<i>mkdir</i>	171
8.6.2	<i>rmdir</i>	171
8.6.3	<i>rm</i>	172
8.6.4	<i>cp</i>	172
8.6.5	<i>mv</i>	172
8.6.6	<i>chmod</i>	173
8.7	Visualizzare i <i>file</i> di testo	173
8.7.1	<i>more</i>	173
8.7.2	<i>less</i>	173
8.7.3	<i>head</i>	174

8.7.4	tail	174
8.8	Modificare i <i>file</i> di testo	174
8.9	I processi	174
8.9.1	ps	175
8.9.2	top	175
8.9.3	kill	175
8.10	La stampa	176
8.10.1	lpr	176
8.10.2	lpq	176
8.10.3	lprm	176
8.11	Accesso alle periferiche	176
8.11.1	mount	177
8.11.2	umount	178
9	SCRIVERE DOCUMENTI SCIENTIFICI: L ^A T _E X	179
9.1	Introduzione	179
9.2	Dalla stesura alla stampa	180
9.3	Il primo documento L ^A T _E X	182
9.4	Un documento realistico	184
9.5	Elenchi	187
9.6	Tabelle	188
9.7	L ^A T _E X e la matematica	190
9.8	Inserire le figure	192
10	VISUALIZZARE ED ANALIZZARE DATI: GNUPLOT	195
10.1	Lanciare gnuplot	196
10.2	Due cose semplici (ma utili)...	197
10.3	... e due cose meno ovvie	199
10.4	Visualizzare una serie di dati	201
10.5	Realizzare un istogramma	204
10.6	Salvare i grafici sotto forma di immagini	205
10.7	Il concetto di macro	206
10.8	Eseguire un fit ad una serie di dati	207
10.9	Operazioni con le colonne	212
A	TAVOLE NUMERICHE	213
A.1	Definizione della funzione erf(<i>x</i>)	213
A.2	Integrale normale degli errori - I	215
A.3	Integrale normale degli errori - II	216
A.4	Distribuzione del χ^2 - I	217
A.5	Distribuzione del χ^2 - II	218
A.6	Distribuzione del χ^2 - III	219
A.7	Distribuzione t di Student - I	220
A.8	Distribuzione t di Student - II	221
B	UN SEMPLICE DOCUMENTO L ^A T _E X	223
C	SIMBOLI MATEMATICI IN L ^A T _E X	227
D	BREVE GLOSSARIO DI GNUPLOT	229

E UN ESEMPIO DI FIT CON SCILAB	233
E.1 Minimo χ^2 nel caso lineare	233
E.2 Implementare il fit con scilab	234
Bibliografia	237

Parte I

STATISTICA ED ANALISI DEI DATI

INTRODUZIONE

Quando un fisico vuole misurare qualche cosa si pone sostanzialmente due problemi:

- (1) Come si misura, cioè come si può ottenere il risultato.
- (2) Quanto è giusto il risultato, o (il che è lo stesso) quanto è sbagliato.

Si noti che il secondo punto è essenziale perché altrimenti anche il primo perde gran parte del suo significato. Infatti il risultato di una misura è soltanto una approssimazione o stima del valore della grandezza specifica che si vuole misurare (il *misurando*) ed il risultato è completo quando è accompagnato da una quantità che definisce l'incertezza. Ci sono due situazioni tipiche che limitano la precisione delle misure:

(1) Lo strumento ha dei limiti. Si parla allora di *incertezza strumentale* ed il risultato delle misure in genere è costante. Per esempio nella misura di un tavolo con un metro a nastro si ottiene sempre lo stesso risultato entro la risoluzione dello strumento cioè entro il millimetro.

(2) Lo strumento è molto raffinato, ma il fenomeno che si studia sfugge al controllo completo (almeno in qualche sua caratteristica secondaria). Allora si hanno risultati diversi da misura a misura: si dice che si hanno *fluttuazioni casuali*.

È chiaro che nella fisica moderna ha molta più importanza la seconda situazione perché quasi sempre si lavora al limite delle tecniche strumentali. Dovremo quindi studiare come ci si deve comportare quando i risultati delle misure fluttuano in modo imprevedibile e casuale: gli strumenti matematici necessari sono la probabilità e la statistica.

PROPAGAZIONE DEGLI ERRORI. PARTE I

INDICE

1.1	Terminologia	5	
1.2	Propagazione dell'incertezza massima (errore massimo)		7
1.2.1	Somma	8	
1.2.2	Differenza	9	
1.2.3	Prodotto	10	
1.2.4	Quoziente	12	
1.2.5	Caso generale	13	
1.3	Errore relativo	16	
1.3.1	Propagazione dell'errore relativo	16	
1.4	Cifre significative e convenzioni di scrittura		17
1.5	Esercizi	22	

Dopo un breve paragrafo di terminologia, in questo capitolo si introducono i fatti fondamentali relativi alla propagazione dell'errore massimo nelle misure indirette ed alle regole e convenzioni di scrittura per i dati sperimentali.

1.1 TERMINOLOGIA

Vi sono alcuni termini che entrano comunemente nel linguaggio del lavoro sperimentale: accuratezza, ripetitività, riproducibilità, errore, errore casuale, errore sistematico, incertezza, precisione. È opportuno che questi termini vengano usati in modo consistente secondo la consuetudine della comunità scientifica internazionale. Pertanto riportiamo qui alcune definizioni seguendo il VIM (*International Vocabulary of Basic and General Terms in Metrology*).

ACCURATEZZA (*della misura*). Una misura è tanto più *accurata* quanto maggiore è l'accordo tra il risultato della misura stessa ed il valore del misurando.

RIPETITIVITÀ (*dei risultati di misura*). Le misure si dicono ripetitive quanto più sono vicini i risultati di successive misure dello stesso misurando nelle stesse condizioni di misura. Queste condizioni si chiamano condizioni di ripetitività ed includono: lo stesso procedimento di misura, lo stesso osservatore, lo stesso strumento di misura usato nelle stesse condizioni, lo stesso posto, ripetizione della misura in brevi periodi di

tempo. La dispersione caratteristica dei risultati permette di definire quantitativamente l'incertezza dei risultati.

RIPRODUCIBILITÀ (*dei risultati di misura*). Le misure si dicono riproducibili quanto più sono in accordo i risultati di misure fatte in condizioni diverse. Questo significa che possono essere diversi, per esempio, il principio di misura, il metodo di misura, l'osservatore, gli strumenti di misura, il luogo, il tempo...

ERRORE (*di misura*). Il risultato di una misura meno il valore del misurando. Questo è un concetto qualitativo: in generale l'errore di misura non è noto perché non è noto il valore del misurando. Tuttavia può essere valutata l'incertezza del risultato di una misura.

ERRORE CASUALE. Il risultato di una misura meno la media che dovrebbe risultare da un infinito numero di misure dello stesso misurando fatte in condizioni di ripetitività.

ERRORE SISTEMATICO. Differenza tra la media che dovrebbe risultare da un infinito numero di misure dello stesso misurando fatte in condizioni di ripetitività ed il valore del misurando. L'errore sistematico è pertanto l'errore di misura meno l'errore casuale. L'errore sistematico e le sue cause non possono essere completamente noti. Spesso gli errori *sistematici* sono quelli che derivano da errori di taratura degli strumenti o dell'apparato sperimentale o dalla mancanza di imparzialità dell'osservatore.

L'organizzazione internazionale di metrologia legale (OIML, 1968) definisce l'errore sistematico come un errore che nel corso di un certo numero di misure, fatte nelle stesse condizioni, dello stesso valore di una data quantità, rimane costante in valore assoluto e segno o varia secondo una legge definita quando cambiano le condizioni. È evidente che gli errori sistematici limitano l'accuratezza della misura: quanto più si riescono ad eliminare gli errori sistematici, tanto più la misura è accurata.

► **ESEMPIO 1.1.** Si supponga di misurare il periodo di un pendolo con un cronometro che ha la risoluzione di $\frac{1}{5}$ di secondo. Se il cronometro va avanti (cioè le lancette girano più in fretta di quanto dovrebbero) l'errore associato al risultato che si legge è potenzialmente molto più grande di $\frac{1}{5}$ di secondo. In ogni caso non è facile stabilire se ci sono errori sistematici a meno di non disporre di un orologio campione che segna l'ora esatta (per definizione). □

► **ESEMPIO 1.2.** Si supponga di voler verificare la legge di Boyle misurando a temperatura ambiente pressione e volume dell'aria contenuta nel ramo chiuso di un tubo ad U contenente Hg. Le misure di pressione non sono in torr perché la temperatura non è a 0 °C. Quindi le stime della costante PV, pur essendo tutte consistenti tra di loro, non servono a calcolare la costante dei gas R. □

INCERTEZZA (*di misura*). È un intervallo al quale contribuiscono varie componenti che possono sostanzialmente essere distinte in due categorie: quelle che sono valutate con metodi statistici e quelle che sono valutate con altri metodi.

Bisognerebbe sempre tenere in mente la differenza tra *errore* ed *incertezza*. Per esempio il risultato di una misura dopo correzioni per compensare effetti sistematici individuati può essere molto vicino (anche se non si può sapere di quanto) al valore incognito del misurando, ed avere così un errore trascurabile, ma avere una incertezza grande.

PRECISIONE (*dei risultati di misura*). Per questa parola si trovano in letteratura varie definizioni, ed il VIM non la definisce. Di solito il concetto di precisione viene legato sia alla ripetitività che alla riproducibilità delle misure, cioè alla consistenza interna dell'insieme delle misure. Quando l'errore sistematico è trascurabile (o si ritiene che lo sia) l'incertezza viene solo dalle fluttuazioni dei risultati o dalla risoluzione degli strumenti e la precisione di una misura viene identificata con l'incertezza relativa, cioè il rapporto tra l'incertezza e il risultato di misura.

Il termine *precisione* non dovrebbe mai essere usato come sinonimo di *accuratezza*, che è un concetto soltanto qualitativo.

1.2 PROPAGAZIONE DELL'INCERTEZZA MASSIMA (ERRORE MASSIMO)

Quando la dispersione delle misure fatte in condizioni di ripetitività è nulla, l'incertezza deriva soltanto dalla risoluzione degli strumenti. Se la misura è diretta l'incertezza è data immediatamente dalla risoluzione dello strumento, se la misura è indiretta l'incertezza è data da una opportuna combinazione delle risoluzioni degli strumenti utilizzati come preciseremo tra breve.

► ESEMPIO 1.3. Si supponga di voler stimare il volume V di un cilindro di raggio r ed altezza h ; siano r_0 ed h_0 le quantità misurate, con incertezze Δr e Δh rispettivamente.

$$\begin{aligned} r &= r_0 \pm \Delta r \\ h &= h_0 \pm \Delta h \end{aligned}$$

Ovviamente vorremmo scrivere il risultato nella forma $V = V_0 \pm \Delta V$. Si pone allora il problema di vedere come le incertezze sulle misure dirette influenzino l'errore ΔV da associare a V . Supponiamo, per fissare le idee, che:

$$\begin{aligned} r &= (10.2 \pm 0.1) \text{ mm} \\ h &= (23.4 \pm 0.1) \text{ mm} \end{aligned}$$

È estremamente semplice calcolare i valori massimo e minimo che il volume può assumere, date le incertezze sul raggio di base e sull'altezza; essi corrisponderanno alle situazioni in cui r ed h assumono i valori massimi e minimi, rispettivamente:

$$\begin{aligned} V_{\max} &= \pi(r_0 + \Delta r)^2 \cdot (h_0 + \Delta h) \approx 7830 \text{ mm}^3 \\ V_{\min} &= \pi(r_0 - \Delta r)^2 \cdot (h_0 - \Delta h) \approx 7470 \text{ mm}^3 \end{aligned}$$

Possiamo allora stimare V_0 e ΔV_0 come media e semidispersione di tali valori:

$$\begin{aligned} V_0 &= \frac{V_{\max} + V_{\min}}{2} \approx 7650 \text{ mm}^3 \\ \Delta V_0 &= \frac{V_{\max} - V_{\min}}{2} \approx 180 \text{ mm}^3 \end{aligned}$$

Si tratta adesso di sviluppare un formalismo che ci permetta di ricavare questo risultato nel modo più semplice e generale possibile, senza passare ogni volta attraverso il calcolo esplicito dei valori massimo e minimo della grandezza cui siamo interessati. In altre parole si tratta di capire come Δh e Δr si combinino per dare ΔV . \square

Nei paragrafi che seguono studieremo la propagazione dell'errore nelle quattro operazioni e, di seguito, per una arbitraria combinazione funzionale delle misure sperimentali. Supporremo sempre che $a_0, b_0, c_0 \dots$ siano i valori *misurati* delle grandezze $a, b, c \dots$ cui siamo interessati e che $\Delta a, \Delta b, \Delta c \dots$ siano le incertezze associate (nel senso che $a, b, c \dots$ possono assumere *solamente* valori entro $a_0 \pm \Delta a, b_0 \pm \Delta b, c_0 \pm \Delta c \dots$). Supporremo anche che le quantità in questione siano tra loro *indipendenti* (nel senso che la misura di una non influenza la misura delle altre); torneremo nel paragrafo 7.1 a discutere che cosa accade se questa assunzione viene meno. Ci poniamo ora il problema di definire, data una grandezza G funzione delle grandezze che misuriamo:

$$G = f(a, b, c, \dots)$$

un intervallo $[G_0 - \Delta G, G_0 + \Delta G]$ in cui si è *certi* di trovare la grandezza stessa, supponendo noti i rispettivi intervalli in cui si è certi di trovare $a, b, c \dots$.

Notiamo esplicitamente che tutti i risultati che otterremo in questo capitolo si riferiscono a quello che viene solitamente definito *errore massimo*. Nel capitolo 4 presenteremo una formulazione del problema statisticamente più corretta.

1.2.1 Somma

Sia S la somma:

$$S = a + b$$

Costruiamo esplicitamente S mettendoci nei due casi peggiori, in cui sia Δa che Δb contribuiscono nello stesso verso:

$$\begin{aligned} S_{\max} &= (a_0 + \Delta a) + (b_0 + \Delta b) = (a_0 + b_0) + (\Delta a + \Delta b) \\ S_{\min} &= (a_0 - \Delta a) + (b_0 - \Delta b) = (a_0 + b_0) - (\Delta a + \Delta b) \end{aligned}$$

Ne segue che possiamo scrivere:

$$S = (a_0 + b_0) \pm (\Delta a + \Delta b)$$

per cui assumeremo come miglior valore S_0 per la somma la quantità:

$$\boxed{S_0 = a_0 + b_0} \quad (1)$$

e come errore *massimo* ΔS le quantità:

$$\Delta S = \Delta a + \Delta b \quad (2)$$

► ESEMPIO 1.4. Supponiamo di misurare due lunghezze:

$$a = (2.50 \pm 0.05) \text{ cm}$$

$$b = (2.34 \pm 0.05) \text{ cm}$$

Date la (1) e la (2), scriveremo la somma come:

$$a + b = (4.84 \pm 0.10) \text{ cm}$$

□

1.2.2 Differenza

Sia D la differenza:

$$D = a - b$$

Come prima consideriamo i casi più sfavorevoli:

$$D_{max} = (a_0 + \Delta a) - (b_0 - \Delta b) = (a_0 - b_0) + (\Delta a + \Delta b)$$

$$D_{min} = (a_0 - \Delta a) - (b_0 + \Delta b) = (a_0 - b_0) - (\Delta a + \Delta b)$$

Ne segue che:


$$D = (a_0 - b_0) \pm (\Delta a + \Delta b)$$

e, come prima, la miglior stima per la differenza sarà:

$$D_0 = a_0 - b_0 \quad (3)$$

mentre prenderemo come errore:

$$\Delta D = \Delta a + \Delta b \quad (4)$$

 Per coloro che, in analogia con la (2), fossero tentati di scrivere $\Delta D = \Delta a - \Delta b$, notiamo esplicitamente come questo sia sbagliato. Per convincersene basti pensare che, nel caso di due misure con la stessa incertezza, questa espressione darebbe zero, il che è chiaramente assurdo. In effetti il modo giusto di propagare l'errore massimo nella differenza è dato dalla 4; purtroppo, in generale, gli errori si sommano e non si sottraggono.

► ESEMPIO 1.5. Supponiamo di misurare due lunghezze:

$$a = (2.50 \pm 0.05) \text{ cm}$$

$$b = (2.34 \pm 0.05) \text{ cm}$$

e di essere interessati alla differenza. Seguendo la (3) e la (4), la scriveremo come:

$$a - b = (0.16 \pm 0.10) \text{ cm}$$

Notiamo che, benché le due misure di a e b siano abbastanza precise, la differenza risulta sostanzialmente indeterminata (nel senso che l'incertezza sulla differenza è dello stesso ordine di grandezza della differenza stessa). Si tratta di un problema comune quando si sottraggono misure vicine tra loro. \square

1.2.3 Prodotto

Sia P il prodotto:

$$P = a \cdot b$$

I casi peggiori sono, al solito:

$$P_{max} = (a_0 + \Delta a)(b_0 + \Delta b) = a_0 \cdot b_0 + (a_0 \Delta b + b_0 \Delta a + \Delta a \Delta b)$$

$$P_{min} = (a_0 - \Delta a)(b_0 - \Delta b) = a_0 \cdot b_0 - (a_0 \Delta b + b_0 \Delta a - \Delta a \Delta b)$$

Se valgono le condizioni:

$$\Delta a \ll a$$

$$\Delta b \ll b$$

che è la situazione tipica in laboratorio, possiamo trascurare il prodotto $\Delta a \Delta b$, come illustrato nell'esempio seguente.

► ESEMPIO 1.6. Torniamo all'esempio 1.5 e supponiamo adesso di voler calcolare il prodotto tra le due grandezze misurate. Per valutare l'incertezza associata ci occorrono:

$$a_0 \Delta b = 0.125 \text{ cm}^2$$

$$b_0 \Delta a = 0.117 \text{ cm}^2$$

$$\Delta a \Delta b = 0.0025 \text{ cm}^2$$

Dunque, effettivamente, il terzo termine è trascurabile rispetto agli altri due. \square

In queste condizioni si ha:

$$P_{max} \simeq a_0 \cdot b_0 + (a_0 \Delta b + b_0 \Delta a)$$

$$P_{min} \simeq a_0 \cdot b_0 - (a_0 \Delta b + b_0 \Delta a)$$

e dunque possiamo scrivere l'espressione per la migliore stima del prodotto come:

$$P_0 = a_0 \cdot b_0 \quad (5)$$

e quella dell'errore associato come:

$$\Delta P = a_0 \Delta b + b_0 \Delta a \quad (6)$$



È importante sottolineare come la (6) non sia una relazione esatta, ma valga solamente nell'ipotesi in cui i termini di ordine superiore al primo siano, come precisato prima, trascurabili.

La (6) può essere utilizzata per calcolare l'incertezza massima in un certo numero di casi interessanti, come mostrato negli esempi seguenti.

► **ESEMPIO 1.7.** Supponiamo di misurare una grandezza a ottenendo il valore $a_0 \pm \Delta a$. Propaghiamo l'errore su a^2 . Notando che $a^2 = a \cdot a$ e sfruttando le equazioni (5) e (6) scriveremo:

$$\begin{aligned} a^2 &= a_0 \cdot a_0 = a_0^2 \\ \Delta(a^2) &= a_0 \Delta a + a_0 \Delta a = 2a_0 \Delta a \end{aligned}$$

□

► **ESEMPIO 1.8.** Ripetiamo l'esempio precedente propagando l'errore su a^3 . Esattamente come prima, notando che $a^3 = a^2 \cdot a$:

$$\begin{aligned} a^3 &= a_0^2 \cdot a_0 = a_0^3 \\ \Delta(a^3) &= a_0^2 \Delta a + a_0 \Delta(a^2) = a_0^2 \Delta a + a_0 \cdot 2a_0 \Delta a = 3a_0^2 \Delta a \end{aligned}$$

□

Si può dimostrare (come vedremo tra breve) che l'errore massimo sulla potenza n -esima a^n si scrive come:

$$\Delta(a^n) = n a_0^{n-1} \Delta a \quad (7)$$

Notiamo infine un'ultima conseguenza particolarmente interessante della (6), nel caso in cui la grandezza misurata a sia moltiplicata per un numero reale c non affetto da errore (cioè $\Delta c = 0$):

$$\Delta(c \cdot a) = c \Delta a + a \Delta c = c \Delta a \quad (8)$$

Ne vediamo un'applicazione nell'esempio che segue.

► **ESEMPIO 1.9.** Si vuole misurare lo spessore s di un foglio di carta disponendo solo di una riga millimetrata. A tale scopo si misura lo spessore h di una risma di 500 fogli identici ottenendo il valore:

$$h = (48 \pm 1) \text{ mm}$$

Abbiamo:

$$s = \frac{h}{500}$$

e, seguendo la 8:

$$\Delta s = \frac{\Delta h}{500}$$

per cui la nostra stima dello spessore del singolo foglio sarà:

$$s = (0.096 \pm 0.002) \text{ mm}$$

Non deve sorprendere che l'errore su s sia più piccolo della risoluzione del nostro strumento (la riga millimetrata) in quanto non si tratta di una misura diretta. \square

1.2.4 Quoziente

Sia Q il quoziente:

$$Q = \frac{a}{b}$$

I casi peggiori sono:

$$\begin{aligned} Q_{max} &= \frac{a_0 + \Delta a}{b_0 - \Delta b} \\ Q_{min} &= \frac{a_0 - \Delta a}{b_0 + \Delta b} \end{aligned}$$

Concentriamoci sulla prima delle due equazioni; moltiplicando numeratore e denominatore per la quantità $(b_0 + \Delta b)$ si ha:

$$Q_{max} = \frac{(a_0 + \Delta a) \cdot (b_0 + \Delta b)}{(b_0 - \Delta b) \cdot (b_0 + \Delta b)} = \frac{a_0 b_0 + a_0 \Delta b + b_0 \Delta a + \Delta a \Delta b}{b_0^2 - (\Delta b)^2}$$

Se, come prima, trascuriamo i termini del second'ordine (cioè quelli del tipo $\Delta a \Delta b$ e $(\Delta b)^2$) possiamo scrivere:

$$Q_{max} \approx \frac{a_0 b_0 + a_0 \Delta b + b_0 \Delta a}{b_0^2} = \frac{a_0}{b_0} + \frac{a_0 \Delta b + b_0 \Delta a}{b_0^2}$$

Lo stesso vale, ovviamente, per Q_{min} ; ne segue che, con la notazione consueta, possiamo scrivere:

$$\boxed{Q_0 = \frac{a_0}{b_0}} \quad (9)$$

e ancora:

$$\boxed{\Delta Q = \frac{a_0 \Delta b + b_0 \Delta a}{b_0^2}} \quad (10)$$

1.2.5 Caso generale

Iniziamo la trattazione della propagazione degli errori/incertezze nel caso generale dal caso semplice di funzioni di una sola variabile. Richiamiamo sinteticamente (senza addentrarci nei dettagli delle ipotesi necessarie alla dimostrazione) la formula di Taylor per lo sviluppo in serie di potenze di una funzione di una variabile reale x attorno ad un generico punto x_0 :

$$f(x) = f(x_0) + \sum_{k=1}^n \frac{1}{k!} \cdot \frac{d^k f}{dx^k}(x_0) \cdot (x - x_0)^k + \mathcal{O}(n+1) \quad (11)$$

in cui con il simbolo

$$\frac{d^k f}{dx^k}(x_0)$$

si intende la derivata k -esima della funzione *calcolata* nel punto x_0 ed il resto $\mathcal{O}(n+1)$ è infinitesimo di ordine $(x - x_0)^{n+1}$. Per completezza riscriviamo esplicitamente la (11) limitata al prim'ordine:

$$f(x) = f(x_0) + \frac{df}{dx}(x_0) \cdot (x - x_0) + \mathcal{O}(2) \quad (12)$$

ed al secondo:

$$f(x) = f(x_0) + \frac{df}{dx}(x_0) \cdot (x - x_0) + \frac{1}{2} \cdot \frac{d^2 f}{dx^2}(x_0) \cdot (x - x_0)^2 + \mathcal{O}(3) \quad (13)$$

Nel seguito non avremo bisogno di utilizzare la formula di Taylor oltre il second'ordine.

► ESEMPIO 1.10. Scriviamo esplicitamente lo sviluppo in serie di Taylor per la funzione

$$f(x) = \sin x$$

attorno al punto $x = 0$, fino al terzo ordine incluso. Utilizzando la (11) si ha:

$$\sin x = \sin 0 + \cos 0 \cdot x - \frac{1}{2} \cdot \sin 0 \cdot x^2 - \frac{1}{6} \cos 0 \cdot x^3 + \mathcal{O}(4) = x - \frac{x^3}{6} + \mathcal{O}(4)$$

Nel punto $x = 0.4$ rad, tanto per fare un esempio, abbiamo $\sin x \approx 0.3894$. Lo sviluppo al prim'ordine dà semplicemente $x = 0.4$ (che è già un'approssimazione ragionevole), mentre al second'ordine abbiamo $x - \frac{x^3}{6} \approx 0.3893$. In generale l'ordine dello sviluppo richiesto da una determinata applicazione dipende dal grado di approssimazione richiesta e dalla distanza dal punto attorno a cui sviluppiamo. \square

Supponiamo dunque, al solito, di misurare una certa grandezza a ottenendo un valore a_0 con un'incertezza Δa :

$$a = a_0 \pm \Delta a$$

e di essere interessati ad una grandezza $G = f(a)$. Formalmente possiamo sviluppare la funzione in serie di Taylor al prim'ordine intorno al valore misurato:

$$f(a_0 \pm \Delta a) = f(a_0) \pm \frac{df}{da}(a_0) \cdot \Delta a + \mathcal{O}(2)$$

Ammettendo che la derivata prima sia non nulla e i termini di ordine superiore al primo siano trascurabili nello sviluppo possiamo scrivere, analogamente a prima, i casi più sfavorevoli come¹:

$$\begin{aligned} G_{max} &= f(a_0) + \left| \frac{df}{da}(a_0) \right| \cdot \Delta a \\ G_{min} &= f(a_0) - \left| \frac{df}{da}(a_0) \right| \cdot \Delta a \end{aligned}$$

Seguendo la notazione introdotta nei paragrafi precedenti scriveremo allora:

$$\boxed{G_0 = f(a_0)} \quad (14)$$

e ancora:

$$\boxed{\Delta G = \left| \frac{df}{da}(a_0) \right| \cdot \Delta a} \quad (15)$$

Si tratta di un risultato fondamentale che permette di propagare l'errore (massimo) su una funzione arbitraria della variabile misurata.



Ricordiamo che la (15) è valida solo nell'ipotesi che il prim'ordine sia il termine dominante nello sviluppo in serie (cfr. l'esempio 1.12).

► ESEMPIO 1.11. Supponiamo di aver misurato un certo angolo θ ottenendo il valore:

$$\theta = \theta_0 \pm \Delta\theta = (1.38 \pm 0.01) \text{ rad}$$

e di essere interessati alla grandezza $G = \sin \theta$. Seguendo la (14) e la (15) scriveremo:

$$G_0 = \sin \theta_0 \approx 0.982$$

ed ancora:

$$\Delta G = |\cos \theta_0| \cdot \Delta\theta \approx 0.002$$

Per cui alla fine si avrà:

$$\sin \theta = 0.982 \pm 0.002$$

□

► ESEMPIO 1.12. Ripetiamo l'esercizio precedente supponendo che questa volta il valore di θ misurato e la relativa incertezza siano:

$$\theta = \theta_0 \pm \Delta\theta = (90 \pm 1)^\circ$$

Procedendo esattamente come nel caso precedente otterremmo:

$$\Delta G = |\cos \theta_0| \cdot \Delta\theta = 0$$

¹ L'uso del modulo tiene conto del fatto che la funzione può essere crescente o decrescente in un intorno di a_0 e, corrispondentemente, la derivata può essere positiva o negativa.

che è chiaramente assurdo. In questo caso la (15) non si può applicare in quanto la derivata prima della funzione si annulla nel punto misurato (facendo cadere una delle ipotesi nella derivazione della formula) e l'incertezza va calcolata esplicitamente considerando i casi più sfavorevoli come nell'esempio 1.3 oppure considerando i termini del second'ordine. \square

Possiamo generalizzare le equazioni (14) e (15) al caso in cui la funzione cui siamo interessati dipenda da più di una variabile:

$$G = f(a, b, c \dots)$$

Lo sviluppo in serie al prim'ordine si scrive adesso come:

$$f(a_0 \pm \Delta a, b_0 \pm \Delta b, c_0 \pm \Delta c \dots) = f(a_0, b_0, c_0 \dots) \pm \frac{\partial f}{\partial a}(a_0, b_0, c_0 \dots) \cdot \Delta a \pm \frac{\partial f}{\partial b}(a_0, b_0, c_0 \dots) \cdot \Delta b \pm \frac{\partial f}{\partial c}(a_0, b_0, c_0 \dots) \cdot \Delta c + \dots + \mathcal{O}(2)$$

in cui, vale la pena notarlo esplicitamente, tutte le derivate sono calcolate nel punto misurato $a_0, b_0, c_0 \dots$. Con ovvia notazione, ed analogamente al caso di una sola variabile, scriviamo dunque:

$$G_0 = f(a_0, b_0, c_0 \dots) \quad (16)$$

e ancora:

$$\Delta G = \left| \frac{\partial f}{\partial a}(a_0, b_0, c_0 \dots) \right| \cdot \Delta a + \left| \frac{\partial f}{\partial b}(a_0, b_0, c_0 \dots) \right| \cdot \Delta b + \left| \frac{\partial f}{\partial c}(a_0, b_0, c_0 \dots) \right| \cdot \Delta c \dots \quad (17)$$

in cui abbiamo, come prima, trascurato i termini di ordine superiore al primo (cioè tutti i termini del tipo $(\Delta a)^2, \Delta a \Delta b$, etc.).

► **ESEMPIO 1.13.** Torniamo all'esempio 1.3 e calcoliamo l'incertezza sul volume V del cilindro con il formalismo appena sviluppato. Dato che $V = \pi r^2 h$, le derivate parziali si scrivono come:

$$\begin{aligned} \frac{\partial V}{\partial r}(r, h) &= 2\pi r h \\ \frac{\partial V}{\partial h}(r, h) &= \pi r^2 \end{aligned}$$

da calcolarsi nei punti misurati:

$$\begin{aligned} \left| \frac{\partial V}{\partial r}(r_0, h_0) \right| &= 2\pi r_0 h_0 \\ \left| \frac{\partial V}{\partial h}(r_0, h_0) \right| &= \pi r_0^2 \end{aligned}$$

per cui scriveremo:

$$\begin{aligned} V_0 &= \pi r_0^2 h_0 \approx 7650 \text{ mm}^3 \\ \Delta V &= \pi r_0^2 \Delta h + 2\pi r_0 h_0 \Delta r \approx 180 \text{ mm}^3 \end{aligned}$$

che è esattamente il risultato ottenuto nell'esempio 1.3. \square

1.3 ERRORE RELATIVO

Quasi sempre quando si misura una generica grandezza a è fisicamente più significativo il rapporto tra l'errore e la grandezza stessa, $\frac{\Delta a}{a}$, piuttosto che l'errore assoluto Δa . Tale rapporto viene chiamato *errore relativo*.

► ESEMPIO 1.14. Se misuriamo una lunghezza l_0 di 1 m con un errore assoluto $\Delta l = 1$ mm, l'errore relativo che commettiamo è:

$$\frac{\Delta l}{l_0} = 10^{-3} = 0.1\%$$

Se adesso, a parità di errore assoluto, l_0 vale 10 mm:

$$\frac{\Delta l}{l_0} = 10^{-1} = 10\%$$

In questo secondo caso lo stesso errore assoluto di 1 mm è dunque più grave. □

1.3.1 Propagazione dell'errore relativo

La propagazione dell'errore relativo non presenta niente di concettualmente nuovo. Si tratta semplicemente di propagare l'errore assoluto e dividere per il valore della grandezza misurata.

► ESEMPIO 1.15. Torniamo al cilindro dell'esempio 1.13. Avevamo ottenuto:

$$\begin{aligned} V_0 &= \pi r_0^2 h_0 \\ \Delta V &= \pi r_0^2 \Delta h + 2\pi r_0 h_0 \Delta r \end{aligned}$$

Dividendo membro a membro si ottiene:

$$\frac{\Delta V}{V_0} = \frac{\Delta h}{h_0} + 2 \frac{\Delta r}{r_0}$$
□

L'esempio appena mostrato è in realtà solo un caso particolare di una classe vasta di problemi in cui propagare direttamente l'errore relativo è più semplice che non passare attraverso il calcolo dell'errore assoluto. Consideriamo a questo proposito una grandezza G del tipo:

$$G = a^n b^m c^p \dots$$

dove $a, b, c \dots$ al solito, sono grandezze misurate *indipendentemente* con incertezze $\Delta a, \Delta b, \Delta c \dots$ rispettivamente, mentre $m, n, p \dots$ sono numeri interi o reali, positivi o negativi. Dalla (17) si legge direttamente:

$$\Delta G = \left| n a^{n-1} b^m c^p \dots \right| \cdot \Delta a + \left| m a^n b^{m-1} c^p \dots \right| \cdot \Delta b + \left| p a^n b^m c^{p-1} \dots \right| \cdot \Delta c + \dots$$

(notiamo ancora una volta che, al solito, le derivate sono calcolate in corrispondenza dei valori misurati). Dividendo per G_0 , che vale:

$$G_0 = a_0^n b_0^m c_0^p \dots$$

otteniamo:

$$\frac{\Delta G}{G_0} = \left| n \frac{\Delta a}{a_0} \right| + \left| m \frac{\Delta b}{b_0} \right| + \left| p \frac{\Delta c}{c_0} \right| + \dots \quad (18)$$

Andiamo a *leggere* la formula: dice che gli errori relativi, quando si abbia a che fare solo con prodotti, quozienti e potenze, si sommano con un peso che è uguale al modulo dell'esponente. Da questo si deduce che è bene che le grandezze che entrano in una misura siano valutate tutte con errori relativi dello stesso ordine di grandezza. O, in altre parole, è inutile cercare di ridurre l'errore relativo su una delle grandezze nel caso in cui esso sia già molto più piccolo degli errori relativi sulle altre (questi ultimi domineranno comunque sull'errore complessivo).

1.4 CIFRE SIGNIFICATIVE E CONVENZIONI DI SCRITTURA

La precisione di un risultato sperimentale è implicita nel modo in cui il risultato è scritto. Il numero di *cifre significative* in un risultato è determinato dalle seguenti regole:

- (1) La cifra più significativa è quella più a sinistra diversa da zero.
- (2) La cifra meno significativa è quella più a destra.
- (3) Tutte le cifre comprese fra la più e la meno significativa sono cifre significative.

► **ESEMPIO 1.16.** Vediamo alcuni esempi che illustrano i punti appena elencati. Indicando con una *sottolineatura* la cifra meno significativa, e con una *sopralineatura* la cifra più significativa, si ha:

$$\begin{array}{c} \overline{3}11\underline{5} \\ \overline{3}212.\underline{5} \\ 0.0\underline{3}2\underline{5} \\ \overline{3}21.\underline{0}\underline{5} \\ 0.0\underline{3}0\underline{0} \\ \overline{3}00.\underline{0}\underline{0} \\ \overline{3}0.\underline{0}3\underline{0} \end{array}$$

Si tenga presente che, come già detto, le cifre significative sono quelle comprese fra la più e la meno significativa. □

È bene aggiungere che esiste una seconda convenzione secondo la quale, se non c'è la virgola decimale, allora la cifra meno significativa è quella più a destra diversa da zero.

► ESEMPIO 1.17. Secondo la convenzione appena menzionata, il numero 10 (così come è scritto qui) ha una sola cifra significativa; se lo volessimo con due cifre significative, dovremmo scrivere 10. per evitare l'ambiguità. □

Per sapere quante cifre significative devono essere riportate nel risultato di un esperimento bisogna valutare gli errori commessi: in generale *si riportano tutte le cifre fino alla prima influenzata dall'errore inclusa*. Questo è vero specialmente se si propagano gli errori utilizzando la nozione di errore massimo. Trattando gli errori in modo statisticamente appena più sofisticato (cfr. capitolo 4) si usa spesso scrivere la misura fino alla seconda cifra affetta da errore e, corrispondentemente, l'errore con due cifre significative.



L'incertezza relativa ad una generica misura non si scrive mai con più di due cifre significative.

Quando le cifre *non* significative vengono tagliate via da un numero, le rimanenti cifre devono essere arrotondate per una maggior accuratezza. Per arrotondare un numero lo si tronca fino ad ottenere il numero di cifre significative desiderato, poi si prendono le cifre in più e si trattano come una frazione decimale. Allora:

- (1) Se la frazione è *maggiore* di $\frac{1}{2}$ si incrementa di una unità l'ultima cifra significativa.
- (2) Se la frazione è *minore* di $\frac{1}{2}$ il numero è già pronto.
- (3) Se la frazione è *uguale* ad $\frac{1}{2}$, allora se l'ultima cifra è pari si procede come al punto 1, mentre se è dispari si procede come al punto 2.

L'ultima regola viene usata perché incrementando sempre o non facendo mai niente si incorrerebbe in un errore sistematico. Andrebbe bene anche fare l'opposto o tirare una moneta: testa incremento, croce lascio. Tuttavia il resto del mondo preferisce seguire la regola 3.

Molto spesso quando si riporta il valore di un numero molto maggiore o molto minore dell'unità di misura, si usa scrivere la grandezza in questione mediante un numero con il punto decimale dopo la prima cifra significativa, moltiplicato per un'opportuna potenza di 10.

► ESEMPIO 1.18. *Le grandezze:*

0.000015 m

677000 m

si scrivono più convenientemente come:

1.5×10^{-5} m

6.77×10^5 m

rispettivamente. Sono molto usate anche le scritture:

0.15×10^{-4} m

0.677×10^6 m

□

Quando i risultati di misure vengono usati per calcolare altre grandezze, bisogna fare attenzione agli errori che si possono compiere nei calcoli numerici.

► ESEMPIO 1.19. Consideriamo due quantità che, in qualche unità di misura, valgono:

$$\begin{aligned} a &= 1.25432 \\ b &= 9.35 \end{aligned}$$

Se esse vengono arrotondate, rispettivamente, a:

$$\begin{aligned} a_0 &= 1.25 \\ b_0 &= 9.4 \end{aligned}$$

quale errore massimo si commette nel prodotto?

Possiamo calcolare il prodotto con i numeri originali (cioè con tutte le cifre significative):

$$p = a \cdot b = 11.727892$$

e con i numeri arrotondati:

$$p_0 = a_0 \cdot b_0 = 11.750$$

La differenza tra i due prodotti vale:

$$p - p_0 = 0.022$$

L'errore è sulla quarta cifra significativa, quindi

$$p_0 = 11.75 \pm 0.02$$

oppure anche

$$p_0 = 11.750 \pm 0.022$$

Se non si riporta l'errore si scrive:

$$p_0 = 11.7$$

perché solo tre cifre sono sicure. Con questa scrittura si intende:

$$11.65 < p_0 < 11.75$$

□

► ESEMPIO 1.20. Consideriamo le due quantità:

$$\begin{aligned} a &= 3.658 \\ b &= 24.763 \end{aligned}$$

Con quante cifre significative è ragionevole scrivere il prodotto?

Il risultato dell'operazione è:

$$p = 90.583054$$

ma, per come si conoscono i numeri, il prodotto può essere un numero qualunque compreso fra i due seguenti:

$$\begin{aligned} p_{\max} &= 3.6585 \cdot 24.7635 = 90.59726457 \\ p_{\min} &= 3.6575 \cdot 24.7625 = 90.56884375 \end{aligned}$$

Allora il prodotto va scritto con 4 cifre significative; l'errore è sulla quarta cifra. □

► ESEMPIO 1.21. Consideriamo le due grandezze:

$$\begin{aligned} a &= 56.434 \\ b &= 251.37 \end{aligned}$$

Con quante cifre significative si scrive la somma $S = a + b$?

Conviene scrivere i numeri in questo modo:

$$\begin{aligned} a &= 0.056434 \times 10^3 \\ b &= 0.25137 \times 10^3 \end{aligned}$$

Facendo la somma:

$$\begin{array}{r} 0.056434 \times 10^3 \quad + \\ 0.25137 \quad \times 10^3 \quad = \\ \hline 0.307804 \times 10^3 \end{array}$$

risulta subito evidente che la sesta cifra non è significativa, come si ottiene pure dalla (2):

$$\Delta S = \Delta a + \Delta b = (5 \times 10^{-7} + 5 \times 10^{-6}) \times 10^3 \approx 5 \times 10^{-6} \times 10^3$$

□

► ESEMPIO 1.22. Consideriamo le due grandezze:

$$\begin{aligned} a &= 86.67 \\ b &= 7.0228 \end{aligned}$$

Con quante cifre significative si scrive la differenza?

Conviene scrivere i numeri in questo modo:

$$\begin{aligned} a &= 0.8667 \times 10^2 \\ b &= 0.070228 \times 10^2 \end{aligned}$$

Facendo la sottrazione:

$$\begin{array}{r} 0.8667 \quad \times 10^2 \quad - \\ 0.070228 \quad \times 10^2 \quad = \\ \hline 0.796472 \times 10^2 \end{array}$$

risulta subito evidente che la quinta e la sesta cifra non sono significative. □

► ESEMPIO 1.23. Consideriamo le due quantità:

$$a = 0.738$$

$$b = 0.736$$

La differenza è:

$$D = a - b = 0.002$$

che ha una sola cifra significativa. Cioè la differenza di due numeri quasi uguali ha una precisione molto minore di quella dei due numeri stessi. Quindi è bene fare attenzione nell'arrotondare numeri di cui poi si deve calcolare la differenza. □

Dagli esempi considerati si può trarre la seguente conclusione: se si vuole che in una operazione tra grandezze misurate gli errori di calcolo numerico siano piccoli rispetto agli errori di misura, occorre, in generale, tenere una cifra significativa in più di quelle consentite dall'errore.

Quando un'operazione coinvolge, tra le altre, grandezze che hanno a priori infinite cifre significative, il numero di cifre significative al quale si arrotondano tali grandezze deve essere tale da rendere trascurabili gli errori di calcolo numerico rispetto agli errori di misura. Vale anche la pena notare, d'altra parte, che usando un calcolatore non c'è nulla di male a prendere π con nove cifre significative (anziché quattro) e moltiplicarlo per un numero con tre cifre significative: basta tagliare le cifre non significative dal risultato.

1.5 ESERCIZI

► **ESERCIZIO 1.1.** Ricavare le formule di propagazione dell'errore massimo nelle quattro operazioni elementari (somma, differenza, prodotto, quoziente) a partire dalla formula generale (17).

► **ESERCIZIO 1.2.** Calcolare l'errore massimo sul volume del cilindro, ricavato nell'esempio 1.13 a partire dalla formula di propagazione dell'errore massimo sul prodotto (6).

► **ESERCIZIO 1.3.** Applicare la formula di propagazione dell'errore massimo sul prodotto (6) all'esempio 1.19 e discutere il risultato.

Suggerimento: si consideri, ad esempio, $a = 1.25$, $\Delta a = 0.005$, $b = 9.4$, $\Delta b = 0.005$.

► **ESERCIZIO 1.4.** Applicare la formula di propagazione dell'errore massimo sul prodotto (6) all'esempio 1.20.

Suggerimento: si consideri, ad esempio, $\Delta a = \Delta b = 5 \times 10^{-4}$.

► **ESERCIZIO 1.5.** Si considerino le due grandezze:

$$a = 10.3484$$

$$b = 6.24$$

Con quante cifre significative deve essere preso il quoziente? Verificare che si ottiene lo stesso risultato dalla propagazione dell'errore massimo (10).

► **ESERCIZIO 1.6.** Nel prodotto tra le due grandezze:

$$a = 1.25432$$

$$b = 9.3$$

a quante cifre significative conviene arrotondare a ?

Suggerimento: eseguire i prodotti arrotondando successivamente a 5, 4, 3, 2 ... cifre significative.

► **ESERCIZIO 1.7.** Una certa attenzione va fatta quando i prodotti sono più di due. Consideriamo ad esempio il caso:

$$p = 1.234 \times 56.789 \times 11.9955 \times 5.1248$$

Mostrare che potrebbe essere dannoso arrotondare prima della moltiplicazione ciascun fattore a quattro cifre. Conviene partire dai fattori con più cifre significative o da quelli con meno?

► **ESERCIZIO 1.8.** A quante cifre significative va arrotondato $\pi = 3.14159265$ affinché le potenze π^2 , π^3 , π^4 , π^5 , π^6 siano corrette al livello dell'1%? E allo 0.1%?

► **ESERCIZIO 1.9.** Consideriamo la grandezza:

$$G = \frac{\sqrt{2} + \pi^2}{\sqrt{10} - \pi}$$

Con quante cifre significative dobbiamo prendere $\sqrt{2}$, $\sqrt{10}$ e π se vogliamo che l'errore relativo su G sia 1%?

► ESERCIZIO 1.10. Ripetere l'esercizio precedente nel caso in cui l'errore relativo richiesto sia lo 0.1%

PROBABILITÀ E STATISTICA

INDICE

2.1	Definizione di probabilità	26
2.1.1	Prima definizione (Laplace, 1812)	26
2.1.2	Seconda definizione (Venn, 1866)	27
2.1.3	Terza definizione (Kolmogorov, 1933)	27
2.2	Proprietà e leggi della probabilità	28
2.2.1	Addizione	28
2.2.2	Moltiplicazione	29
2.3	Variabili casuali e funzioni di distribuzione	32
2.3.1	Variabili casuali discrete	33
2.3.2	Variabili casuali continue	33
2.4	Rappresentazione di fenomeni casuali	35
2.5	Caratteristiche comuni alle distribuzioni	40
2.5.1	Premessa	40
2.5.2	Notazioni	41
2.5.3	Definizioni di base: normalizzazione e valor medio	41
2.5.4	Valore di aspettazione o di previsione	46
2.5.5	Dispersione intorno alla media: varianza e deviazione standard	46
2.5.6	La semilarghezza a metà altezza	49
2.5.7	Momenti di una distribuzione	51
2.5.8	Discussione e cenno al Teorema di Tschebyscheff	52
2.5.9	Cenno alla teoria dei campioni	53
2.6	Esercizi	55

Questo capitolo raccoglie alcuni fatti fondamentali relativi al calcolo elementare delle probabilità ed alla rappresentazione di fenomeni casuali. Si introduce il concetto di funzione di distribuzione (per variabili casuali discrete e continue) e si definiscono le caratteristiche fondamentali (media, varianza e deviazione standard) comuni a tutte le distribuzioni.

2.1 DEFINIZIONE DI PROBABILITÀ

2.1.1 *Prima definizione (Laplace, 1812)*

La probabilità di un evento E , che si indica solitamente con $P(E)$, è uguale al rapporto tra il numero dei casi favorevoli e quello dei casi possibili:

$$P(E) = \frac{\text{casi favorevoli}}{\text{casi possibili}} \quad (19)$$

Questa viene anche chiamata *probabilità a priori* di un evento. È evidente, dalla (19), che:

$$P(E) \leq 1 \quad (20)$$

Questa definizione, limitata al campo dei numeri razionali ed applicabile solo in situazioni in cui il numero di casi possibili è finito, è sufficiente per i casi più elementari.

► **ESEMPIO 2.1.** *Qual è la probabilità che, nel lancio di una moneta, esca testa (o croce)? Indichiamo con T e C l'uscita di testa o croce, rispettivamente. Il numero di casi possibili è due, mentre il numero di casi favorevoli è uno, per cui:*

$$P(T) = P(C) = \frac{1}{2}$$

□

► **ESEMPIO 2.2.** *Qual è la probabilità che, nel lancio di un dado, esca un numero fissato (ad esempio il 3)?*

Vi è un solo caso favorevole su sei possibili, per cui:

$$P(3) = \frac{1}{6}$$

□

► **ESEMPIO 2.3.** *Qual è la probabilità che, lanciando due monete, esca almeno una testa? Indichiamo con T l'uscita di una testa e con C l'uscita di una croce. Le realizzazioni elementari del nostro esperimento si possono identificare univocamente con coppie ordinate (u_1, u_2) dove u_1 indica l'uscita della prima moneta e u_2 quella della seconda. Nel nostro caso avremo esattamente 4 casi possibili: (T, T) , (T, C) , (C, T) e (C, C) . Tra di essi quelli favorevoli sono 3: (T, T) , (T, C) e (C, T) ; per cui:*

$$P(\text{almeno una testa}) = \frac{3}{4}$$

□

Non appena la situazione si complica diventa necessario tenere conto che alcuni casi possono essere più o meno *favorevoli* di altri. Dobbiamo quindi modificare la nostra definizione aggiungendo esplicitamente la richiesta che i casi siano *ugualmente possibili*. Vedremo tra breve una definizione alternativa e più generale.

► ESEMPIO 2.4. Supponiamo di lanciare due dadi a sei facce; qual è la probabilità che la somma delle uscite sia 4? Se applicassimo acriticamente la 19, essendo i casi possibili esattamente 11 (la somma delle uscite di due dadi può assumere tutti i valori interi da 2 a 12) dovremmo concludere che $P(4) = \frac{1}{11}$ che è falso. Vedremo tra breve come calcolare correttamente questa probabilità; per inciso anticipiamo che $P(4) = \frac{1}{12}$. \square

2.1.2 Seconda definizione (Venn, 1866)

Si tratta di una definizione operativa. La probabilità di un evento E coincide con la frequenza relativa dell'evento stesso quando il numero totale N di volte in cui si ripete l'esperimento è molto grande:

$$P(E) = \lim_{N \rightarrow \infty} \frac{n}{N} \quad (21)$$

in cui n è il numero di volte in cui si registra l'evento E . Poiché $n \leq N$ vale ovviamente la (20).

L'equazione (21) non ha chiaramente il significato che i matematici attribuiscono di solito al concetto di limite. Essa significa sostanzialmente che, compiendo una serie di prove, al crescere di N il rapporto $\frac{n}{N}$ tende a stabilizzarsi intorno ad un certo valore, con fluttuazioni che tendono a decrescere sempre di più.

Questa definizione è utile per le applicazioni pratiche ma non per una rigorosa costruzione matematica della teoria della probabilità.

2.1.3 Terza definizione (Kolmogorov, 1933)

Si tratta di una definizione assiomatica fondata sulla teoria degli insiemi. Sia E un generico evento elementare associato ad un certo esperimento e chiamiamo \mathcal{S} l'insieme di tutti gli eventi elementari E . Si consideri adesso un generico sottoinsieme s di \mathcal{S} :

$$s \subseteq \mathcal{S}$$

e sia \mathcal{E} l'insieme di tutti i sottoinsiemi di \mathcal{S} :

$$\mathcal{E} = \{s : s \subseteq \mathcal{S}\}$$

Notiamo esplicitamente che l'insieme vuoto \emptyset e l'insieme \mathcal{S} appartengono ad \mathcal{E} :

$$\begin{aligned} \emptyset &\in \mathcal{E} \\ \mathcal{S} &\in \mathcal{E} \end{aligned}$$

Notiamo altresì che, dati due qualsiasi insiemi \mathcal{A} e \mathcal{B} appartenenti ad \mathcal{E} , anche la loro unione, la loro intersezione ed i loro complementari in \mathcal{E} appartengono ad \mathcal{E} . In formule:

$$\begin{aligned} \mathcal{A} \cup \mathcal{B} &\in \mathcal{E} & \forall \mathcal{A}, \mathcal{B} \in \mathcal{E} \\ \mathcal{A} \cap \mathcal{B} &\in \mathcal{E} & \forall \mathcal{A}, \mathcal{B} \in \mathcal{E} \\ \overline{\mathcal{A}} &\in \mathcal{E} & \forall \mathcal{A} \in \mathcal{E} \end{aligned}$$

Sia dunque $P(s)$ una funzione definita nell'insieme \mathcal{E} che associa ad ogni sottoinsieme s di S un numero reale non negativo. Diciamo che $P(s)$ è una probabilità se valgono le seguenti proprietà:

1. $P(S) = 1$
2. $P(A \cup B) = P(A) + P(B) - P(A \cap B) \quad \forall A, B \in \mathcal{E}$
3. $P(\emptyset) = 0$

Da queste tre proprietà deriva di nuovo la (20).

2.2 PROPRIETÀ E LEGGI DELLA PROBABILITÀ

2.2.1 Addizione

Se A e B sono due eventi *disgiunti* (cioè *non* possono verificarsi contemporaneamente nello stesso esperimento o, equivalentemente, i sottoinsiemi che li rappresentano hanno intersezione vuota), la probabilità che si verifichi l'evento A oppure l'evento B (in altre parole l'unione degli eventi A e B) è:

$$P(A \circ B) = P(A) + P(B) \quad (22)$$

► **ESEMPIO 2.5.** Si lanci un dado a sei facce. Qual è la probabilità che esca il 2 oppure il 3? Con il consueto significato dei simboli:

$$P(2 \circ 3) = P(2) + P(3) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}$$

□

Se A e B sono *compatibili*, cioè non sono disgiunti, possono in generale presentarsi dei casi favorevoli ad entrambi gli eventi, che vengono dunque contati sia nel calcolo di $P(A)$ che in quello di $P(B)$. In questo caso se applicassimo la (22), questi casi verrebbero contati due volte; ne segue che la legge dell'addizione delle probabilità deve essere così generalizzata:

$$P(A \circ B) = P(A) + P(B) - P(A \text{ e } B) \quad (23)$$

► **ESEMPIO 2.6.** Supponiamo di estrarre una carta da un mazzo di 52 carte e consideriamo i seguenti eventi possibili:

A = la carta è di fiori

B = la carta è un re

Le probabilità che ciascuno di questi due eventi si realizzi si calcola facilmente come rapporto tra casi favorevoli e casi possibili:

$$\begin{aligned}P(A) &= \frac{13}{52} \\P(B) &= \frac{4}{52}\end{aligned}$$

Vogliamo calcolare la probabilità dell'evento C che la carta estratta sia una carta di fiori oppure un re:

C = la carta è di fiori oppure è un re

Se applicassimo la (22) otterremmo $P(C) = P(A) + P(B) = \frac{17}{52}$, che è errato. I due eventi A e B , infatti, non sono disgiunti (il re di fiori è contemporaneamente una carta di fiori ed un re), per cui:

$$P(C) = P(A) + P(B) - P(A \text{ e } B) = \frac{13}{52} + \frac{4}{52} - \frac{1}{52} = \frac{16}{52}$$

che è il risultato corretto, come è immediato verificare. □

2.2.2 Moltiplicazione

Se A e B sono due eventi *indipendenti* (cioè non causalmente connessi, nel senso che il fatto che si verifichi A non influenza in alcun modo il fatto che B si verifichi, e viceversa) la probabilità che sia A che B si verifichino è data da:

$$P(A \text{ e } B) = P(A) \cdot P(B) \quad (24)$$

Infatti se indichiamo con n_A ed n_B i casi favorevoli degli eventi A e B e con N_A e N_B i corrispondenti casi possibili, si ha:

$$\begin{aligned}P(A) &= \frac{n_A}{N_A} \\P(B) &= \frac{n_B}{N_B}\end{aligned}$$

I casi possibili dell'evento $C = (A \text{ e } B)$ sono $N_A \cdot N_B$, mentre i casi favorevoli sono $n_A \cdot n_B$, per cui:

$$P(A \text{ e } B) = \frac{n_A \cdot n_B}{N_A \cdot N_B} = P(A) \cdot P(B)$$

► **ESEMPIO 2.7.** Si lancino un dado ed una moneta. Qual è la probabilità che escano il 2 e la testa?

Seguendo la (24) scriviamo:

$$P(2 \text{ e testa}) = \frac{1}{6} \cdot \frac{1}{2} = \frac{1}{12}$$

□

► **ESEMPIO 2.8.** Supponiamo di lanciare due dadi a sei facce. Qual è la probabilità che il primo dado dia 2 ed il secondo 3? Identifichiamo i nostri eventi elementari:

$A = \text{il primo dado dà } 2$
 $B = \text{il secondo dado dà } 3$

$$P(A \text{ e } B) = P(A) \cdot P(B) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$$

□

► **ESEMPIO 2.9.** Supponiamo di estrarre una carta da un mazzo di 52 carte e consideriamo di nuovo gli eventi:

$A = \text{la carta è di fiori}$
 $B = \text{la carta è un re}$

Le corrispondenti probabilità, come già detto nell'esempio 2.6 sono $P(A) = \frac{13}{52}$ e $P(B) = \frac{4}{52}$. Questa volta vogliamo calcolare la probabilità dell'evento C che la carta estratta sia il re di fiori:

$C = \text{la carta è il re di fiori}$

Essendo gli eventi indipendenti possiamo applicare la (24):

$$P(C) = P(A) \cdot P(B) = \frac{13}{52} \cdot \frac{4}{52} = \frac{1}{52}$$

che è il risultato corretto. Notiamo esplicitamente che l'esempio non è così banale come può apparire a prima vista; vi sono casi in cui non è affatto ovvio stabilire se due eventi siano indipendenti o meno (cfr. in particolare l'esempio 2.11). □

Quando gli eventi *non* sono indipendenti la (24) non vale più. I due esempi che seguono chiariscono questa affermazione.

► **ESEMPIO 2.10.** Supponiamo di lanciare un dado. Qual è la probabilità che l'uscita sia 2 e 3? Gli eventi elementari sono:

$A = \text{l'uscita è } 2$
 $B = \text{l'uscita è } 3$

Se applicassimo di nuovo la (24) otterremmo, esattamente come nell'esempio 2.8, una probabilità pari ad $\frac{1}{36}$, il che è chiaramente assurdo, in quanto il 2 ed il 3 non possono uscire contemporaneamente. In particolare la probabilità cercata è nulla.

In effetti i due eventi sono incompatibili e, di conseguenza, non sono indipendenti, per cui la (24) non vale in questo caso. □

► **ESEMPIO 2.11.** Un caso meno ovvio. Consideriamo di nuovo l'esempio 2.9 ma questa volta aggiungiamo un jolly al mazzo di 52 carte. Sia di nuovo:

A = la carta è di fiori
 B = la carta è un re
 C = la carta è il re di fiori

Adesso il mazzo è di 53 carte per cui le probabilità degli eventi A e B divengono:

$$\begin{aligned}
 P(A) &= \frac{13}{53} \\
 P(B) &= \frac{4}{53}
 \end{aligned}$$

Se applicassimo di nuovo la (24) otterremmo

$$P(C) = P(A) \cdot P(B) = \frac{13}{53} \cdot \frac{4}{53} = \frac{212}{2809}$$

mentre sappiamo che la probabilità di estrarre il re di fiori è esattamente $\frac{1}{53}$.

In questo caso non vale la legge della moltiplicazione per eventi indipendenti, in quanto gli eventi non sono più indipendenti: se esce una carta di fiori questo esclude che sia un jolly, quindi modifica la probabilità che sia un re.

Impareremo tra breve come la (24) possa essere generalizzata al caso di eventi non indipendenti; nel frattempo l'esempio merita un commento un poco più articolato, dato che la questione dell'indipendenza tra eventi tende talvolta ad essere sfuggente. Supponiamo, nelle ipotesi dell'esempio, di estrarre una carta, e supponiamo ancora che questa carta sia di fiori. La probabilità che essa sia un re è esattamente $\frac{1}{13}$, dato che vi sono 13 carte di fiori e, tra queste, solo una è un re. Supponiamo viceversa che la carta non sia di fiori. In questo caso la probabilità che essa sia un re è $\frac{3}{40}$ (vi sono esattamente $53 - 13 = 40$ carte che non sono di fiori e, tra queste, solo 3 sono re). Dunque sapere se la carta sia di fiori o meno influenza la probabilità che essa sia un re; che è come dire che i due eventi non sono indipendenti.

Se ripetiamo lo stesso esercizio nel caso dell'esempio 2.9 le cose vanno diversamente: se estraiamo una carta di fiori la probabilità che essa sia un re è $\frac{1}{13}$, e se estraiamo una carta non di fiori la probabilità che essa sia un re è $\frac{3}{39}$ (vi sono esattamente $52 - 13 = 39$ carte che non sono di fiori e, tra queste, solo 3 sono re) cioè ancora $\frac{1}{13}$. In quel caso possiamo affermare che gli eventi sono effettivamente indipendenti. \square

La legge della moltiplicazione può essere generalizzata al caso di eventi non indipendenti. Definiamo la *probabilità dell'evento B condizionata dall'evento A* , e la indichiamo con $P(B|A)$, come la probabilità che si verifichi B una volta che si è verificato A . La nozione di probabilità condizionata permette di caratterizzare in modo semplice l'indipendenza tra eventi. Se A e B sono indipendenti vale infatti:

$$P(B|A) = P(B|\bar{A}) = P(B) \quad (25)$$

e viceversa gli eventi A e B sono indipendenti se la (25) è verificata. Nel nostro nuovo linguaggio la legge di moltiplicazione delle probabilità diviene:

$$P(A \text{ e } B) = P(A) \cdot P(B|A) \quad (26)$$

Alla luce di questa relazione possiamo riesaminare criticamente gli esempi 2.10 e 2.11.

► ESEMPIO 2.12. *Riesaminiamo l'esempio 2.10. La probabilità che esca il 3, una volta che è uscito il 2, è nulla poiché i due eventi sono incompatibili:*

$$P(B|A) = 0$$

per cui:

$$P(C) = P(A) \cdot P(B|A) = \frac{1}{6} \cdot 0$$

□

► ESEMPIO 2.13. *Torniamo per l'ultima volta sull'esempio 2.11 e consideriamo di nuovo i nostri tre eventi elementari:*

A = la carta è di fiori

B = la carta è un re

C = la carta è il re di fiori

È facile convincersi (lo abbiamo sostanzialmente già dimostrato) che:

$$P(B|A) = \frac{1}{13}$$

$$P(B|\bar{A}) = \frac{3}{40}$$

per cui gli eventi non sono indipendenti. Scriveremo allora:

$$P(C) = P(A) \cdot P(B|A) = \frac{13}{53} \cdot \frac{1}{13} = \frac{1}{53}$$

che è il risultato corretto. Nel caso dell'esempio 2.9 si ha viceversa:

$$P(B|A) = \frac{1}{13}$$

$$P(B|\bar{A}) = \frac{3}{39} = \frac{1}{13}$$

per cui possiamo usare indifferentemente la (24) o la (26).

□

2.3 VARIABILI CASUALI E FUNZIONI DI DISTRIBUZIONE

Sia x il risultato di un determinato esperimento; x dipende in generale dalla particolare realizzazione dell'esperimento stesso, per cui è una variabile casuale. Esistono casi di variabili casuali *discrete*, cioè variabili che possono assumere un numero finito (o al massimo numerabile) di valori in un intervallo finito, e casi di variabili *continue* su intervalli finiti o infiniti.

► ESEMPIO 2.14. *Si consideri il lancio di un dado a sei facce. L'uscita del lancio è una variabile casuale discreta che può assumere esattamente sei valori.* □

► ESEMPIO 2.15. *Supponiamo di sparare con una carabina su di un bersaglio: la distanza dal centro a cui arriva il colpo è una variabile casuale continua su di un intervallo finito.* □

2.3.1 Variabili casuali discrete

Consideriamo dapprima il caso di variabile discreta e supponiamo che $x_1 \dots x_n$ siano gli n possibili esiti; indichiamo con $P(x_i)$ la probabilità che l'esito dell'esperimento in questione sia proprio x_i . Definiamo *funzione di distribuzione* della variabile x l'insieme dei valori di $P(x_i)$.

► **ESEMPIO 2.16.** Sia la variabile casuale x l'uscita di un dado a sei facce; essa può assumere sei valori x_i , che coincidono con i numeri interi tra 1 e 6. Se il dado è equo ciascuna delle facce ha la stessa probabilità $\frac{1}{6}$ di uscire e la funzione di distribuzione è:

x_i	1	2	3	4	5	6
$P(x_i)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

□

► **ESEMPIO 2.17.** Consideriamo adesso il lancio di due dadi e sia x la somma delle uscite dei dadi stessi; x può assumere tutti i valori interi tra 2 e 12, ma questa volta gli x_i non sono equiprobabili anche se i dadi sono equi. I casi possibili sono in totale $6 \cdot 6 = 36$ ed ognuno può essere univocamente identificato con la coppia (u_1, u_2) , in cui u_1 è l'uscita del primo dado e u_2 quella del secondo. Mentre il 2 ha solamente un modo per realizzarsi, $(1, 1)$, il 4, tanto per fare un esempio, può realizzarsi in tre modi diversi: $(1, 3)$, $(2, 2)$ o $(3, 1)$. La probabilità di uscita può allora essere calcolata come rapporto tra il numero di casi favorevoli ed il numero totale di casi possibili:

x_i	2	3	4	5	6	7	8	9	10	11	12
$P(x_i)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

□

2.3.2 Variabili casuali continue

Mentre nel caso discreto la funzione di distribuzione è una funzione che associa ad ogni valore della variabile x la sua probabilità, nel caso di una variabile continua questa definizione non è più applicabile poiché la probabilità che la variabile x assuma un valore esattamente definito è zero (vedremo tra poco che si tratta essenzialmente di un integrale su un dominio di misura nulla). È invece sensato chiedersi quale sia la probabilità che la x assuma un valore compreso in un certo intervallo assegnato. Dividiamo allora l'intervallo in cui può variare la x in n intervallini di ampiezza Δx centrati attorno ai punti x_i , con $i = 1 \dots n$. Possiamo calcolare la probabilità che la variabile vada a cadere in ciascun intervallino, cioè la probabilità che essa assuma un valore compreso fra $x_i - \frac{\Delta x}{2}$ e $x_i + \frac{\Delta x}{2}$:

$$P(x_i, \Delta x) \equiv P\left(x_i - \frac{\Delta x}{2} \leq x \leq x_i + \frac{\Delta x}{2}\right)$$

Per avere una grandezza indipendente dalla larghezza dell'intervallo si può calcolare il rapporto tra $P(x_i, \Delta x)$ e Δx . Quello che otteniamo è qualcosa di simile ad una probabilità specifica o probabilità per unità di intervallo. Se consideriamo il limite:

$$\lim_{\Delta x \rightarrow 0} \frac{P(x_i, \Delta x)}{\Delta x}$$

otteniamo la *densità di probabilità* $p(x)$ che è per definizione la funzione di distribuzione della variabile casuale x . La probabilità che la nostra variabile casuale assuma un valore compreso nell'intervallo *infinitesimo* dx centrato attorno ad un generico valore x si scrive pertanto:

$$P(x, dx) = p(x) dx$$

e, corrispondentemente, la probabilità che essa assuma un valore compreso nell'intervallo *finito* $[a, b]$ è data da:

$$P(a \leq x \leq b) = \int_a^b p(x) dx \quad (27)$$

► **ESEMPIO 2.18.** Consideriamo la funzione di distribuzione:

$$p(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & x < 0; x > 1 \end{cases}$$

Ci si chiede quale sia la probabilità che x sia maggiore di 0.5.

Seguendo la (27), scriveremo:

$$P(0.5 \leq x \leq 1) = \int_{0.5}^1 1 dx = [x]_{0.5}^1 = 1 - 0.5 = 0.5$$

□

► **ESEMPIO 2.19.** Supponiamo di dover prendere un autobus. Sappiamo che passa da una certa fermata ogni 10 minuti ma non sappiamo a che ora. Qual è la probabilità di dover aspettare per più di 3 minuti, se ci presentiamo a caso alla fermata? L'istante in cui ci rechiamo alla fermata è una variabile casuale del tutto scorrelata dall'istante di arrivo dell'autobus, per cui il tempo di attesa sarà pure una variabile casuale la cui funzione di distribuzione sarà uniforme tra un valore minimo (0 minuti, nel caso in cui arriviamo esattamente insieme all'autobus) ed un valore massimo (10 minuti, nel caso in cui arriviamo appena dopo la partenza dell'autobus stesso). Possiamo scrivere esplicitamente questa funzione di distribuzione come:

$$p(x) = \begin{cases} \frac{1}{10} & 0 \leq x \leq 10 \\ 0 & x < 0; x > 10 \end{cases}$$

in cui il valore $\frac{1}{10}$, come vedremo in dettaglio nella sezione 2.5.3, è conseguenza diretta di quella che va sotto il nome di condizione di normalizzazione. La probabilità cercata, dunque, si scrive come:

$$P(x \geq 3) = \int_3^{10} \frac{1}{10} dx = \left[\frac{x}{10} \right]_3^{10} = 1 - \frac{3}{10} = 0.7$$

□

2.4 RAPPRESENTAZIONE DI FENOMENI CASUALI

In questo paragrafo ci occupiamo del problema fondamentale della rappresentazione dei dati sperimentali raccolti nello studio di un certo fenomeno casuale.

Se la grandezza che stiamo misurando può assumere un numero finito di valori (o meglio se la variabile casuale che stiamo studiando è discreta) allora il modo più immediato di rappresentare i dati raccolti è quello di creare una tabella contenente, per ognuno dei valori possibili di x , il numero di volte (numero di *occorrenze* o *frequenza*) in cui abbiamo registrato il valore in questione.

► **ESEMPIO 2.20.** *Supponiamo di lanciare per 200 volte due dadi e di registrare, per ogni lancio, la somma delle due uscite, che chiamiamo x . La seguente tabella contiene le occorrenze misurate dei possibili valori di x ; come è ovvio, la somma degli $n(x_i)$ è proprio 200.*

x_i	2	3	4	5	6	7	8	9	10	11	12
$n(x_i)$	7	5	18	18	31	36	29	24	12	12	8

□

Una rappresentazione alternativa, e per certi versi più *espressiva*, dei dati è costituita da un grafico (cfr. figura 1) in cui, in corrispondenza di ognuno dei valori possibili della variabile casuale si traccia un segmento verticale di lunghezza proporzionale al numero di volte in cui si è registrato il valore in questione, ossia al numero di *occorrenze* $n(x_i)$ di tale valore.

In una rappresentazione di questo tipo, che si chiama generalmente *rappresentazione di frequenza*, sull'asse delle ordinate si leggono direttamente le occorrenze (registrate nelle N volte in cui si ripete l'esperimento) di ognuno dei valori che la variabile casuale può assumere. È evidente che in questa rappresentazione la somma dei valori sull'asse delle ordinate è uguale proprio ad N .

Talvolta può essere più utile rappresentare le frequenze relative, cioè i rapporti tra le singole occorrenze ed il numero totale di volte in cui si ripete l'esperimento:

$$f(x_i) = \frac{n(x_i)}{N}$$

Uno dei vantaggi di questo approccio è che, dato che la frequenza relativa di un evento tende alla probabilità dell'evento stesso al tendere di N all'infinito, come si legge nella (21), la tabella delle frequenze relative può essere confrontata direttamente con la funzione di distribuzione.

► **ESEMPIO 2.21.** *Torniamo ai dati dell'esempio 2.20. In questo caso la tabella delle frequenze relative si scrive come:*

x_i	2	3	4	5	6	7	8	9	10	11	12
$f(x_i)$	0.035	0.025	0.09	0.09	0.155	0.18	0.145	0.12	0.06	0.06	0.004

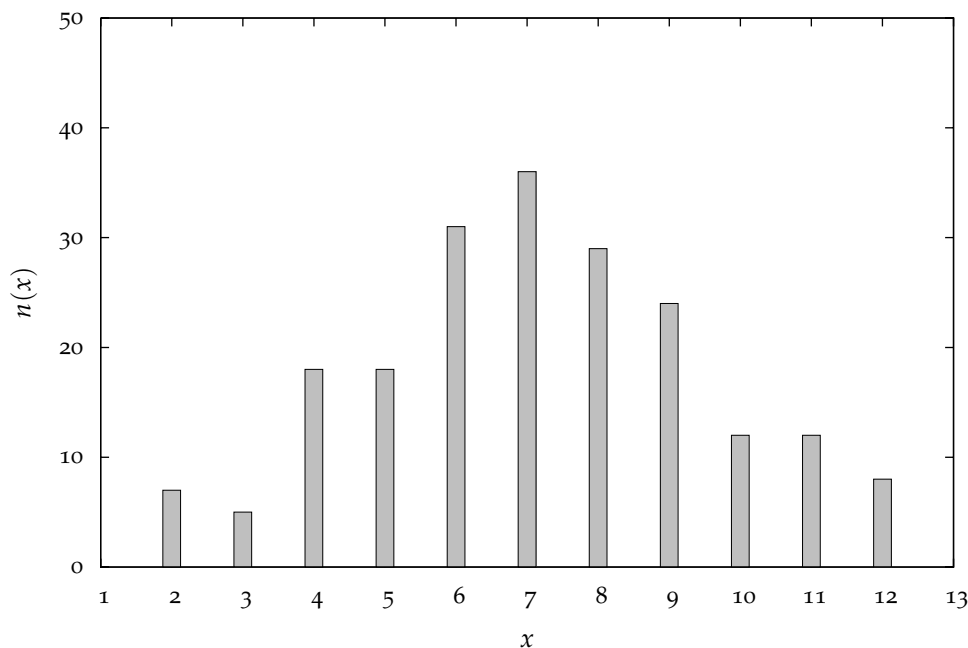


FIGURA 1.. Rappresentazione di frequenza della somma delle uscite di due dadi, relativa ai dati dell'esempio 2.20. Sull'asse delle ascisse si ha la variabile casuale x e, in corrispondenza di ognuno dei valori possibili x_i (che in questo caso sono tutti gli interi da 2 a 12), si traccia un segmento verticale la cui lunghezza è proporzionale al numero di volte in cui si è ottenuto il valore in questione.

e la somma dei valori degli $f(x_i)$ è adesso 1 anziché 200. □

In figura 2 si ha la *rappresentazione di frequenza relativa* riferita all'esempio 2.20. Uno degli svantaggi della rappresentazione di frequenza relativa è costituito dal fatto che l'informazione su N è completamente perduta. Questi due diversi approcci sono dunque entrambi utili e l'uno o l'altro deve essere privilegiato a seconda del particolare aspetto che si vuole mettere in evidenza.

► **ESEMPIO 2.22.** *Supponiamo di avere un campione di 70 chiodini e di aver misurato con un calibro centesimale la lunghezza di ciascuno di essi ottenendo la seguente tabella¹.*

LUNGHEZZA (cm)									
1.731	1.739	1.698	1.700	1.697	1.728	1.696	1.714	1.697	1.697
1.724	1.729	1.746	1.703	1.696	1.679	1.756	1.689	1.743	1.747
1.742	1.723	1.712	1.729	1.739	1.681	1.657	1.663	1.694	1.693
1.737	1.717	1.736	1.709	1.705	1.724	1.706	1.699	1.701	1.708
1.712	1.731	1.729	1.703	1.698	1.661	1.685	1.695	1.703	1.740
1.705	1.730	1.723	1.716	1.711	1.722	1.724	1.729	1.695	1.721
1.707	1.713	1.722	1.695	1.649	1.706	1.726	1.732	1.703	1.707

¹ Dati realmente ottenuti il 6 Agosto 2005.

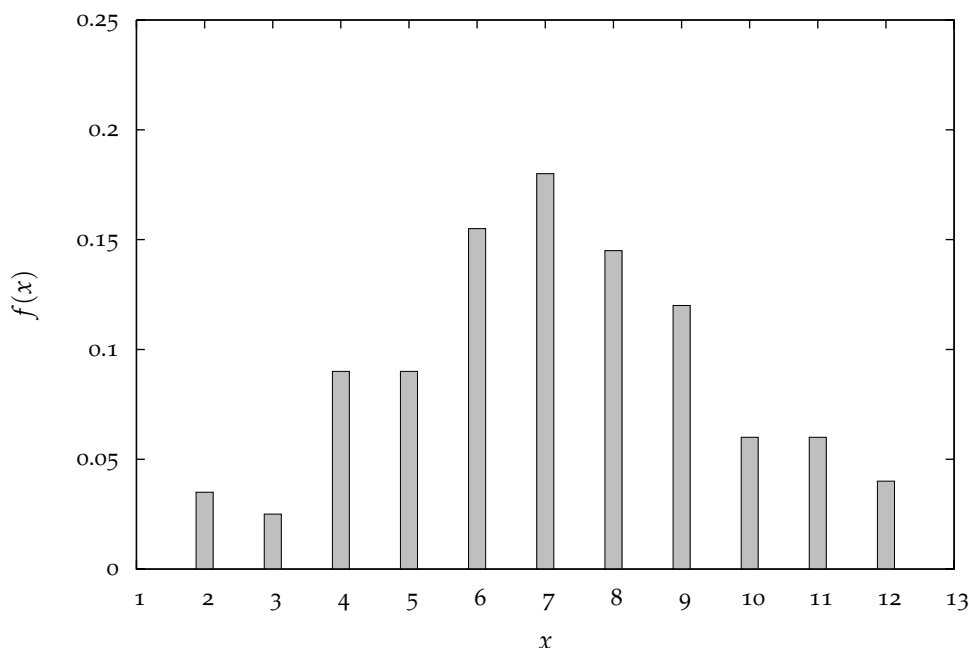


FIGURA 2.. Rappresentazione di frequenza relativa della somma delle uscite di due dadi (i dati sono quelli dell'esempio 2.20).

È facile verificare che le misure assumono valori compresi tra 1.649 cm e 1.756 cm; data la risoluzione dello strumento, 0.001 cm, il numero di valori che la nostra variabile può, in linea di principio, assumere, è

$$\frac{(1.756 - 1.649)}{0.001} = 107$$

Non c'è dubbio, allora, che una tabella con 107 colonne, analoga a quella dell'esempio 2.20 in questo caso sarebbe poco leggibile. □

Questo tipo di rappresentazione è molto utile per variabili che possono assumere valori discreti ed in numero non molto grande. Quando queste condizioni non sono soddisfatte si usa tipicamente un altro tipo di rappresentazione detto ad *istogramma*. Nel caso di variabili continue l'istogramma è sostanzialmente l'unica rappresentazione possibile. Un istogramma è di nuovo un grafico in cui si pone sull'asse delle ascisse la variabile casuale x da rappresentare. L'intervallo di variabilità di x (cioè, operativamente, l'intervallo entro cui si dispongono le misure di x effettuate) viene suddiviso in un certo numero di intervallini di uguale ampiezza (che si chiamano anche *canali* o *bin* dell'istogramma) ed ognuno di questi è assunto come base di un rettangolo di area *proporzionale* al numero di misure che cadono nel canale stesso². In base alla costruzione appena descritta l'area al di sotto dell'istogramma è pari al numero N delle

² Si osservi che, in un istogramma, le grandezze riportate sull'asse delle ordinate devono avere le dimensioni fisiche dell'inverso della variabile x , di modo che l'area di ogni rettangolo abbia le dimensioni di un numero puro.

prove effettuate: si dice che l'istogramma è *normalizzato* al numero N delle misure o semplicemente normalizzato ad N .

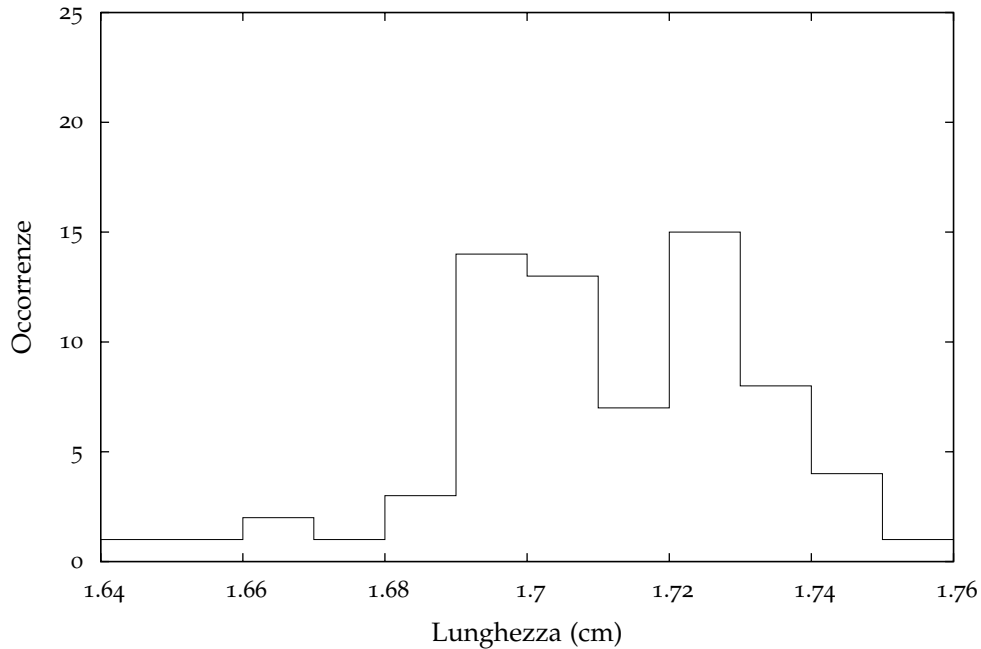


FIGURA 3.. Iistogramma normalizzato ad N relativo ai dati dell'esempio 2.22. La larghezza di ogni canale dell'istogramma è fissata a 0.01 cm (pari a 10 volte la risoluzione dello strumento).

Anche in questo caso si può riportare in ogni canale dell'istogramma la frazione (anziché il numero) di prove in cui il risultato cade entro il canale stesso; operativamente è sufficiente dividere l'altezza di ogni rettangolo per N . Così facendo l'area al di sotto dell'istogramma diviene uguale ad 1 e si dice che l'istogramma è normalizzato ad 1.

Benché si perda completamente l'informazione relativa alla distribuzione dei dati dentro ciascun intervallo, questa rappresentazione è molto utile per una prima analisi di un fenomeno. In generale si deve cercare un compromesso tra la quantità totale dell'informazione e la sua chiarezza: non conviene fare gli intervalli troppo piccoli perché altrimenti si mettono in evidenza le fluttuazioni statistiche; d'altra parte, se essi sono troppo grandi, al limite non si ha più informazione.

Per determinare l'ampiezza degli intervalli risulta particolarmente conveniente usare la seguente regola empirica che troverà negli argomenti trattati in seguito la sua giustificazione: l'ampiezza degli intervallini deve essere circa uguale all'intervallo di variabilità della variabile casuale diviso per la radice quadrata del numero di prove effettuate:

$$\Delta x \approx \frac{x_{\max} - x_{\min}}{\sqrt{N}} \quad (28)$$

Tale quantità risulta essere una stima grossolana ma veloce della deviazione standard della distribuzione [8, 9].

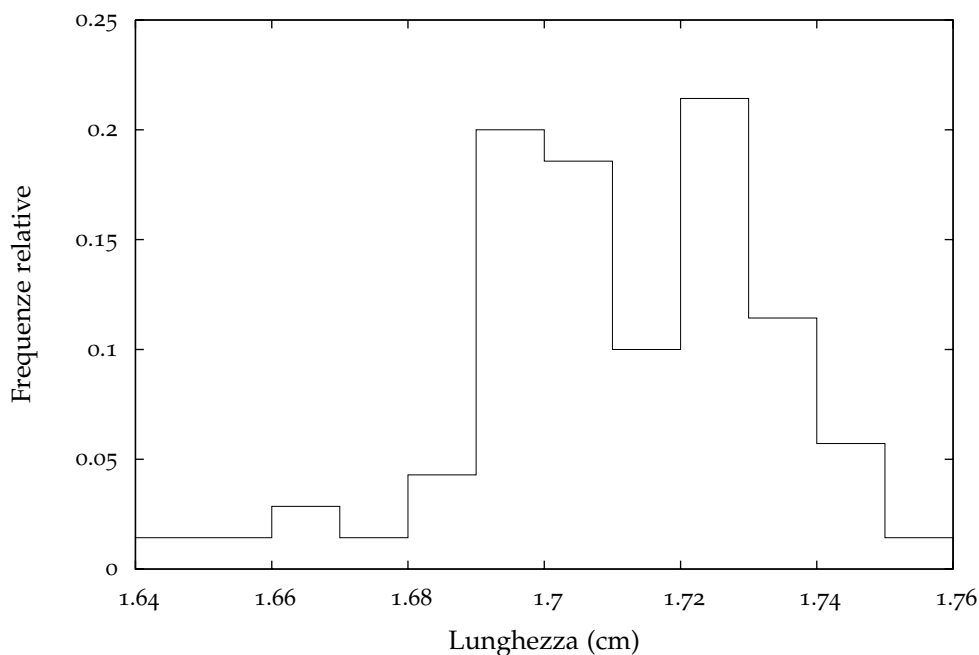
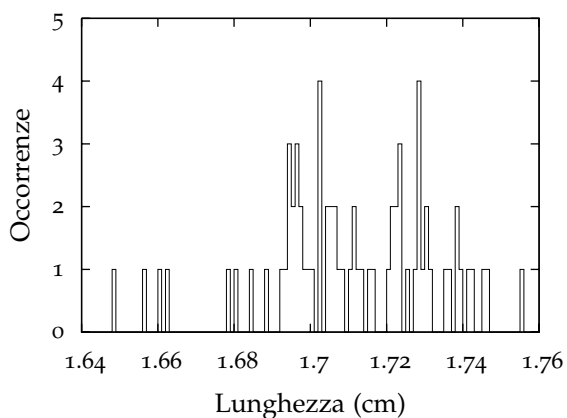
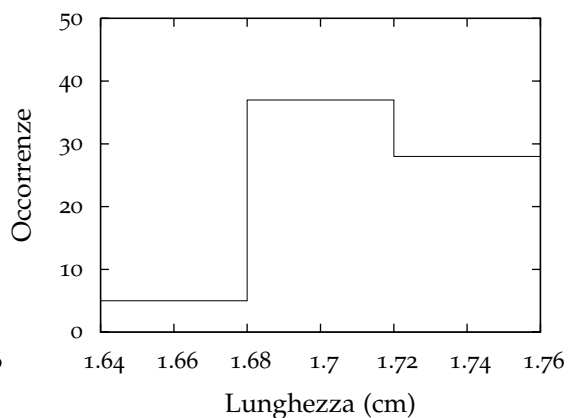


FIGURA 4.. Istogramma normalizzato ad 1 dei dati dell'esempio 2.22. La larghezza di ciascuno dei canali dell'istogramma è pari a 0.010 cm



(5.1) Istogramma con 107 canali (di larghezza pari a 0.001 cm).



(5.2) Istogramma con 3 canali (di larghezza pari a 0.04 cm).

FIGURA 5.. Istogrammi normalizzati ad N dei dati dell'esempio 2.22 per diverse scelte dell'ampiezza dei canali. In un caso (5.1) l'istogramma è sostanzialmente equivalente ad una normale rappresentazione in frequenza e dominato dalle fluttuazioni statistiche; nell'altro (5.2) il numero di canali è così piccolo che la quantità di informazione è molto ridotta rispetto alla tabella dei dati.



La (28) non è una regola assoluta: se il risultato della divisione è un numero scomodo per

i calcoli lo si può tranquillamente approssimare per ottenere un valore più comodo.

► **ESEMPIO 2.23.** Nell'esempio 2.22 si ha $x_{\min} = 1.649$, $x_{\max} = 1.756$, con $N = 70$, per cui la (28) fornisce $\Delta x \approx 0.013$ cm. Per confronto, il valore scelto per gli istogrammi riportati in figure 3 e 4 è 0.01 cm. □

Per inciso, notiamo che talvolta si usa rappresentare i dati con istogrammi anche se i valori assunti dalla variabile sono discreti ed in numero limitato. Tale rappresentazione si usa solo nel caso in cui il numero dei dati a disposizione non è molto maggiore dei possibili valori assunti dalla variabile casuale, per cui è necessario raggruppare i dati a disposizione se si vuole ottenere una rappresentazione degli stessi abbastanza regolare.

2.5 CARATTERISTICHE COMUNI ALLE DISTRIBUZIONI

2.5.1 Premessa

Premettiamo qui alcune definizioni che ci saranno utili nel seguito. Definiamo:

$$\sum_{i=1}^n a_i = a_1 + a_2 + a_3 + \cdots + a_n \quad (29)$$

Elenchiamo di seguito alcune delle proprietà formali di cui gode la sommatoria.

(1) Il significato della (29) non cambia se si cambia nome all'indice su cui si esegue la somma:

$$\sum_{i=1}^n a_i = \sum_{\ell=1}^n a_\ell = \sum_{b=1}^n a_b = \sum_{\varsigma=1}^n a_\varsigma \quad (30)$$

che si esprime solitamente dicendo che l'indice è *muto*.

(2) Per la proprietà distributiva del prodotto rispetto all'addizione si ha:

$$\sum_{i=1}^n c \cdot a_i = c \cdot \sum_{i=1}^n a_i \quad (31)$$

cioè una costante (leggi: una grandezza che non dipende dall'indice su cui si somma) può essere *portata fuori* dalla sommatoria.

(3) Il quadrato di una sommatoria (lo useremo in seguito) si scrive come:

$$\left(\sum_{i=1}^n a_i \right)^2 = \sum_{i=1}^n a_i \cdot \sum_{j=1}^n a_j = \sum_{i=1}^n \sum_{j=1}^n a_i a_j = \sum_{i,j=1}^n a_i a_j \quad (32)$$

► **ESEMPIO 2.24.** *Scriviamo esplicitamente il quadrato di una sommatoria di tre termini. Usando la notazione appena introdotta, si ha:*

$$\left(\sum_{i=1}^3 a_i \right)^2 = (a_1 + a_2 + a_3)^2 = a_1 a_1 + a_2 a_1 + a_3 a_1 + a_1 a_2 + a_2 a_2 + a_3 a_2 + a_1 a_3 + a_2 a_3 + a_3 a_3$$

□

2.5.2 Notazioni

Prima di andare avanti ci soffermiamo un attimo per fissare le notazioni che useremo nel seguito. Abbiamo definito la funzione di distribuzione di una variabile casuale discreta come l'insieme delle probabilità delle realizzazioni elementari del nostro fenomeno $\{P(x_i)\}$ e quella di una variabile casuale continua come la densità di probabilità $p(x)$. Nel seguito seguiremo sempre la convenzione di indicare con una lettera maiuscola (non necessariamente la P) la funzione di distribuzione di una variabile discreta e con una lettera minuscola (non necessariamente la p) la funzione di distribuzione di una variabile continua.

Avremo spesso a che fare con *famiglie* di funzioni di distribuzioni dipendenti da uno o più parametri. In tal caso, per convenzione, scriveremo i nomi dei parametri (quando necessario) separati dalla variabile tramite un punto e virgola; $B(n; p_1, p_2 \dots p_n)$ o $u(x; p_1, p_2 \dots p_n)$, tanto per fare un esempio.

► **ESEMPIO 2.25.** *Abbiamo incontrato nell'esempio 2.18 il primo caso di una distribuzione, quella uniforme, di cui parleremo diffusamente nel seguito (cfr. sezione 3.3). Nel caso in questione si trattava di una distribuzione uniforme tra 0 ed 1, ma è chiaro che, presi due numeri reali a e b (con $a < b$) possiamo definire la funzione di distribuzione uniforme tra a e b . Seguendo le nostre convenzioni ci riferiremo ad essa con la notazione $u(x; a, b)$ o, quando gli estremi saranno chiari dal contesto, semplicemente $u(x)$. La lettera minuscola indica che si tratta di una variabile continua (u sta per uniforme) ed i parametri della famiglia (in questo caso i due estremi a e b) figurano accanto al nome della variabile (x) separati da un punto e virgola.*

□

Senza eccedere in fantasia, riserveremo la lettera P maiuscola per indicare la probabilità che una variabile casuale discreta assuma un valore fissato, ad esempio $P(x = 3)$ o semplicemente $P(3)$, ove non vi siano ambiguità, o che una variabile casuale continua assuma un valore compreso in un certo intervallo, ad esempio $P(1 \leq x \leq 2)$.

2.5.3 Definizioni di base: normalizzazione e valor medio

Sia x una variabile casuale, discreta o continua e sia $P(x)$ la distribuzione di probabilità associata (nel caso in cui x sia discreta), oppure $p(x)$ la densità di probabilità (se x è

continua). Elenchiamo di seguito alcune proprietà e definizioni che sono comuni a tutte le distribuzioni e che utilizzeremo ampiamente nel seguito.

Per prima cosa vogliamo che, in accordo alle nostre definizioni di probabilità, la somma su tutti i casi possibili della probabilità sia uguale ad uno. Questo si esprime matematicamente imponendo la *condizione di normalizzazione*:

$$1 = \begin{cases} \sum_{i=1}^n P(x_i) & \text{per una variabile discreta} \\ \int_{-\infty}^{\infty} p(x) dx & \text{per una variabile continua} \end{cases} \quad (33)$$

Definiamo la *media* μ di una variabile casuale x come :

$$\mu = \begin{cases} \sum_{i=1}^n x_i \cdot P(x_i) & \text{per una variabile discreta} \\ \int_{-\infty}^{\infty} x \cdot p(x) dx & \text{per una variabile continua} \end{cases} \quad (34)$$

Possiamo giustificare la plausibilità di questa definizione applicandola al caso particolare di una variabile casuale discreta in cui i valori possibili $x_1 \dots x_n$ siano equiprobabili:

$$P(x_1) = P(x_2) = \dots = P(x_n) = P_0$$

In questo caso la condizione di normalizzazione:

$$\sum_{i=1}^n P(x_i) = \sum_{i=1}^n P_0 = nP_0 = 1$$

impone che:

$$P_0 = \frac{1}{n}$$

e la media si scrive, secondo la (34), come:

$$\mu = \sum_{i=1}^n x_i \cdot P(x_i) = \sum_{i=1}^n x_i \cdot P_0 = \sum_{i=1}^n x_i \cdot \frac{1}{n} = \frac{1}{n} \sum_{i=1}^n x_i$$

In questo caso, dunque, la media della variabile coincide proprio con la media aritmetica dei valori che essa può assumere. Notiamo ancora che, nel caso in cui i valori x_i non siano equiprobabili, essi vengono correttamente pesati in modo diverso nella (34) a seconda della probabilità corrispondente.

► **ESEMPIO 2.26.** Sia la variabile casuale x l'uscita di un dado equo a sei facce, di cui abbiamo già scritto esplicitamente la funzione di distribuzione nell'esempio 2.16. Quanto vale la media? Si ha:

$$\mu = \sum_{i=1}^6 x_i \cdot P(x_i) = \sum_{i=1}^6 x_i \cdot \frac{1}{6} = \frac{1}{6} \cdot (1 + 2 + 3 + 4 + 5 + 6) = \frac{21}{6} = 3.5$$

Notiamo esplicitamente che in questo caso il valor medio non coincide con nessuno dei valori possibili della variabile x . \square

► **ESEMPIO 2.27.** Ripetiamo l'esercizio precedente nel caso in cui la variabile x sia la somma delle uscite nel lancio di due dadi (cfr. esempio 2.17).

$$\begin{aligned} \mu &= \sum_{i=1}^{11} x_i \cdot P(x_i) = \\ &= 2 \cdot \frac{1}{36} + 3 \cdot \frac{2}{36} + 4 \cdot \frac{3}{36} + 5 \cdot \frac{4}{36} + 6 \cdot \frac{5}{36} + 7 \cdot \frac{6}{36} + 8 \cdot \frac{5}{36} + \\ &+ 9 \cdot \frac{4}{36} + 10 \cdot \frac{3}{36} + 11 \cdot \frac{2}{36} + 12 \cdot \frac{1}{36} = \frac{252}{36} = 7 \end{aligned}$$

che è esattamente il doppio di quanto calcolato poco fa per un singolo dado. Vedremo nel seguito che si tratta in effetti di un caso particolare di una relazione la cui validità è molto più generale. \square

► **ESEMPIO 2.28.** Sia x una variabile casuale definita nell'intervallo $[0, 2]$ con densità di probabilità (cfr. figura 6):

$$p(x) = \begin{cases} 1 - \frac{1}{2}x & 0 \leq x \leq 2 \\ 0 & x < 0; x > 2 \end{cases}$$

La media della distribuzione (che è un caso particolare di distribuzione triangolare) è data da:

$$\begin{aligned} \mu &= \int_{-\infty}^{\infty} x \cdot p(x) dx = \int_0^2 x \cdot \left(1 - \frac{1}{2}x\right) dx = \int_0^2 x dx - \frac{1}{2} \int_0^2 x^2 dx = \\ &= \left[\frac{x^2}{2}\right]_0^2 - \left[\frac{x^3}{6}\right]_0^2 = 2 - \frac{4}{3} = \frac{2}{3} \approx 0.666 \end{aligned}$$

Notiamo esplicitamente come la media non coincida, in questo caso particolare, con il valor medio dell'intervallo in cui la variabile è definita (cioè 1): essendo i valori prossimi allo 0 più probabili degli altri, μ tende corrispondentemente a spostarsi verso lo 0. \square

Definiamo *mediana* di una distribuzione quel valore $\mu_{\frac{1}{2}}$ della variabile casuale tale che:

$$\begin{cases} P\left(x \leq \mu_{\frac{1}{2}}\right) = P\left(x \geq \mu_{\frac{1}{2}}\right) & \text{per una variabile discreta} \\ \int_{-\infty}^{\mu_{\frac{1}{2}}} p(x) dx = \int_{\mu_{\frac{1}{2}}}^{\infty} p(x) dx & \text{per una variabile continua} \end{cases} \quad (35)$$

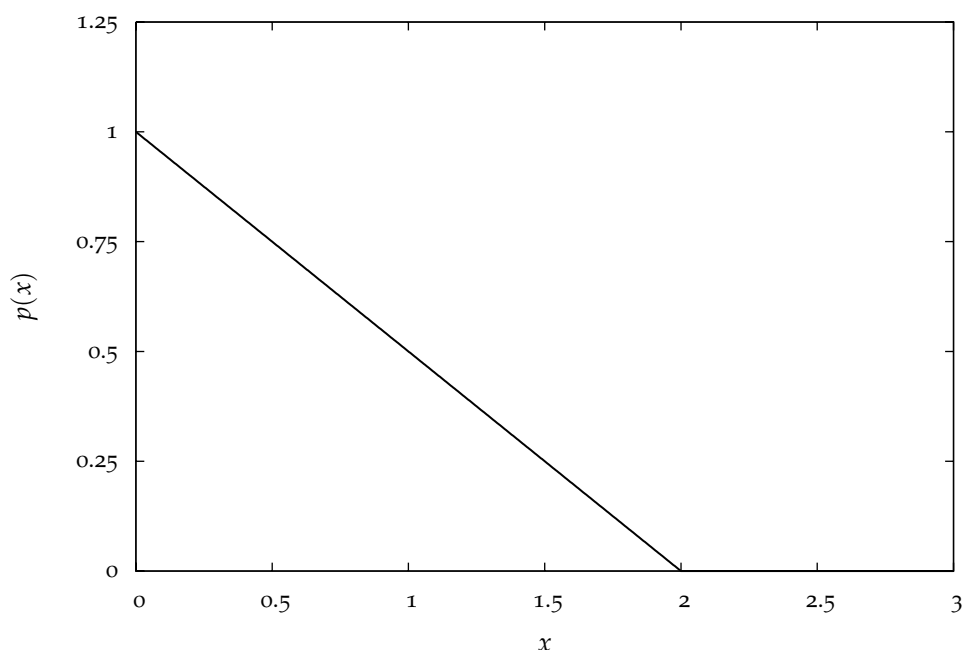



FIGURA 6.. Grafico della distribuzione triangolare definita nell'esempio 2.28. Si può verificare facilmente che la condizione di normalizzazione (33) è verificata.

Per una distribuzione di variabile casuale continua la mediana in generale esiste ed in più vale la condizione:

$$\int_{-\infty}^{\mu_{\frac{1}{2}}} p(x) dx = \int_{\mu_{\frac{1}{2}}}^{\infty} p(x) dx = \frac{1}{2}$$

 Nel caso di una variabile discreta non è detto che questo valore esista (cfr. esempio 2.29) per cui la trattazione è delicata. In letteratura si trovano diverse convenzioni tra le quali molto comune è la seguente: se le variabili costituiscono una popolazione finita a $(2N + 1)$ elementi, dove N è un intero positivo, la mediana è il valore del $(N + 1)^{\text{mo}}$ elemento ordinato. Se ci sono $2N$ elementi ogni valore dell'intervallo tra i due elementi di mezzo N ed $N + 1$ soddisfa la 2.17 e questa non unicità è convenzionalmente rimossa definendo la mediana essere a metà tra i valori dell' N^{mo} e dell' $(N + 1)^{\text{mo}}$ elemento.

Notiamo che se la funzione di distribuzione è simmetrica, allora la mediana (se esiste) e la media coincidono.

► ESEMPIO 2.29. Consideriamo, al solito, l'uscita di un dado equo a sei facce (cfr. esempio 2.16).

In questo caso, se seguiamo la definizione (35), la mediana non esiste perché non esiste alcun valore della variabile casuale per cui essa sia soddisfatta, come si può verificare facilmente:

$$\begin{aligned}\frac{1}{6} &= P(x \leq 1) \neq P(x \geq 1) = 1 \\ \frac{2}{6} &= P(x \leq 2) \neq P(x \geq 2) = \frac{5}{6} \\ \frac{3}{6} &= P(x \leq 3) \neq P(x \geq 3) = \frac{4}{6} \\ \frac{4}{6} &= P(x \leq 4) \neq P(x \geq 4) = \frac{3}{6} \\ \frac{5}{6} &= P(x \leq 5) \neq P(x \geq 5) = \frac{2}{6} \\ 1 &= P(x \leq 6) \neq P(x \geq 6) = \frac{1}{6}\end{aligned}$$

□

► ESEMPIO 2.30. Nel caso della distribuzione triangolare dell'esempio 2.28 la mediana è data dall'equazione:

$$\begin{aligned}\frac{1}{2} &= \int_{-\infty}^{\mu_{\frac{1}{2}}} p(x) dx = \int_0^{\mu_{\frac{1}{2}}} \left(1 - \frac{1}{2}x\right) dx = \left[x - \frac{x^2}{4}\right]_0^{\mu_{\frac{1}{2}}} = \\ &= \mu_{\frac{1}{2}} - \frac{\mu_{\frac{1}{2}}^2}{4}\end{aligned}$$

da cui:

$$\mu_{\frac{1}{2}} = 2 - \sqrt{2} \approx 0.586$$

□

Definiamo il *valore più probabile* (che alcuni autori chiamano anche *moda*) come il valore della variabile casuale (se esiste) in corrispondenza del quale la funzione di distribuzione ha un massimo.

► ESEMPIO 2.31. Nel caso dell'uscita di un dado equo i sei valori possibili della variabile casuale sono equiprobabili per cui il valore più probabile non è definito. □

► ESEMPIO 2.32. Nel caso della distribuzione triangolare dell'esempio 2.28 il valore più probabile è $x = 0$. □

2.5.4 *Valore di aspettazione o di previsione*

Consideriamo una funzione $f(x)$ di una variabile casuale x ; definiamo il *valore di aspettazione* o *valore di previsione* di $f(x)$ come:

$$E[f(x)] = \begin{cases} \sum_{i=1}^n f(x_i) \cdot P(x_i) & \text{per una variabile discreta} \\ \int_{-\infty}^{\infty} f(x) \cdot p(x) dx & \text{per una variabile continua} \end{cases} \quad (36)$$

Si può vedere facilmente che, nel caso particolare $f(x) = x$, si ha:

$$E[x] = \mu$$

cioè il valor medio di una variabile x , definito dalla (34), non è altro che il valore di aspettazione della variabile stessa. Per questo motivo il valore di aspettazione di una funzione $f(x)$ viene anche detto *valor medio* di $f(x)$. Elenchiamo di seguito alcune proprietà dell'operatore E .

(1) Il valore di aspettazione di una costante c è uguale alla costante stessa:

$$E[c] = c$$

(2) L'operatore E è un operatore lineare, nel senso che, se c_1 e c_2 sono due costanti e $f(x)$ e $g(x)$ sono due funzioni di una variabile casuale x , allora:

$$E[c_1 \cdot f(x) + c_2 \cdot g(x)] = c_1 \cdot E[f(x)] + c_2 \cdot E[g(x)]$$

Sono casi particolari di quanto appena detto le seguenti espressioni, che ci saranno utili nel seguito:

$$\begin{aligned} E[c_1 \cdot f(x)] &= c_1 \cdot E[f(x)] \\ E[f(x) + g(x)] &= E[f(x)] + E[g(x)] \end{aligned}$$

La dimostrazione di queste proprietà sfrutta sostanzialmente la linearità della sommatoria e dell'integrale.

2.5.5 *Dispersione intorno alla media: varianza e deviazione standard*

Ci poniamo in questo paragrafo il problema di determinare, data una variabile casuale x , una funzione $f(x)$ tale che il valore di aspettazione $E[f(x)]$ indichi quanto in media la variabile stessa tenda ad essere lontana dal suo valor medio (con lieve abuso di termini si usa anche dire: quanto la distribuzione sia *larga* o *stretta*). Come vedremo in seguito questo problema è strettamente legato alla valutazione dell'errore statistico da associare ad una serie di misure sperimentali di una variabile casuale.

Ingenualmente potremmo pensare di prendere, quale funzione adatta allo scopo:

$$f(x) = x - \mu$$

cioè la funzione che rappresenta lo scarto dalla media; tuttavia si ha:

$$E[x - \mu] = E[x] - E[\mu] = \mu - \mu = 0$$

Quindi questo valore di aspettazione di fatto non fornisce informazione alcuna. D'altra parte era logico aspettarselo in quanto gli scarti per difetto tendono statisticamente a compensare quelli per eccesso. Una seconda possibilità è data da

$$f(x) = |x - \mu|$$

Si tratta di una quantità assolutamente *ragionevole* il cui difetto maggiore risiede nel non essere particolarmente comoda per i calcoli. In effetti quello che si usa di solito è

$$f(x) = (x - \mu)^2$$

Definiamo allora *varianza* di una distribuzione la quantità:

$$\sigma^2 = E[(x - \mu)^2] = \begin{cases} \sum_{i=1}^n (x_i - \mu)^2 \cdot P(x_i) & \text{per una variabile discreta} \\ \int_{-\infty}^{\infty} (x - \mu)^2 \cdot p(x) dx & \text{per una variabile continua} \end{cases} \quad (37)$$

e la *deviazione standard* come la radice quadrata della varianza:

$$\sigma = \sqrt{\sigma^2} \quad (38)$$

► **ESEMPIO 2.33.** Sia la variabile casuale x l'uscita di un dado equo a sei facce (cfr. esempio 2.16 e 2.26). La varianza della funzione di distribuzione si scrive:

$$\begin{aligned} \sigma^2 &= \sum_{i=1}^6 (x_i - \mu)^2 \cdot P(x_i) = \sum_{i=1}^6 \left(x_i - \frac{7}{2}\right)^2 \cdot \frac{1}{6} = \\ &= \left(\frac{25}{4} + \frac{9}{4} + \frac{1}{4} + \frac{1}{4} + \frac{9}{4} + \frac{25}{4}\right) \cdot \frac{1}{6} = \frac{35}{2} \cdot \frac{1}{6} = \frac{35}{12} \approx 2.917 \end{aligned}$$

e la deviazione standard:

$$\sigma = \sqrt{\frac{35}{12}} \approx 1.708$$

□

► ESEMPIO 2.34. Sia la variabile x la somma delle uscite nel lancio di due dadi (cfr. esempio 2.17 e 2.27). la varianza è:

$$\begin{aligned}\sigma^2 &= \sum_{i=1}^{11} (x_i - \mu)^2 \cdot P(x_i) = \sum_{i=1}^{11} (x_i - 7)^2 \cdot P(x_i) = \\ &= 25 \cdot \frac{1}{36} + 16 \cdot \frac{2}{36} + 9 \cdot \frac{3}{36} + 4 \cdot \frac{4}{36} + 1 \cdot \frac{5}{36} + 0 \cdot \frac{6}{36} + 1 \cdot \frac{5}{36} + \\ &+ 4 \cdot \frac{4}{36} + 9 \cdot \frac{3}{36} + 16 \cdot \frac{2}{36} + 25 \cdot \frac{1}{36} = \frac{210}{36} = \frac{35}{6} \approx 5.833\end{aligned}$$

che è esattamente il doppio di quanto calcolato poco fa per un dado. Vedremo nel seguito che, come per la media (cfr. esempio 2.27), si tratta in effetti di un caso particolare di un risultato importante della statistica la cui validità è molto più generale. La deviazione standard è:

$$\sigma = \sqrt{\frac{35}{6}} \approx 2.415$$

□

► ESEMPIO 2.35. Nel caso della distribuzione triangolare dell'esempio 2.28 la varianza è:

$$\begin{aligned}\sigma^2 &= \int_{-\infty}^{\infty} (x - \mu)^2 \cdot p(x) dx = \int_0^2 \left(x - \frac{2}{3}\right)^2 \cdot \left(1 - \frac{1}{2}x\right) dx = \\ &= \int_0^2 \left(-\frac{1}{2}x^3 + \frac{5}{3}x^2 - \frac{14}{9}x + \frac{4}{9}\right) dx = \left[-\frac{1}{8}x^4 + \frac{5}{9}x^3 - \frac{14}{18}x^2 + \frac{4}{9}x\right]_0^2 = \\ &= \left(-2 + \frac{40}{9} - \frac{28}{9} + \frac{8}{9}\right) = \frac{2}{9} \approx 0.222\end{aligned}$$

e la deviazione standard:

$$\sigma = \sqrt{\frac{2}{9}} \approx 0.471$$

□

Partendo dalla definizione di varianza data in (37) possiamo scrivere:

$$\begin{aligned}\sigma^2 &= E[(x - \mu)^2] = E[(x^2 + \mu^2 - 2x\mu)] = \\ &= E[x^2] + E[\mu^2] - E[2x\mu] = \\ &= E[x^2] + \mu^2 - 2\mu \cdot E[x] = \\ &= E[x^2] + \mu^2 - 2\mu^2 = E[x^2] - \mu^2\end{aligned}$$

Siamo arrivati così all'importante risultato:

$$\boxed{\sigma^2 = E[x^2] - \mu^2} \quad (39)$$

Anticipiamo sin da ora che la (39) risulta spesso utile nei calcoli e la useremo sovente nel seguito.

► **ESEMPIO 2.36.** Ricalcoliamo la varianza della distribuzione triangolare dell'esempio 2.28 utilizzando la (39):

$$\begin{aligned} E[x^2] &= \int_{-\infty}^{\infty} x^2 \cdot p(x) dx = \int_0^2 x^2 \cdot \left(1 - \frac{1}{2}x\right) dx = \int_0^2 \left(x^2 - \frac{1}{2}x^3\right) dx = \\ &= \left[\frac{1}{3}x^3 - \frac{1}{8}x^4\right]_0^2 = \left(\frac{8}{3} - 2\right) = \frac{2}{3} \end{aligned}$$

per cui, ricordando che $\mu = \frac{2}{3}$, come mostrato nell'esempio 2.28:

$$\sigma^2 = E[x^2] - \mu^2 = \frac{2}{3} - \frac{4}{9} = \frac{2}{9}$$

che è proprio il risultato trovato nell'esempio 2.35. □

2.5.6 La semilarghezza a metà altezza

Sebbene la nozione di semilarghezza a metà altezza non si applichi in modo banale a tutte le distribuzioni, si tratta di una quantità così spesso usata in letteratura che vale la pena, a questo punto, accennarvi con un certo dettaglio. Supporremo, nel seguito di questo paragrafo, di avere a che fare con una variabile casuale continua x la cui funzione di distribuzione $p(x)$ presenti un solo massimo locale (in caso contrario la trattazione si complicherebbe). Si tratta di un'assunzione assolutamente ragionevole, che come vedremo sarà soddisfatta dalla maggior parte delle distribuzioni che incontreremo nel seguito. Nelle nostre ipotesi ci sono esattamente due valori di x (che chiameremo qui x_+ e x_-) per cui il valore della funzione di distribuzione è esattamente la metà del valore che la funzione stessa assume nel massimo (che sappiamo essere unico). In altre parole possiamo dire che la retta orizzontale che interseca l'asse y in corrispondenza della metà dell'altezza della funzione di distribuzione intersecherà il grafico della funzione stessa proprio nei due punti x_+ e x_- (cfr. figura 7).

Definiamo allora la semilarghezza a metà altezza, che si trova spesso indicata con l'acronimo HWHM³, come:

$$\text{HWHM} = \frac{|x_+ - x_-|}{2} \quad (40)$$

Spesso, in letteratura, si usa anche la larghezza intera a metà altezza, che si indica con l'acronimo FWHM⁴ e si definisce banalmente come:

$$\text{FWHM} = |x_+ - x_-| \quad (41)$$

³ HWHM sta qui per l'espressione inglese *Half Width at Half Maximum*, che significa appunto semilarghezza a metà altezza.

⁴ Corrispondentemente FWHM sta per *Full Width at Half Maximum*.

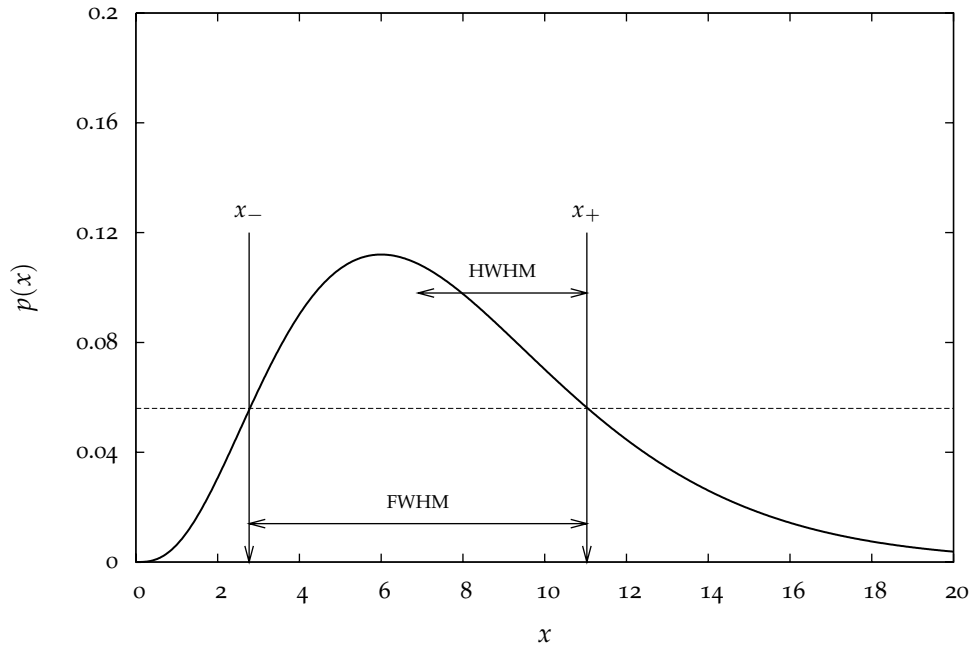


FIGURA 7.. Significato geometrico della semilarghezza a metà altezza. Notiamo esplicitamente che non è necessario che la funzione di distribuzione sia simmetrica rispetto al valor medio.

► ESEMPIO 2.37. Torniamo ancora una volta alla distribuzione triangolare definita nell'esercizio 2.28. In questo caso abbiamo banalmente:

$$\begin{aligned}x_- &= 0 \\ x_+ &= 1\end{aligned}$$

per cui la semilarghezza a metà altezza è:

$$\text{HWHM} = \frac{1}{2}$$

Utilizzando il risultato dell'esempio 2.35 possiamo anche calcolare il rapporto:

$$\frac{\text{HWHM}}{\sigma} = \frac{1}{2} \cdot \sqrt{\frac{9}{2}} = \sqrt{\frac{9}{8}} \approx 1.06$$

che è estremamente vicino all'unità. □

Il fatto che, nell'esempio precedente, il rapporto tra la semilarghezza e la deviazione standard della distribuzione fosse così prossimo ad 1 non deve stupire ed è solamente un caso particolare di un risultato che, per lo meno a livello qualitativo, ha una validità molto più generale. In effetti abbiamo già detto che la deviazione standard (proprio come la HWHM), costituisce una stima di quanto una data distribuzione sia *larga*, in un senso che abbiamo precisato nel paragrafo precedente. Se scriviamo allora:

$$\text{HWHM} = \varepsilon \sigma \tag{42}$$

(cosa che è sempre possibile, a patto di scegliere opportunamente la costante ε) possiamo precisare quanto ottenuto fino ad ora dicendo che ε , pur dipendendo dalla particolare forma della distribuzione, è dell'ordine dell'unità. Notiamo esplicitamente che, nota la forma analitica della funzione di distribuzione, ε può essere calcolato come nell'esempio 2.37; in ogni caso, comunque, la semilarghezza a metà altezza costituisce una stima più o meno accurata (dipende dai dettagli della funzione di distribuzione) della deviazione standard.

2.5.7 Momenti di una distribuzione

Normalizzazione, media e varianza sono casi particolari dei più generali *momenti centrati di ordine n* così definiti:

$$\mathcal{M}_n(a) = E[(x - a)^n] \quad (43)$$

Normalizzazione e media sono rispettivamente i momenti *centrati nell'origine* di ordine zero ed uno. Infatti il momento di ordine n centrato nell'origine si scrive come:

$$\mathcal{M}_n(0) = E[(x - 0)^n] = E[x^n]$$

ed è immediato verificare che

$$\begin{aligned} \mathcal{M}_0(0) &= 1 \\ \mathcal{M}_1(0) &= \mu \end{aligned}$$

I momenti di ordine generico rispetto al valor medio μ (*momenti centrati nel valor medio*) sono così spesso usati da meritare una notazione dedicata. Nel seguito ogni volta che ometteremo il valore attorno a cui un momento è centrato, intenderemo il momento stesso centrato attorno al valor medio:

$$\mathcal{M}_n \equiv \mathcal{M}_n(\mu) = E[(x - \mu)^n] \quad (44)$$

Con questo significato dei termini si ha, come è immediato verificare:

$$\begin{aligned} \mathcal{M}_0 &= E[1] = 1 \\ \mathcal{M}_1 &= E[(x - \mu)^1] = 0 \\ \mathcal{M}_2 &= E[(x - \mu)^2] = \sigma^2 \end{aligned}$$

Mediante il momento di ordine tre si costruisce il parametro adimensionale:

$$\gamma_1 = \frac{\mathcal{M}_3}{\sigma^3} \quad (45)$$

chiamato coefficiente di asimmetria o *skewness*. Infatti \mathcal{M}_3 è nullo quando la distribuzione è simmetrica rispetto a μ .

► ESEMPIO 2.38. Torniamo ancora una volta sulla distribuzione triangolare dell'esempio 2.28 e calcoliamone il coefficiente di asimmetria. Il momento terzo centrato nel valor medio è:

$$\begin{aligned}
 \mathcal{M}_3 &= \int_{-\infty}^{\infty} (x - \mu)^3 \cdot p(x) \, dx = \int_0^2 \left(x - \frac{2}{3}\right)^3 \cdot \left(1 - \frac{1}{2}x\right) \, dx = \\
 &= \int_0^2 \left(-\frac{1}{2}x^4 + 2x^3 - \frac{8}{3}x^2 + \frac{40}{27}x - \frac{8}{27}\right) \, dx = \\
 &= \left[-\frac{1}{10}x^5 + \frac{1}{2}x^4 - \frac{8}{9}x^3 + \frac{20}{27}x^2 - \frac{8}{27}x\right]_0^2 = \\
 &= \left(-\frac{16}{5} + 8 - \frac{64}{9} + \frac{80}{27} - \frac{16}{27}\right) = \frac{8}{135}
 \end{aligned}$$

e la skewness:

$$\gamma_1 = \frac{\mathcal{M}_3}{\sigma^3} = \frac{8}{135} \cdot \left(\frac{9}{2}\right)^{\frac{3}{2}} = \frac{2^{\frac{3}{2}}}{5} \approx 0.566$$

Come ci aspettavamo, non essendo la distribuzione simmetrica rispetto alla media, il coefficiente di asimmetria risulta diverso da zero. Il segno (positivo) ci dice che la zona con minore densità di probabilità sta a destra del valor medio. \square

Mediante il momento di ordine quattro \mathcal{M}_4 si costruisce il parametro adimensionale:

$$\gamma_2 = \frac{\mathcal{M}_4}{\sigma^4} \quad (46)$$

chiamato anche *curtosi*, che misura quanto è piatta una distribuzione (quanto più è piatta tanto minore è γ_2).

Nota la funzione di distribuzione, si possono conoscere i momenti di tutti gli ordini, purché esistano. Per inciso notiamo che è vero anche il viceversa: se si conoscono tutti i momenti di una certa variabile casuale è possibile ottenere la funzione di distribuzione di tale variabile.

2.5.8 Discussione e cenno al Teorema di Tschebyscheff

Quando si hanno misure i cui risultati fluttuano, il risultato della misura è una variabile casuale funzione delle condizioni sperimentali. Come vedremo in dettaglio più avanti la media di questa variabile casuale è il valore più significativo del risultato. La deviazione standard misura l'incertezza da attribuire a questo risultato. Il fatto che si possa assumere come valore buono la media e come misura dell'incertezza la deviazione standard è fondato su di un importante teorema di statistica dovuto a Tschebyscheff.

TEOREMA 2.1 (DI TSCHEBYSCHIEFF). Sia x una variabile casuale tale che esistano finiti μ e σ . Detto k un numero positivo, si ha:

$$P(|x - \mu| \geq k\sigma) \leq \frac{1}{k^2} \quad (47)$$

cioè è poco probabile avere grandi deviazioni dalla media e tali deviazioni sono convenientemente misurate in unità di σ . Il teorema si dimostra notando che, se nella definizione di σ^2 (che per comodità scriviamo nel caso continuo):

$$\sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 \cdot p(x) dx$$

mettiamo al posto di $(x - \mu)^2$ zero se esso è minore di $k^2\sigma^2$ (dove k è una costante assegnata) e $k^2\sigma^2$ negli altri casi, otteniamo:

$$\sigma^2 \geq k^2\sigma^2 \int_{|x-\mu| \geq k\sigma} p(x) dx = k^2\sigma^2 \cdot P(|x - \mu| \geq k\sigma)$$

che è esattamente la relazione (47).

► **ESEMPIO 2.39.** Sia x una variabile casuale distribuita secondo la funzione triangolare dell'esempio 2.28. Ci chiediamo quale sia la probabilità che x disti più di 2 deviazioni standard dal suo valor medio. Notiamo anzitutto che x non può essere minore di $\mu - 2\sigma$ poiché questo valore è escluso dall'intervallo su cui la funzione di distribuzione è definita. Il valore cercato è perciò:

$$\begin{aligned} P(|x - \mu| \geq 2\sigma) &= \int_{\mu+2\sigma}^{\infty} p(x) dx = \int_{\frac{2}{3}+2\sqrt{\frac{2}{9}}}^2 \left(1 - \frac{1}{2}x\right) dx = \\ &= \left[x - \frac{1}{4}x^2\right]_{\frac{2}{3}+2\sqrt{\frac{2}{9}}}^2 = \frac{6 - 4\sqrt{2}}{9} \approx 0.038 \end{aligned}$$

Il teorema di Tschebyscheff fornisce in questo caso il limite superiore:

$$P(|x - \mu| \geq 2\sigma) \leq \left(\frac{1}{2}\right)^2 = \frac{1}{4} = 0.25$$

che è ampiamente soddisfatto. □

L'esempio 2.39 mette in luce come, molto spesso, il teorema di Tschebyscheff fornisca un limite superiore piuttosto *blando* alla probabilità che una variabile casuale si discosti più di una quantità fissata dal proprio valor medio (nel caso in questione è, a 2σ , il 25% contro un valore effettivo di meno del 4%). In effetti, nota la funzione di distribuzione, la quantità $P(|x - \mu| \geq k\sigma)$ può essere calcolata esplicitamente per qualsiasi valore di k . D'altra parte il teorema è assolutamente generale, nel senso che vale per una funzione di distribuzione arbitraria.

2.5.9 Cenni alla teoria dei campioni

Vediamo quali sono le ipotesi che si fanno applicando i concetti statistici fino ad ora trattati ai risultati delle misure sperimentali:

1. L'insieme delle misure sperimentali (che non è che una piccola parte di tutte le misure effettuabili) è un *campione* di tutte le possibili misure.

2. Si postula l'esistenza di una distribuzione dei risultati di tutte le misure effettuabili: questa viene chiamata *distribuzione generatrice* (*parent distribution*).
3. Dal campione dobbiamo ricavare la migliore stima della media μ e della varianza σ^2 della distribuzione generatrice, che sono ignote.

Questo rientra in un discorso generale nell'ambito della teoria dei campioni, che qui non tratteremo. Per completezza diamo qui le espressioni per la migliore stima (a partire da un campione finito di misure) della media e della deviazione standard della distribuzione generatrice, che saranno ricavate nel paragrafo 4.2:

$$m = \frac{1}{n} \sum_{i=1}^n x_i$$

$$s = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - m)^2}$$

Notiamo che molti testi postulano l'esistenza di un *valore vero* della grandezza da misurare. Questo postulato è non solo non necessario, ma anche sbagliato: non esiste il *valore vero di una grandezza fisica*. Quanto più si raffinano gli strumenti tanto meglio si scopre che le grandezze fisiche fluttuano: al limite in cui si arriva alla struttura atomica il principio di indeterminazione afferma che il risultato è inevitabilmente governato da leggi probabilistiche. Quindi poniamo il seguente postulato: non esiste un valore vero, ma esiste il valore medio della distribuzione generatrice. Questo è dunque il valore che è più ragionevole accettare per rappresentare la grandezza in questione.

2.6 ESERCIZI

► ESERCIZIO 2.1. Verificare *sperimentalmente* la (21) nel caso del lancio di una moneta (verificare, cioè, che il rapporto tra il numero di teste n ed il numero di lanci N tende ad $\frac{1}{2}$ al crescere di N).

Suggerimento: operativamente si può costruire un grafico cartesiano in cui, dopo ogni lancio, si pone N sull'asse delle x e n/N su quello delle y .

► ESERCIZIO 2.2. Sia data una scatola contenente 6 palline rosse, 3 blu e 5 bianche. Supponiamo di estrarre una pallina a caso dalla scatola; calcolare la probabilità che essa sia blu o bianca.

Risposta: $\left[\frac{4}{7} \right]$

► ESERCIZIO 2.3. Nel lancio di una moneta qual è la probabilità di avere una successione di cinque teste, seguita da una croce?

Risposta: $\left[\frac{1}{64} \right]$

► ESERCIZIO 2.4. Nel lancio di una moneta, qual è la probabilità che la *prima* croce esca esattamente al k -esimo tentativo (cioè che escano $k - 1$ teste in successione, seguite da una testa)?

Risposta: $\left[\frac{1}{2^k} \right]$

► ESERCIZIO 2.5. Consideriamo il lancio di una dado a sei facce. Qual è la probabilità che un numero fissato (ad esempio il 3) esca per la prima volta al k -esimo tentativo (cioè che le prime $k - 1$ uscite siano diverse da 3 e la k -esima sia 3)?

Risposta: $\left[\frac{1}{6} \cdot \left(\frac{5}{6} \right)^{k-1} \right]$

► ESERCIZIO 2.6. Con riferimento all'esempio 2.19, quanti minuti dovremo aspettare l'autobus alla fermata, in media?

Suggerimento: si tratta semplicemente di calcolare la media della distribuzione scritta esplicitamente nell'esempio usando la definizione (34).

Risposta: $[5]$

► ESERCIZIO 2.7. Provate a rappresentare le funzioni di distribuzione per la somma delle uscite di uno, due e tre dadi e confrontate i grafici.

► ESERCIZIO 2.8. Un gruppo di studenti (ognuno dei quali è identificato da un numero progressivo i) misura ad occhio ed indipendentemente la lunghezza della lavagna dell'aula A. Vengono registrati i seguenti risultati (in metri)⁵:

⁵ Dati realmente ottenuti il 9 Gennaio 1990.

LUNGHEZZA (cm)									
6	4.8	6.5	5.62	3	7.5	5.5	4	5.5	5
4.5	5	4.75	4	5.8	5	5	5.5	4.6	4.7
4.5	5.3	4.75	2.80	4.50	5	6.1	4	5.5	4.6
6.60	5.80	5.4	5.5	7.5	5.9	10	5.8	3.7	5.4
5.5	5	5.5	5.8	5	5.5	5.7	5	6.2	6.1
3.45	7.5	5.40	3.5	6.2	5.18	4.5	4.80	5.30	5.5
6	5.3	6	5	6.20	6	5	5	6.5	5.5
5	5.5	5.35	5.2	5.8	4.5	5.45	6.5	6	5
6	5.75	5.25	5.6	5.6	5.9	5	6.5	8	5.40
4	3.2	5.60	6.05	5.5	6.5				

Rappresentare questi dati mediante un istogramma.

► ESERCIZIO 2.9. Dimostrare che:

$$\sum_{k=1}^n c = nc$$

► ESERCIZIO 2.10. Dimostrare che:

$$\sum_{k=1}^n k = \frac{n(n+1)}{2}$$

Suggerimento: sommare i termini a due a due (il primo con l'ultimo, il secondo con il penultimo e così via). Che cosa succede?

► ESERCIZIO 2.11. Dimostrare che:

$$\sum_{k=1}^n k^2 = \frac{n(n+1)(2n+1)}{6}$$

Suggerimento: procedere per induzione.

► ESERCIZIO 2.12. Si consideri un numero reale positivo $0 < a < 1$. Calcolare esplicitamente la quantità:

$$\sum_{k=1}^n a^k = a + a^2 + a^3 + \dots + a^n$$

Suggerimento: si scriva esplicitamente la somma e si moltiplichino numeratore e denominatore per $1 - a$. Si svolga dunque il prodotto al numeratore (un numero cospicuo di termini dovrebbe elidersi).

$$\text{Risposta: } \left[a \cdot \frac{(1-a^n)}{(1-a)} \right]$$

► ESERCIZIO 2.13. Si consideri un numero reale positivo $0 < a < 1$. Utilizzando il risultato dell'esercizio precedente si calcoli la quantità:

$$\sum_{k=1}^{\infty} a^k$$

Suggerimento: si calcoli il limite, per n che tende ad infinito, dell'espressione ottenuta nell'esercizio precedente.

$$\text{Risposta: } \left[\frac{a}{(1-a)} \right]$$

► ESERCIZIO 2.14. Si consideri un numero reale positivo $0 < a < 1$. Calcolare esplicitamente la quantità:

$$\sum_{k=1}^n ka^k = a + 2a^2 + 3a^3 + \dots + na^n$$

Suggerimento: si applichi il suggerimento dell'esercizio 2.12. Questa volta i termini non si elidono, ma applicando il risultato ottenuto nell'esercizio stesso...

$$\text{Risposta: } \left[a \cdot \frac{(1-a^n)}{(1-a)^2} - \frac{na^{n+1}}{(1-a)} \right]$$

► ESERCIZIO 2.15. Si consideri un numero reale positivo $0 < a < 1$. Utilizzando il risultato dell'esercizio precedente si calcoli la quantità:

$$\sum_{k=1}^{\infty} ka^k$$

$$\text{Risposta: } \left[\frac{a}{(1-a)^2} \right]$$

► ESERCIZIO 2.16. Verificare che la particolare distribuzione triangolare definita nell'esempio 2.28 è normalizzata.

► ESERCIZIO 2.17. Si dimostri che il valore di aspettazione di una costante è uguale al valore della costante stessa.

Suggerimento: si scriva esplicitamente il valore di aspettazione e si sfrutti la condizione di normalizzazione della funzione di distribuzione.

► ESERCIZIO 2.18. Si consideri di nuovo l'esercizio 2.5. Quanti lanci di un dado a sei facce dobbiamo aspettare, in media, perché esca un 3?

Suggerimento: sappiamo già, dall'esercizio 2.5, la probabilità che il primo 3 esca esattamente al k -esimo lancio. Si tratta allora di valutare il valore di aspettazione $E[k]$ per la specifica funzione di distribuzione. Sarà utile anche il risultato dell'esercizio 2.15.

$$\text{Risposta: } [6]$$

► ESERCIZIO 2.19. Sia la variabile casuale x l'uscita di un dado equo a sei facce. Si calcoli il valore di aspettazione $E[x^2]$.

$$\text{Risposta: } \left[\frac{91}{6} \right]$$

► ESERCIZIO 2.20. Si verifichi la validità della (39) sfruttando il risultato dell'esercizio precedente.

Suggerimento: si usino gli esempi 2.26 e 2.33.

► ESERCIZIO 2.21. Si ripeta l'esercizio 2.19 nel caso in cui la variabile casuale x sia la somma delle uscite nel lancio di due dadi (si calcoli il valore di aspettazione $E[x^2]$).

Risposta: $\left[\frac{329}{6}\right]$

► ESERCIZIO 2.22. Si verifichi la validità della (39) sfruttando il risultato dell'esercizio precedente.

Suggerimento: si sfruttino in questo caso gli esempi 2.27 e 2.34.

► ESERCIZIO 2.23. Dimostrare che per una distribuzione arbitraria il momento di ordine 3 centrato nel valor medio vale:

$$\mathcal{M}_3 = E[x^3] - 3\mu\sigma^2 - \mu^3$$

Suggerimento: scrivere esplicitamente la definizione di \mathcal{M}_3 e sfruttare la linearità dell'operatore valore di aspettazione, insieme alla (39).

► ESERCIZIO 2.24. Dimostrare che per una funzione di distribuzione simmetrica rispetto al valor medio il coefficiente di asimmetria (45) è nullo.

DISTRIBUZIONI

INDICE

3.1	Distribuzione binomiale	60
3.1.1	Premessa: i coefficienti binomiali	60
3.1.2	La distribuzione binomiale	60
3.1.3	Normalizzazione, media e varianza	63
3.2	Distribuzione di Poisson	65
3.2.1	Premessa: il numero di Nepero	65
3.2.2	La distribuzione di Poisson	66
3.2.3	Normalizzazione, media e varianza	70
3.3	Distribuzione uniforme	72
3.3.1	Normalizzazione, media e varianza	72
3.4	Distribuzione esponenziale	75
3.4.1	Normalizzazione, media e varianza	77
3.5	Distribuzione di Gauss	79
3.5.1	Premessa: la funzione di Gauss	79
3.5.2	La distribuzione di Gauss	80
3.5.3	Normalizzazione, media e varianza	82
3.5.4	Integrazione della distribuzione di Gauss	83
3.5.5	Relazione con altre distribuzioni	84
3.6	Distribuzione del χ^2	88
3.6.1	Premessa: la funzione Γ di Eulero	88
3.6.2	La distribuzione del χ^2	88
3.6.3	Normalizzazione, media e varianza	89
3.7	Distribuzione di Cauchy	90
3.7.1	Normalizzazione, media e varianza	91
3.8	Sommario	95
3.9	Esercizi	96

Questo capitolo raccoglie alcuni risultati fondamentali circa le funzioni di distribuzione più comunemente usate in fisica.

3.1 DISTRIBUZIONE BINOMIALE

3.1.1 Premessa: i coefficienti binomiali

Richiamiamo brevemente alcune definizioni che ci saranno utili nel seguito. Dato un numero intero k , il suo fattoriale è definito come:

$$k! = k \cdot (k-1) \cdot (k-2) \cdots 2 \cdot 1$$

Inoltre:

$$0! = 1! = 1$$

Il numero di possibili combinazioni di n elementi k a k , anche detto coefficiente binomiale n su k si può scrivere come:

$$\binom{n}{k} = \frac{n \cdot (n-1) \cdots (n-k+1)}{k!} = \frac{n!}{k!(n-k)!} \quad (48)$$

Utilizzando i coefficienti binomiali è possibile scrivere in modo estremamente compatto la formula per lo sviluppo della potenza n -esima di un binomio:

$$(x+y)^n = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k} \quad (49)$$

Per inciso notiamo che, nel caso particolare $x = y = 1$, l'equazione precedente fornisce l'interessante risultato:

$$\sum_{k=0}^n \binom{n}{k} = 2^n \quad (50)$$

3.1.2 La distribuzione binomiale

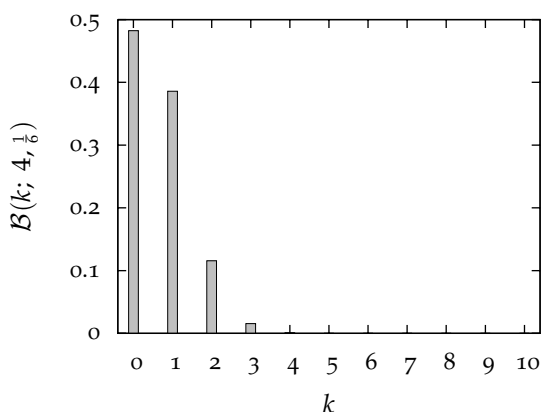
Consideriamo n realizzazioni indipendenti di uno stesso esperimento E che abbia esattamente due esiti distinti possibili, E_1 ed E_2 , con probabilità:

$$\begin{aligned} P(E_1) &= p \\ P(E_2) &= q = 1 - p \end{aligned}$$

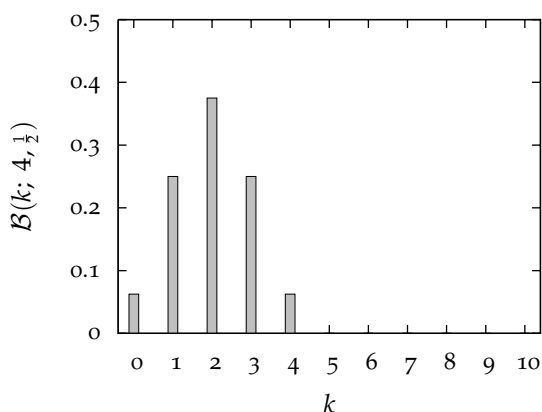
Dato un generico numero intero k con $k \in [0, n]$, la probabilità di ottenere esattamente k volte l'evento E_1 è data dalla distribuzione binomiale

$$\mathcal{B}(k; n, p) = \binom{n}{k} p^k q^{n-k} \quad (51)$$

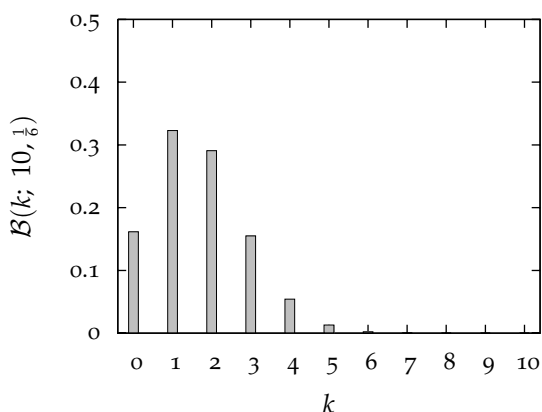
Notiamo esplicitamente che, nella formula, k è la variabile casuale, mentre n e p sono parametri fissati dal particolare problema sotto studio (come di consueto la variabile è separata dai parametri da un punto e virgola); q non è un parametro indipendente in



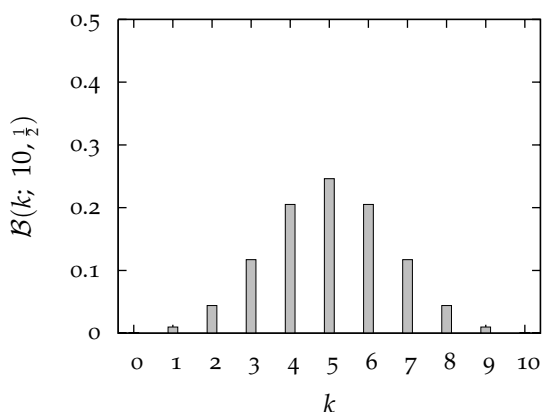
(8.1) Grafico della distribuzione binomiale per $n = 4$ e $p = \frac{1}{6}$.



(8.2) Grafico della distribuzione binomiale per $n = 4$ e $p = \frac{1}{2}$.



(8.3) Grafico della distribuzione binomiale per $n = 10$ e $p = \frac{1}{6}$.



(8.4) Grafico della distribuzione binomiale per $n = 10$ e $p = \frac{1}{2}$.

FIGURA 8.. Grafici della distribuzione binomiale per diversi valori di n e p . Come discuteremo in dettaglio nel seguito, per alti valori di np la distribuzione binomiale tende ad assumere una classica forma a campana.

quanto è univocamente fissato da p . Qui e nel seguito scriviamo q anziché $(1 - p)$ solo per avere formule più compatte.

Supponiamo, per fissare le idee, che l'evento sia il risultato del lancio di una moneta. I modi nei quali l'evento può realizzarsi sono due: testa o croce, che indichiamo qui con T e C . Se la moneta è equa, essi sono equiprobabili:

$$\begin{aligned} P(T) &= p = \frac{1}{2} \\ P(C) &= q = 1 - p = \frac{1}{2} \end{aligned}$$

La probabilità che in n lanci i primi k diano T e gli altri $(n - k)$ diano C è:

$$P \left(\overbrace{TTTTT \dots}^{k \text{ volte}} \overbrace{CCCCC \dots}^{n-k \text{ volte}} \right) = p^k q^{n-k} = \left(\frac{1}{2} \right)^n$$

però avere k teste esattamente in questo ordine è solo uno dei tanti possibili modi nei quali si possono avere k volte T e $(n - k)$ volte C . Potremmo avere, tanto per fare un esempio, C nei primi $(n - k)$ lanci e T nei rimanenti k :

$$P \left(\overbrace{CCCCC \dots}^{n-k \text{ volte}} \overbrace{TTTTT \dots}^{k \text{ volte}} \right) = p^k q^{n-k} = \left(\frac{1}{2} \right)^n$$

insieme a molte altre combinazioni *miste*, tutte con la stessa probabilità $p^k q^{n-k}$. Si tratta dunque di contare queste combinazioni; e il numero totale di modi un cui possiamo *disporre* k teste in n lanci è dato proprio dal numero di combinazioni di n elementi k a k , così come scritto nella (48). Da cui segue immediatamente la distribuzione binomiale (51).

► **ESEMPIO 3.1.** *Supponiamo di lanciare una moneta per 4 volte. Le sequenze possibili sono complessivamente $2^4 = 16$ e sono qui di seguito elencate: TTTT, CTTT, TCTT, CCTT, TTCT, CTCT, TCCT, CCCT, TTTC, CTTC, TCTC, CCTC, TTCC, CTCC, TCCC, CCCC. Se la moneta non è truccata ognuna di queste combinazioni ha la stessa probabilità P di uscire:*

$$P = \left(\frac{1}{2} \right)^4 = \frac{1}{16}$$

Le combinazioni in cui T (testa) figura esattamente due volte sono 6 (per l'esattezza le combinazioni 4, 6, 7, 10, 11 e 13, numerate nell'ordine in cui esse sono riportate sopra). Notiamo esplicitamente che:

$$6 = \binom{4}{2}$$

La probabilità di avere due volte testa in quattro lanci può essere calcolata esplicitamente come il rapporto tra casi favorevoli e casi possibili:

$$P = \frac{6}{16} = \frac{3}{8}$$

Più semplicemente, la (51) si applica direttamente al nostro caso, fornendo (ovviamente) lo stesso risultato:

$$\mathcal{B}(2; 4, \frac{1}{2}) = \binom{4}{2} \cdot \left(\frac{1}{2} \right)^2 \cdot \left(\frac{1}{2} \right)^{4-2} = 6 \cdot \frac{1}{4} \cdot \frac{1}{4} = \frac{6}{16} = \frac{3}{8}$$

□

3.1.3 Normalizzazione, media e varianza

Sfruttando la relazione (49) è semplice dimostrare che la condizione di normalizzazione (33) è soddisfatta:

$$\sum_{k=0}^n \mathcal{B}(k; n, p) = \sum_{k=0}^n \binom{n}{k} p^k q^{n-k} = (p+q)^n = 1^n = 1 \quad (52)$$

La media si calcola secondo la (34):

$$\mu = E[k] = \sum_{k=0}^n k \cdot \mathcal{B}(k; n, p) = \sum_{k=0}^n k \cdot \binom{n}{k} p^k q^{n-k} = \sum_{k=0}^n k \cdot \frac{n!}{k!(n-k)!} \cdot p^k q^{n-k}$$

Il termine con $k = 0$ non contribuisce alla somma perché il k è a moltiplicare, per cui possiamo riscrivere l'espressione precedente facendo partire la somma da 1:

$$\mu = \sum_{k=1}^n k \cdot \frac{n!}{k!(n-k)!} \cdot p^k q^{n-k} = \sum_{k=1}^n k \cdot \frac{np \cdot (n-1)!}{k \cdot (k-1)!(n-k)!} \cdot p^{k-1} q^{n-k}$$

Adesso possiamo portare fuori dal segno di sommatoria il termine np a moltiplicare, dato che non dipende dall'indice k su cui si somma, e semplificare il k che è sia al numeratore che al denominatore:

$$\mu = np \sum_{k=1}^n \frac{(n-1)!}{(k-1)!(n-k)!} \cdot p^{k-1} q^{n-k}$$

Ponendo $h = k - 1$ e $m = n - 1$ si ottiene:

$$\mu = np \sum_{h=0}^m \frac{m!}{h!(m-h)!} \cdot p^h q^{m-h} = np \sum_{h=0}^m \binom{m}{h} p^h q^{m-h}$$

Nell'ultima sommatoria riconosciamo la condizione di normalizzazione (52), per cui concludiamo:

$$\boxed{\mu = np} \quad (53)$$

Per stimare la varianza utilizziamo la (39); il valore di aspettazione di k^2 si calcola come:

$$E[k^2] = \sum_{k=0}^n k^2 \cdot \mathcal{B}(k; n, p) = \sum_{k=0}^n k^2 \cdot \binom{n}{k} p^k q^{n-k} = \sum_{k=0}^n k^2 \cdot \frac{n!}{k!(n-k)!} \cdot p^k q^{n-k}$$

Esattamente come prima possiamo far partire la somma da 1, dato che il primo termine è nullo:

$$\begin{aligned} E[k^2] &= \sum_{k=1}^n k^2 \cdot \frac{n!}{k!(n-k)!} \cdot p^k q^{n-k} = \sum_{k=1}^n k^2 \cdot \frac{np \cdot (n-1)!}{k \cdot (k-1)!(n-k)!} \cdot p^{k-1} q^{n-k} = \\ &= np \sum_{k=1}^n k \cdot \frac{(n-1)!}{(k-1)!(n-k)!} \cdot p^{k-1} q^{n-k} \end{aligned}$$

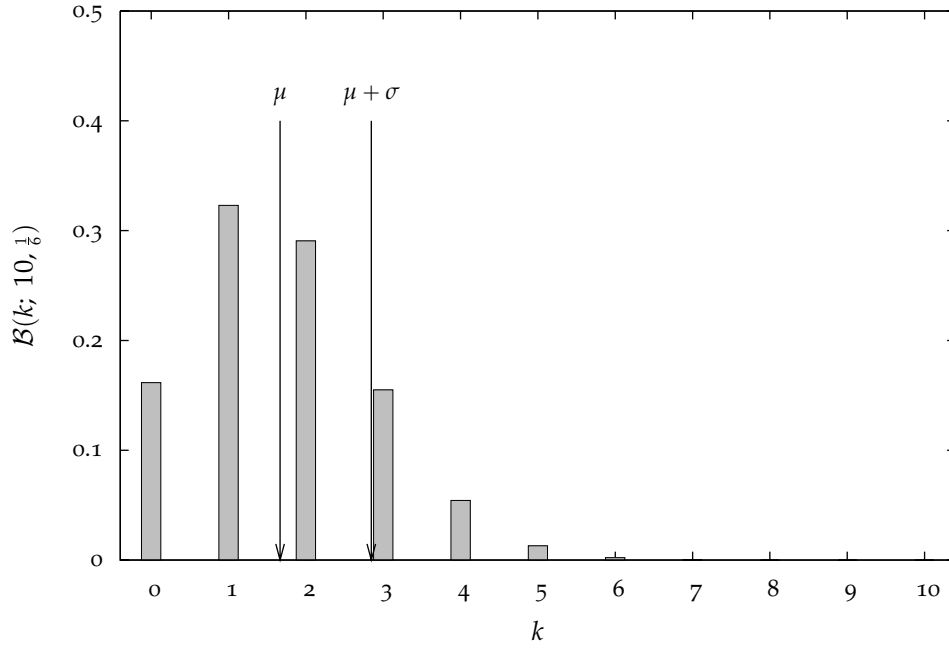


FIGURA 9.. Grafico della distribuzione binomiale per $n = 10$ e $p = \frac{1}{6}$; rappresenta, per fissare le idee, la probabilità che in 10 lanci di un dado un numero dato (ad esempio il 5) esca k volte ($k = 0, \dots, 10$). Per completezza i punti corrispondenti a $x = \mu$ e $x = \mu + \sigma$ sono esplicitamente indicati sul grafico.

E ponendo, come prima, $h = k - 1$ e $m = n - 1$ si ottiene:

$$E[k^2] = np \sum_{h=0}^m (h+1) \cdot \mathcal{B}(h; m, p) = np \left[\sum_{h=0}^m h \cdot \mathcal{B}(h; m, p) + \sum_{h=0}^m \mathcal{B}(h; m, p) \right] = np(mp + 1)$$

Ricordando che $m = n - 1$ si ha alla fine:

$$E[k^2] = np(np - p + 1)$$

che possiamo utilizzare per il calcolo della varianza:

$$\sigma^2 = E[k^2] - \mu^2 = np(np - p + 1) - (np)^2 = np - np^2 = np(1 - p) = npq$$

Concludendo la varianza vale:

$$\sigma^2 = np(1 - p) = npq \quad (54)$$

e la deviazione standard:

$$\sigma = \sqrt{np(1 - p)} = \sqrt{npq} \quad (55)$$

► **ESEMPIO 3.2.** Supponiamo di lanciare un dado per 4 volte. La probabilità che un numero fissato (per esempio il 3) esca 0, 1, ..., 4 volte è dato dalla distribuzione binomiale con $n = 4$, $p = \frac{1}{6}$ e $k = 0, 1, \dots, 4$ rispettivamente. Calcoliamo esplicitamente $\mathcal{B}(k; 4, \frac{1}{6})$ per ogni valore di k (vedi anche figura 8):

$$\begin{aligned} P(0) = \mathcal{B}(0; 4, \tfrac{1}{6}) &= \binom{4}{0} \cdot \left(\tfrac{1}{6}\right)^0 \cdot \left(\tfrac{5}{6}\right)^{4-0} = 1 \cdot 1 \cdot \frac{625}{1296} \approx 0.482 \\ P(1) = \mathcal{B}(1; 4, \tfrac{1}{6}) &= \binom{4}{1} \cdot \left(\tfrac{1}{6}\right)^1 \cdot \left(\tfrac{5}{6}\right)^{4-1} = 4 \cdot \tfrac{1}{6} \cdot \frac{125}{216} \approx 0.386 \\ P(2) = \mathcal{B}(2; 4, \tfrac{1}{6}) &= \binom{4}{2} \cdot \left(\tfrac{1}{6}\right)^2 \cdot \left(\tfrac{5}{6}\right)^{4-2} = 6 \cdot \tfrac{1}{36} \cdot \frac{25}{36} \approx 0.116 \\ P(3) = \mathcal{B}(3; 4, \tfrac{1}{6}) &= \binom{4}{3} \cdot \left(\tfrac{1}{6}\right)^3 \cdot \left(\tfrac{5}{6}\right)^{4-3} = 4 \cdot \tfrac{1}{216} \cdot \frac{5}{6} \approx 1.54 \cdot 10^{-2} \\ P(4) = \mathcal{B}(4; 4, \tfrac{1}{6}) &= \binom{4}{4} \cdot \left(\tfrac{1}{6}\right)^4 \cdot \left(\tfrac{5}{6}\right)^{4-4} = 1 \cdot \tfrac{1}{1296} \cdot 1 \approx 7.72 \cdot 10^{-4} \end{aligned}$$

La media e la deviazione standard di questa distribuzione sono rispettivamente, secondo la (53) e la (55):

$$\begin{aligned} \mu &= np = 4 \cdot \tfrac{1}{6} \approx 0.67, \\ \sigma &= \left(4 \cdot \tfrac{1}{6} \cdot \tfrac{5}{6}\right)^{\frac{1}{2}} \approx 0.74. \end{aligned}$$

□

3.2 DISTRIBUZIONE DI POISSON

3.2.1 Premessa: il numero di Nepero

Definiamo il numero di Nepero e , che ci sarà utile nel seguito, come:

$$e \equiv \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n \approx 2.7182818284590452353602874713526624977572470 \dots$$

Per completezza ricordiamo che si può dimostrare la seguente uguaglianza:

$$\lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n = e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \quad (56)$$

da cui, per $x = 1$, si ottiene una formula grazie alla quale è facile valutare e con il grado di approssimazione desiderato:

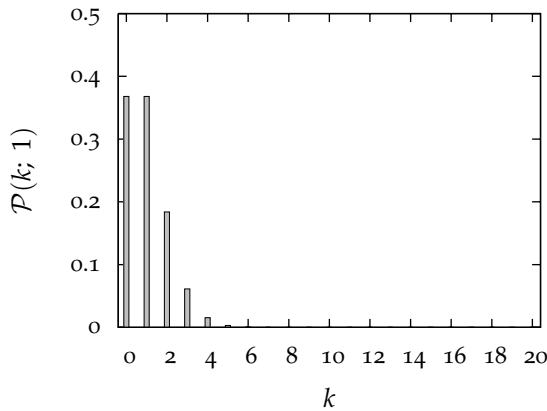
$$e = 1 + 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} + \frac{1}{120} + \frac{1}{720} + \dots$$

3.2.2 La distribuzione di Poisson

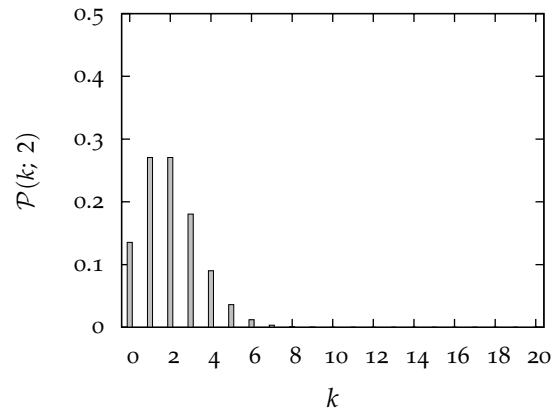
Sotto alcune semplici ipotesi, che verranno chiarite tra breve, la probabilità che si verifichino k eventi in una situazione in cui in media se ne verificano μ è data dalla cosiddetta distribuzione di Poisson:

$$\mathcal{P}(k; \mu) = \frac{\mu^k}{k!} e^{-\mu} \quad (57)$$

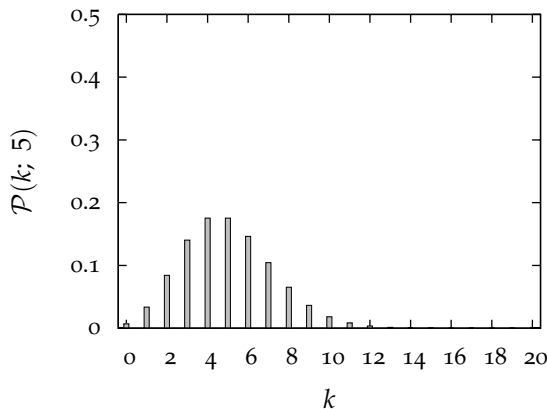
Sottolineiamo di nuovo la notazione utilizzata: k è la variabile casuale e μ è un parametro che, come vedremo, rappresenta la media nel senso della (34).



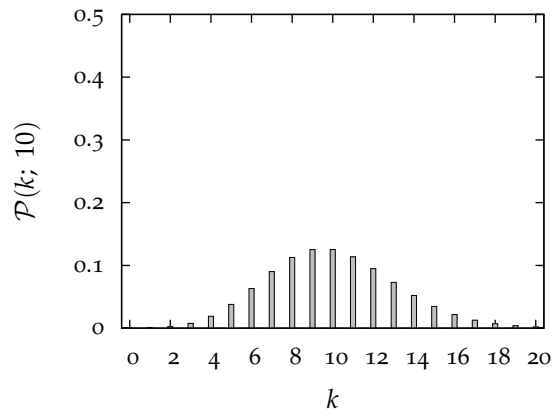
(10.1) Grafico della distribuzione di Poisson per $\mu = 1$.



(10.2) Grafico della distribuzione di Poisson per $\mu = 2$.



(10.3) Grafico della distribuzione di Poisson per $\mu = 5$.



(10.4) Grafico della distribuzione di Poisson per $\mu = 10$.

FIGURA 10.. Grafici della distribuzione di Poisson per diversi valori della media μ . Al crescere di μ la distribuzione di Poisson tende ad assumere via via la classica forma a *campana*, simmetrica rispetto al valor medio.

Formalmente la distribuzione di Poisson si ottiene come limite della distribuzione binomiale nel caso in cui:

$$\begin{aligned} p &\ll 1 \\ n &\gg 1 \\ np &= \mu \text{ (costante)} \end{aligned} \tag{58}$$

cioè quando n tende all'infinito e p tende a zero ma il loro prodotto (che, ricordiamo, rappresenta il valor medio della distribuzione) rimane finito. Infatti:

$$\lim_{n \rightarrow \infty} \mathcal{B}(k; n, p) = \lim_{n \rightarrow \infty} \binom{n}{k} p^k (1-p)^{n-k} = \lim_{n \rightarrow \infty} \binom{n}{k} p^k (1-p)^{n-k}$$

Ricordando che il valor medio della distribuzione binomiale è $\mu = np$, come appare nell'ultima condizione delle (58), possiamo scrivere:

$$p = \frac{\mu}{n}$$

e quindi:

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathcal{B}(k; n, p) &= \lim_{n \rightarrow \infty} \frac{n!}{k!(n-k)!} \cdot \left(\frac{\mu}{n}\right)^k \cdot \left(1 - \frac{\mu}{n}\right)^{n-k} = \\ &= \lim_{n \rightarrow \infty} \frac{n \cdot (n-1) \cdots (n-k+1)}{k!} \cdot \left(\frac{\mu}{n}\right)^k \cdot \left(1 - \frac{\mu}{n}\right)^{n-k} = \\ &= \lim_{n \rightarrow \infty} \frac{n \cdot (n-1) \cdots (n-k+1)}{n^k} \cdot \frac{\mu^k}{k!} \cdot \frac{\left(1 - \frac{\mu}{n}\right)^n}{\left(1 - \frac{\mu}{n}\right)^k} \end{aligned}$$

Notiamo che il primo termine del prodotto dentro il segno di limite:

$$\lim_{n \rightarrow \infty} \overbrace{\frac{n \cdot (n-1) \cdots (n-k+1)}{n^k}}^{k \text{ termini}} = 1$$

tende ad 1 poiché per n grande:

$$(n-1) \simeq (n-2) \simeq \dots \simeq (n-k+1) \simeq n$$

Si ha dunque:

$$\lim_{n \rightarrow \infty} \mathcal{B}(k; n, p) = \lim_{n \rightarrow \infty} \frac{\mu^k}{k!} \cdot \frac{\left(1 - \frac{\mu}{n}\right)^n}{\left(1 - \frac{\mu}{n}\right)^k} = \frac{\mu^k}{k!} e^{-\mu} = \mathcal{P}(k; \mu)$$

Questo significa che ci si deve aspettare una Poissoniana tutte le volte in cui il numero degli eventi possibili è enorme, ma quelli che in media si verificano sono abbastanza pochi.

► **ESEMPIO 3.3.** Supponiamo di lanciare tre dadi per 500 volte; supponiamo anche di registrare, ad ogni lancio, la somma S delle uscite dei tre dadi in questione. Ci chiediamo quale sia la probabilità che il valore $S = 4$ esca per 10 volte.

Il modo formalmente corretto di impostare il problema è quello di utilizzare la statistica binomiale. Con tre dadi a sei facce si hanno $6^3 = 216$ possibili combinazioni in totale; quelle che danno come somma 4 sono esattamente 3: $(1, 1, 2)$, $(1, 2, 1)$ e $(2, 1, 1)$. La probabilità che la somma delle uscite in un lancio di tre dadi sia 4 è dunque $p = \frac{3}{216}$. A questo punto è facile calcolare la probabilità che questo accada 10 volte in 500 lanci:

$$P = \mathcal{B}(10; 500, \frac{3}{216}) = \binom{500}{10} \cdot \left(\frac{3}{216}\right)^{10} \cdot \left(\frac{213}{216}\right)^{490} \simeq 0.06933.$$

Ma alla luce di quanto detto in questo paragrafo possiamo anche applicare la statistica Poissoniana (n è grande e p è piccolo). Su 500 lanci S ammonterà a 4, in media, un numero di volte pari a:

$$\mu = 500 \cdot \frac{3}{216} = \frac{125}{18} \simeq 6.94$$

e la probabilità di avere 10 volte il valore 4 sarà dunque:

$$P = \mathcal{P}(10; \frac{125}{18}) = \frac{6.94^{10}}{10!} \cdot e^{-6.94} \simeq 0.06929$$

che è in ottimo accordo con quanto trovato prima. □

Se l'applicabilità della distribuzione di Poisson fosse limitata a situazioni simili all'esempio 3.3, la sua utilità sarebbe piuttosto modesta, non vi è dubbio. Pur tuttavia, ancor prima di andare avanti, vale la pena di notare come, in questa formulazione, non sia necessario conoscere a priori né il numero n di eventi possibili né la probabilità p di un singolo evento, ma solo il numero medio di eventi osservati np o una sua stima. E in molte situazioni è proprio questo ciò che si misura direttamente in laboratorio.

► **ESEMPIO 3.4.** Nelle misure di radiometria si ha tipicamente a che fare con un campione di materiale radioattivo ed un apparato sperimentale che permette di rivelare le particelle o la radiazione prodotte dai processi di decadimento. In una situazione di questo tipo (come vedremo in seguito si tratta, sotto certe ipotesi, di un tipico processo Poissoniano) ciò che si misura direttamente è proprio il numero medio di decadimenti per unità di tempo e non il numero di atomi del campione o la probabilità di decadimento del singolo nucleo. □

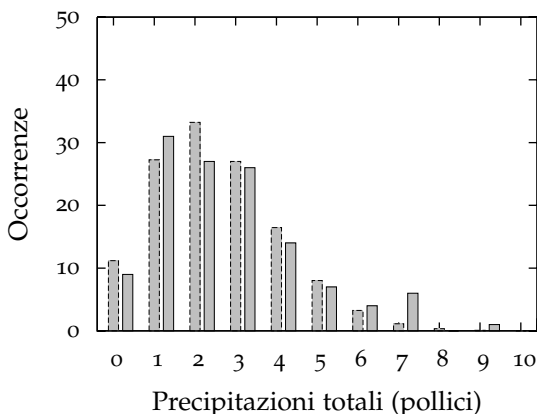
Come accennato prima, la distribuzione di Poisson non si ottiene solamente come limite di una binomiale. E' possibile dimostrare [1] che sono governati dalla statistica di Poisson conteggi casuali che soddisfino le seguenti proprietà:

1. il verificarsi di un evento non è influenzato in nessun modo dal verificarsi o meno degli altri eventi (*indipendenza degli eventi*)
2. per piccoli intervalli di tempo la possibilità che si verifichino k eventi è proporzionale al tempo (*regolarità*),

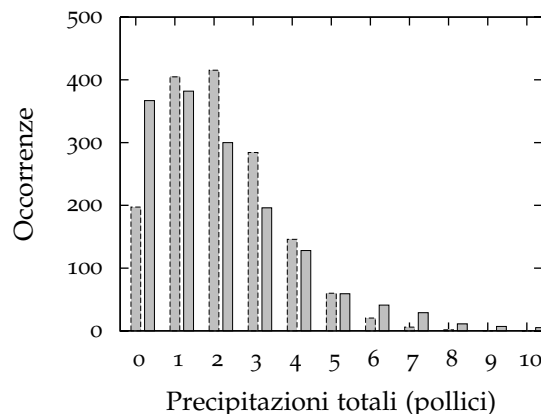
3. la probabilità di contare un certo numero di eventi in un intervallo di tempo τ non dipende dall'istante in cui si effettua il conteggio, ma solo da τ (*stazionarietà*),

Le ipotesi necessarie alla dimostrazione sono così generali che in effetti la statistica di Poisson ha un campo di applicabilità estremamente vasto [10]. Tanto per fare qualche esempio: il numero di persone che passano per una determinata soglia in un intervallo di tempo fissato (per esempio il numero di persone che entrano in una biblioteca in 10 minuti) o il numero di un particolare tipo di automobili che passano per un certo luogo in un intervallo di tempo fissato, o ancora il numero di particelle di un campione radioattivo che decadono in un intervallo di tempo fissato sono tipici esempi di conteggi Poissoniani.

► **ESEMPIO 3.5.** *L'esempio seguente mette in luce l'importanza della stazionarietà nei fenomeni Poissoniani. In figura 11 sono riportati i dati relativi alle precipitazioni mensili registrate da una stazione meteorologica Statunitense¹ nel periodo compreso tra il 1871 ed il 1998. Possiamo chiederci se il livello di precipitazioni mensile sia ben descritto da una distribuzione di Poisson; operativamente ciò che dobbiamo fare è confrontare la distribuzione in frequenza dei dati con le occorrenze attese assumendo una distribuzione di Poisson con media μ pari al livello medio di precipitazioni misurato.*



(11.1) Istogramma delle precipitazioni registrate nel mese di gennaio negli anni 1871 ed il 1998 (128 misure in totale).



(11.2) Istogramma delle precipitazioni mensili (senza alcuna selezione sul mese) per gli anni 1871 ed il 1998 ($128 \cdot 12 = 1536$ misure in totale).

FIGURA 11.. Distribuzione delle precipitazioni mensili (in pollici) registrate dalla stazione meteorologica di Wendover, Utah (Stati Uniti) nel periodo compreso tra il gennaio 1871 ed il 1998. Nel grafico 11.2 sono riportate tutte le $128 \cdot 12 = 1536$ misure, mentre nel grafico 11.1 sono riportate solamente quelle relative ai mesi di gennaio. L'istogramma tratteggiato rappresenta, in entrambi i casi, la distribuzione di Poisson che meglio si adatta ai dati.

¹ I dati sono disponibili alla pagina web <http://www.sci.usq.edu.au/staff/dunn/Datasets/index.html> insieme a numerosi altri *set* di dati potenzialmente interessanti.

Ciò che impariamo da questo semplice esercizio è che (nel caso in questione) la distribuzione delle misure è molto lontana da una Poissoniana se includiamo tutte le osservazioni (figura 11.2) ma è invece sorprendentemente vicina ad una Poissoniana se ci limitiamo ai dati registrati nel mese di gennaio (figura 11.1). È chiaro che siamo di fronte ad un fenomeno che è ragionevolmente stazionario su tempi scala dell'ordine del mese ma non su tempi più lunghi (d'altra parte è ragionevole aspettarsi che le variazioni stagionali siano significative). Semplificando brutalmente possiamo pensare che le distribuzioni delle precipitazioni, limitate a ciascuno dei 12 mesi dell'anno, siano descritte (più o meno) bene da distribuzioni Poissoniane; il punto è che queste 12 distribuzioni avranno medie diverse tra di loro (a causa delle variazioni stagionali) e questo fa sì che la distribuzione integrata sull'intero anno non sia Poissoniana. \square

3.2.3 Normalizzazione, media e varianza

Verifichiamo che la distribuzione (57) è normalizzata:

$$\sum_{k=0}^{\infty} \mathcal{P}(k; \mu) = \sum_{k=0}^{\infty} \frac{\mu^k}{k!} e^{-\mu} = e^{-\mu} \cdot \sum_{k=0}^{\infty} \frac{\mu^k}{k!} = e^{-\mu} \cdot e^{\mu} = 1$$

dove si è fatto uso, nel penultimo passaggio, della (56).

Calcoliamo il valore di aspettazione di k sulla distribuzione, cioè la media, secondo la (34):

$$E[k] = \sum_{k=0}^{\infty} k \cdot \mathcal{P}(k; \mu) = \sum_{k=0}^{\infty} k \cdot \frac{\mu^k}{k!} e^{-\mu} = e^{-\mu} \cdot \sum_{k=0}^{\infty} k \cdot \frac{\mu^k}{k!}$$

Al solito il termine con $k = 0$ non contribuisce alla somma, per cui possiamo far iniziare la somma stessa da $k = 1$:

$$E[k] = e^{-\mu} \cdot \sum_{k=1}^{\infty} k \cdot \frac{\mu^k}{k!} = e^{-\mu} \cdot \sum_{k=1}^{\infty} \frac{\mu^k}{(k-1)!}$$

Ponendo $h = k - 1$ possiamo trasformare l'espressione come segue:

$$E[k] = e^{-\mu} \cdot \sum_{h=0}^{\infty} \frac{\mu^{h+1}}{h!} = e^{-\mu} \cdot \sum_{h=0}^{\infty} \mu \cdot \frac{\mu^h}{h!} = \mu e^{-\mu} \cdot \sum_{h=0}^{\infty} \frac{\mu^h}{h!} = \mu e^{-\mu} \cdot e^{\mu} = \mu$$

Dunque il parametro μ ha proprio il significato della media nel senso della definizione (34), come avevamo anticipato.

Per il calcolo della varianza partiamo, esattamente come abbiamo fatto nel caso della distribuzione binomiale, dal valore di aspettazione di k^2 :

$$E[k^2] = \sum_{k=0}^{\infty} k^2 \cdot \mathcal{P}(k; \mu) = \sum_{k=0}^{\infty} k^2 \cdot \frac{\mu^k}{k!} e^{-\mu}$$

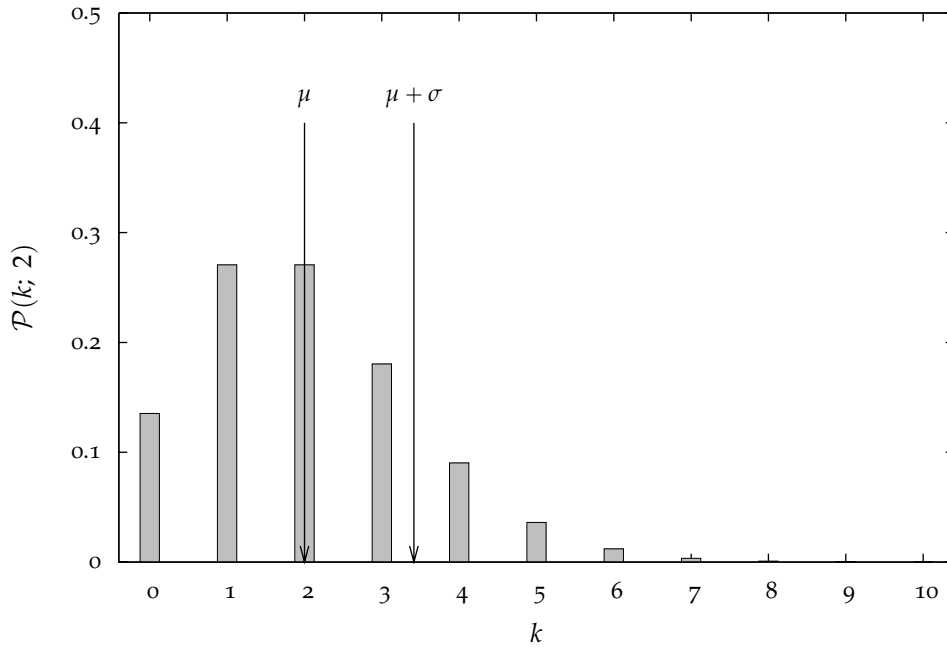


FIGURA 12.. Grafico della distribuzione di Poisson per $\mu = 2$. Per completezza i punti corrispondenti a $x = \mu$ e $x = \mu + \sigma$ sono esplicitamente indicati sul grafico.

Al solito eliminiamo il termine con $k = 0$ e operiamo il cambiamento di variabile $h = k - 1$:

$$\begin{aligned}
 E[k^2] &= \sum_{k=1}^{\infty} k^2 \cdot \frac{\mu^k}{k!} e^{-\mu} = \sum_{k=1}^{\infty} k\mu \cdot \frac{\mu^{k-1}}{(k-1)!} e^{-\mu} = \mu \cdot \sum_{h=0}^{\infty} (h+1) \cdot \frac{\mu^h}{h!} e^{-\mu} = \\
 &= \mu \left[\sum_{h=0}^{\infty} h \cdot \mathcal{P}(h; \mu) + \sum_{h=0}^{\infty} \mathcal{P}(h; \mu) \right] = \mu(\mu + 1) = \mu^2 + \mu
 \end{aligned}$$

da cui:

$$\sigma^2 = E[k^2] - \mu^2 = \mu^2 + \mu - \mu^2 = \mu$$

In conclusione abbiamo l'espressione per la varianza:

$$\boxed{\sigma^2 = \mu} \quad (59)$$

e per la deviazione standard:

$$\boxed{\sigma = \sqrt{\mu}} \quad (60)$$

Questo dice che la larghezza della distribuzione di Poisson è pari alla radice quadrata del valor medio.

► **ESEMPIO 3.6.** Supponiamo per un certo campione di una sostanza radioattiva si abbiano in media 10 decadimenti al secondo. Facciamo l'ulteriore assunzione che questo numero non

cambi significativamente nel corso del nostro esperimento. Ci chiediamo quale sia la probabilità di avere meno di tre decadimenti in un intervallo di 0.5 secondi.

Il numero medio di decadimenti nel nostro intervallo di tempo è $\mu = 0.5 \cdot 10 = 5$. La probabilità di avere 0, 1, 2 decadimenti è, rispettivamente:

$$\begin{aligned} P(0) = \mathcal{P}(0; 5) &= \frac{5^0}{0!} e^{-5} \simeq 0.0067 \\ P(1) = \mathcal{P}(1; 5) &= \frac{5^1}{1!} e^{-5} \simeq 0.0337 \\ P(2) = \mathcal{P}(2; 5) &= \frac{5^2}{2!} e^{-5} \simeq 0.0842, \end{aligned}$$

per cui la probabilità cercata è la somma di queste tre:

$$P = P(0) + P(1) + P(2) \simeq 12\%$$

Notiamo che l'assunzione che il numero di decadimenti al secondo non cambi nel tempo (assunzione per altro tipicamente verificata nelle misure che si eseguono normalmente in laboratorio con sostanze radioattive) garantisce la stazionarietà del processo che, come detto prima, è fondamentale per la derivazione formale della distribuzione di Poisson. In generale il numero di atomi radioattivi contenuti in un certo campione decresce esponenzialmente con il tempo

$$N(t) = N_0 e^{-\lambda t}$$

e con esso anche il numero medio di decadimenti per unità di tempo, per cui su intervalli molto lunghi il processo non è Poissoniano. Per completezza: l'inverso del parametro λ prende il nome di vita media del nucleo radioattivo (si può facilmente dimostrare che dopo un tempo $\frac{1}{\lambda}$ il numero di atomi non ancora decaduti si è ridotto di un fattore $\frac{1}{e} \simeq 0.368$) e la nostra condizione di stazionarietà può essere espressa quantitativamente richiedendo che la durata del nostro esperimento sia molto più breve della vita media del campione. \square

3.3 DISTRIBUZIONE UNIFORME

È il più semplice esempio di funzione di distribuzione di variabile casuale continua (per cui essa rappresenta una densità di probabilità):

$$u(x; a, b) = \begin{cases} \frac{1}{(b-a)} & a \leq x \leq b \\ 0 & x < a; x > b \end{cases} \quad (61)$$

3.3.1 Normalizzazione, media e varianza

La distribuzione, così come è scritta nella (61), è correttamente normalizzata:

$$\int_{-\infty}^{\infty} u(x; a, b) dx = \int_a^b \frac{1}{(b-a)} dx = \frac{1}{(b-a)} \int_a^b dx = \frac{1}{(b-a)} \cdot (b-a) = 1$$

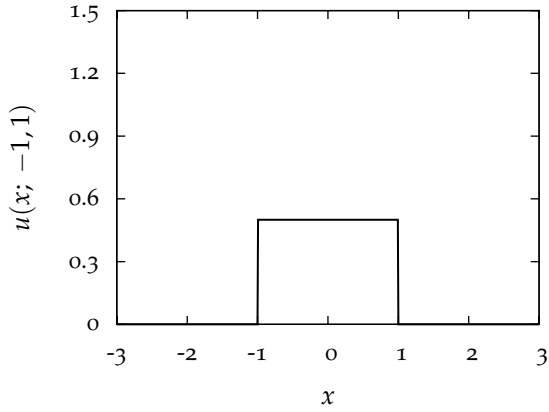
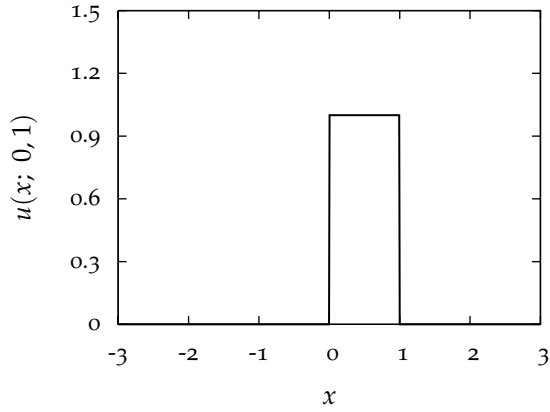
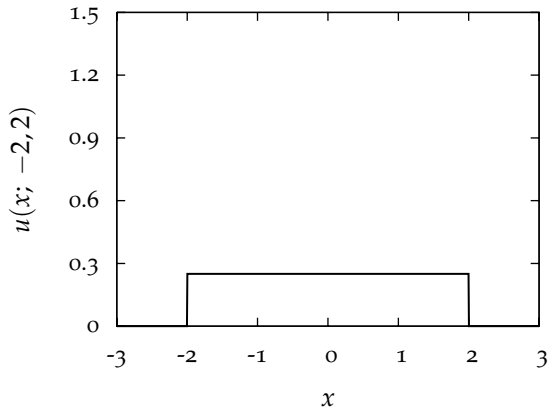
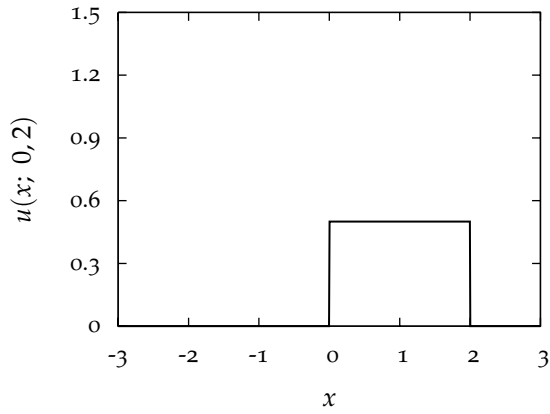
(13.1) Grafico della distribuzione uniforme per $a = -1$ e $b = 1$.(13.2) Grafico della distribuzione uniforme per $a = 0$ e $b = 1$.(13.3) Grafico della distribuzione uniforme per $a = -2$ e $b = 2$.(13.4) Grafico della distribuzione uniforme per $a = 0$ e $b = 2$.

FIGURA 13.. Grafici della distribuzione uniforme per diversi valori di a e b . Dovendo valere la condizione di normalizzazione, il massimo della funzione di distribuzione è tanto più piccolo quanto più grande è la lunghezza dell'intervallo $[a, b]$ su cui essa è diversa da zero.

La media è data da:

$$\begin{aligned}\mu &= E[x] = \int_{-\infty}^{\infty} x \cdot u(x; a, b) dx = \int_a^b \frac{x}{(b-a)} dx = \frac{1}{(b-a)} \int_a^b x dx = \frac{1}{(b-a)} \left[\frac{x^2}{2} \right]_a^b = \\ &= \frac{1}{(b-a)} \cdot \frac{(b^2 - a^2)}{2} = \frac{1}{(b-a)} \cdot \frac{(b-a) \cdot (b+a)}{2} = \frac{(b+a)}{2}\end{aligned}$$

cioè coincide, come è naturale aspettarsi, con il valor medio dell'intervallo su cui la densità di probabilità è non nulla:

$$\boxed{\mu = \frac{(b+a)}{2}} \quad (62)$$

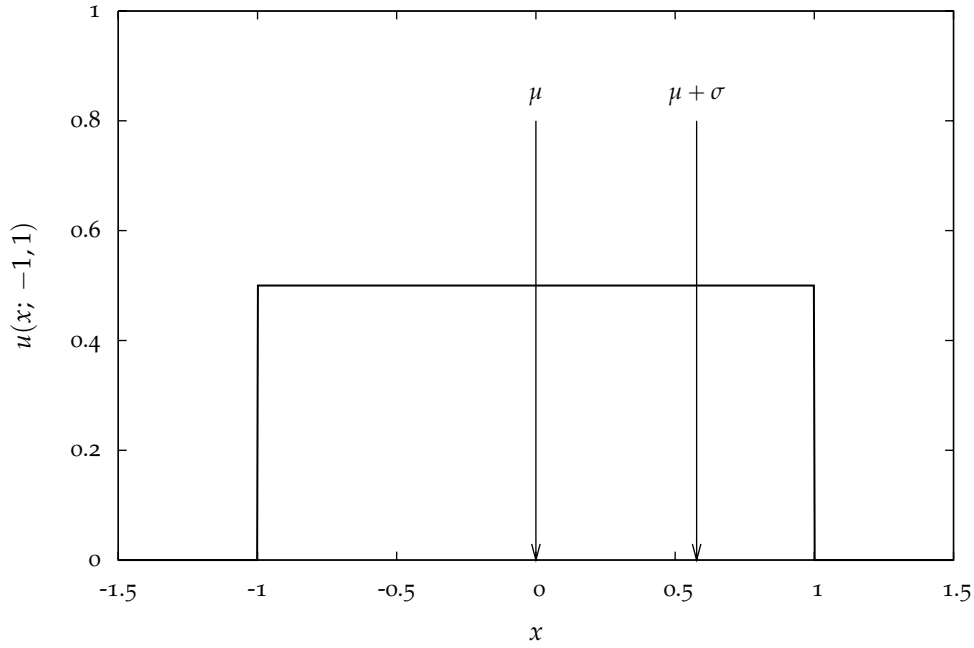


FIGURA 14.. Grafico della distribuzione uniforme per $a = -1$, $b = 1$. Per completezza i punti corrispondenti a $x = \mu$ e $x = \mu + \sigma$ sono esplicitamente indicati sul grafico.

Il valore di aspettazione di x^2 è dato da:

$$\begin{aligned} E[x^2] &= \int_{-\infty}^{\infty} x^2 \cdot u(x; a, b) dx = \int_a^b \frac{x^2}{(b-a)} dx = \frac{1}{(b-a)} \int_a^b x^2 dx = \frac{1}{(b-a)} \left[\frac{x^3}{3} \right]_a^b = \\ &= \frac{1}{(b-a)} \cdot \frac{(b^3 - a^3)}{3} = \frac{1}{(b-a)} \cdot \frac{(b-a) \cdot (b^2 + ab + a^2)}{3} = \frac{(b^2 + ab + a^2)}{3} \end{aligned}$$

e la varianza, al solito:

$$\begin{aligned} \sigma^2 &= E[x^2] - \mu^2 = \frac{(b^2 + ab + a^2)}{3} - \frac{(b+a)^2}{4} = \frac{(4b^2 + 4ab + 4a^2 - 3b^2 - 6ab - 3a^2)}{12} = \\ &= \frac{(b^2 - 2ab + a^2)}{12} = \frac{(b-a)^2}{12} \end{aligned}$$

Per concludere:

$$\boxed{\sigma^2 = \frac{(b-a)^2}{12}} \quad (63)$$

e ancora:

$$\boxed{\sigma = \frac{(b-a)}{\sqrt{12}}} \quad (64)$$

► **ESEMPIO 3.7.** Supponiamo di misurare la massa di un oggetto con una bilancia digitale con la risoluzione di un grammo; sia $m = 58$ g il valore indicato dal display. Se possiamo escludere

la presenza di effetti sistematici è ragionevole ammettere che il misurando sia compreso, con densità di probabilità uniforme, tra 57.5 e 58.5 g. La media della distribuzione sarà 58 g e la deviazione standard $\frac{1}{\sqrt{12}} = 0.289$ g. Se vogliamo attribuire un errore statisticamente corretto alla nostra misura scriveremo, al livello di una deviazione standard:

$$m = 58.0 \pm 0.3 \text{ g}$$

□

3.4 DISTRIBUZIONE ESPONENZIALE

Una variabile casuale la cui funzione di distribuzione sia data da:

$$\varepsilon(x; \lambda) = \begin{cases} \lambda e^{-\lambda x} & 0 \leq x < \infty \\ 0 & x < 0 \end{cases} \quad (65)$$

(con $\lambda \geq 0$) viene chiamata variabile casuale esponenziale con parametro λ (oppure anche esponenzialmente distribuita con parametro λ). Vale la pena di notare che, per ragioni dimensionali, il parametro λ ha le dimensioni di x^{-1} .

La probabilità che il valore della variabile casuale sia compreso in un intervallo fissato può essere calcolato molto facilmente:

$$P(x_1 \leq x \leq x_2) = \int_{x_1}^{x_2} \varepsilon(x; \lambda) dx = e^{-\lambda x_1} - e^{-\lambda x_2} \quad (66)$$

con $x_1 > 0$ e $x_2 > x_1$. Come caso particolare della (66) si può calcolare:

$$P(x \leq x_0) = 1 - e^{-\lambda x_0} \quad (67)$$

che prende il nome di *funzione cumulativa* della distribuzione esponenziale.

Un risultato particolarmente interessante (che non dimostreremo qui) è costituito dal fatto che la funzione di distribuzione esponenziale è la funzione di distribuzione dell'intervallo (di solito temporale) che intercorre tra due conteggi successivi in processi Poissoniani. Tipici esempi di variabili casuali distribuite esponenzialmente sono: il tempo dopo il quale un nucleo radioattivo decade (in questo caso l'inverso del parametro λ prende il nome di *vita media* del nucleo), il tempo che intercorre tra due successive chiamate di telefono sbagliate, ma anche la distanza che una particella percorre in un mezzo omogeneo prima di interagire con il mezzo stesso (in questo caso l'inverso del parametro λ prende il nome di *lunghezza di interazione* o *cammino libero medio*).

Sfruttando la 67 possiamo dimostrare un'altra interessante proprietà della distribuzione esponenziale; scriviamo:

$$P(x \geq x_1 + x_2) = 1 - P(x \leq x_1 + x_2) = e^{-\lambda(x_1+x_2)} = e^{-\lambda x_1} \cdot e^{-\lambda x_2}$$

Per cui si ha:

$$P(x \geq x_1 + x_2) = P(x \geq x_1) \cdot P(x \geq x_2) \quad (68)$$

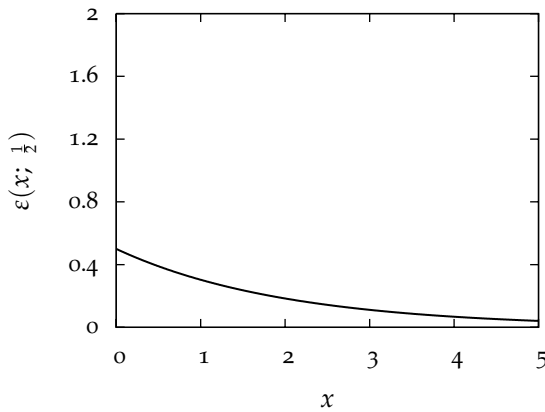
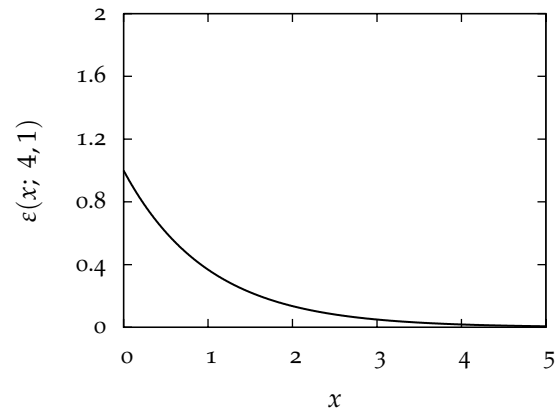
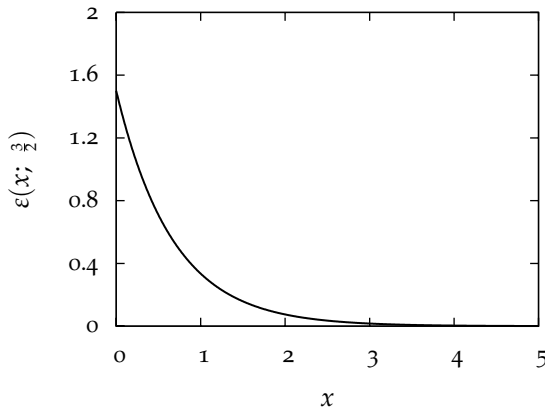
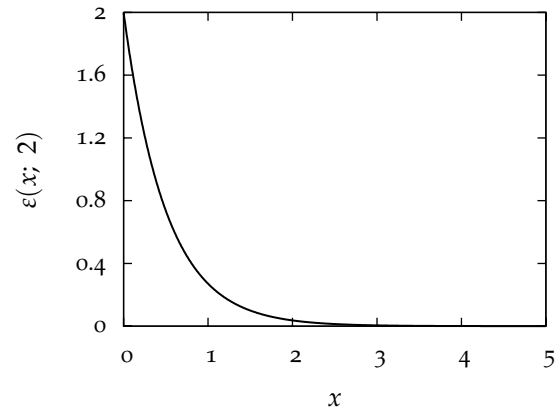
(15.1) Grafico della distribuzione esponenziale per $\lambda = 0.5$.(15.2) Grafico della distribuzione esponenziale per $\lambda = 1$.(15.3) Grafico della distribuzione esponenziale per $\lambda = 1.5$.(15.4) Grafico della distribuzione esponenziale per $\lambda = 2$.

FIGURA 15.. Grafici della funzione $\varepsilon(x; \lambda)$ per diversi valori di λ . Dovendo valere la condizione di normalizzazione, il valore massimo della funzione di distribuzione (in questo caso per $x = 0$) aumenta all'aumentare del parametro λ .

Una variabile casuale che goda di questa proprietà si dice una variabile *senza memoria* (o anche *memory-less*). La condizione di assenza di memoria si può scrivere anche come:

$$P(x \geq x_1 + x_2 | x \geq x_1) = P(x \geq x_2) \quad (69)$$

► **ESEMPIO 3.8.** Supponiamo di osservare, a partire da un istante t_0 , un nucleo radioattivo con vita media (inverso del parametro λ) di 1 s. La probabilità che, dopo un secondo, il nucleo non sia decaduto è:

$$P(x \geq 1) = e^{-1} \simeq 37\%$$

Supponiamo adesso che il nucleo non sia ancora decaduto dopo 10 s, il che è estremamente improbabile ma perfettamente possibile—a proposito: quanto vale la probabilità di questo evento?

Ebbene: la probabilità che il nucleo non decada tra $t_0 + 10$ s e $t_0 + 11$ s è di nuovo il 37%. Tanto per fissare le idee: se la stessa cosa valesse per un'automobile, ad ogni istante una vettura appena uscita dal concessionario ed una con 200000 km alle spalle avrebbero la stessa probabilità di rompersi entro il giorno successivo... Purtroppo, al contrario dei nuclei radioattivi le auto hanno memoria della propria storia! \square

3.4.1 Normalizzazione, media e varianza

La distribuzione, al solito, è normalizzata:

$$\int_0^{\infty} \varepsilon(x; \lambda) dx = \int_0^{\infty} \lambda e^{-\lambda x} dx = \left[-e^{-\lambda x} \right]_0^{\infty} = 1$$

(nel penultimo passaggio si è operato il cambiamento di variabile $t = \lambda x$).

La media della distribuzione, scritta nella forma (65), vale:

$$\mu = E[x] = \int_0^{\infty} x \cdot \varepsilon(x; \lambda) dx = \int_0^{\infty} x \lambda e^{-\lambda x} dx$$

Con il cambiamento di variabile $\lambda x = t$, l'integrale diviene:

$$\mu = \frac{1}{\lambda} \int_0^{\infty} t e^{-t} dt$$

che può essere calcolato per parti:

$$\mu = \frac{1}{\lambda} \left(\left[-t e^{-t} \right]_0^{\infty} + \int_0^{\infty} e^{-t} dt \right) = \frac{1}{\lambda} \left[-e^{-t} \right]_0^{\infty} = \frac{1}{\lambda}$$

Si ha dunque:

$$\boxed{\mu = \frac{1}{\lambda}} \quad (70)$$

Per la varianza, al solito, si parte dal valore di aspettazione di x^2 :

$$E[x^2] = \int_0^{\infty} x^2 \cdot \varepsilon(x; \lambda) dx = \int_0^{\infty} x^2 \lambda e^{-\lambda x} dx$$

Con lo stesso cambiamento di variabile ed integrando per parti due volte si ha:

$$E[x^2] = \frac{1}{\lambda^2} \int_0^{\infty} t^2 e^{-t} dt = \frac{1}{\lambda^2} \left(\left[-t^2 e^{-t} \right]_0^{\infty} + 2 \int_0^{\infty} t e^{-t} dt \right) = \frac{2}{\lambda^2}$$

e, di conseguenza:

$$\sigma^2 = E[x^2] - \mu^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

Riassumendo:

$$\boxed{\sigma^2 = \frac{1}{\lambda^2}} \quad (71)$$

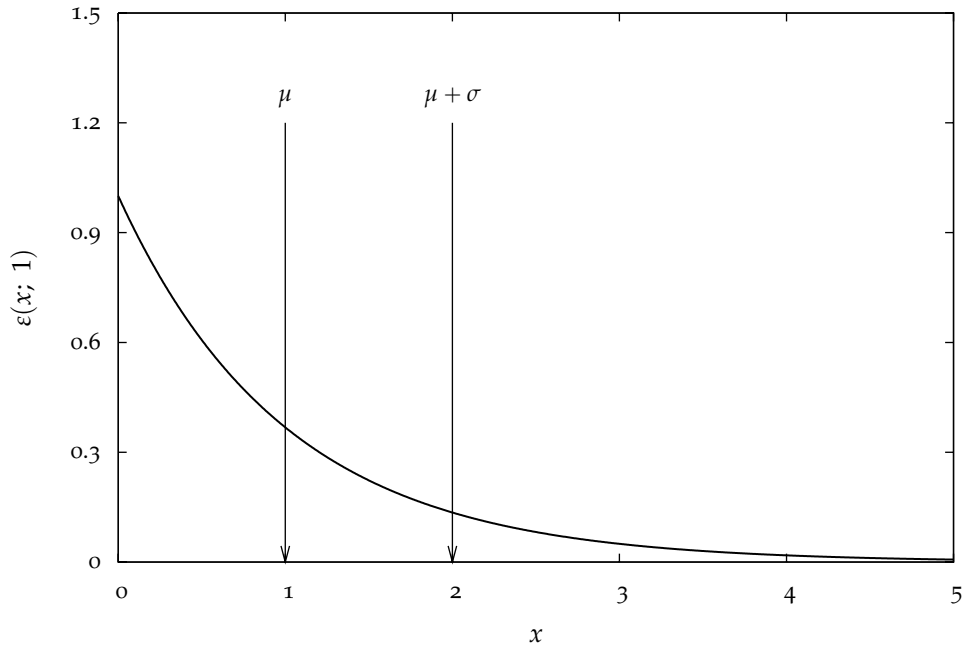


FIGURA 16.. Grafico della funzione $\varepsilon(x; \lambda)$ per $\lambda = 1$. Per completezza i punti corrispondenti a $x = \mu$ e $x = \mu + \sigma$ sono esplicitamente indicati sul grafico.

e ancora:

$$\sigma = \frac{1}{\lambda}$$

(72)

► **ESEMPIO 3.9.** Supponiamo che il cammino libero medio l_0 di una particella in un certo mezzo (omogeneo) sia 1 mm. Quando la particella penetra nel mezzo, in generale percorrerà una certa distanza x (che, a parità di condizioni, varia in modo casuale di volta in volta) prima di interagire; come detto prima, questa distanza è descritta da una densità di probabilità del tipo:

$$p(x) = \frac{1}{l_0} e^{-\frac{x}{l_0}}$$

che è proprio una distribuzione esponenziale con parametro $\lambda = \frac{1}{l_0}$. Come si evince facilmente dalla (70), il valore di aspettazione della variabile x è l_0 . Il che significa, in altre parole, che la nostra particella percorre in media una distanza l_0 prima di interagire (da cui il nome di cammino libero medio). Ci chiediamo quale sia la probabilità che la particella percorra una distanza l di almeno 1 cm nel mezzo prima di interagire col mezzo stesso.

La probabilità cercata si calcola immediatamente sfruttando la funzione cumulativa (67) trovata prima:

$$P(x \geq l) = 1 - P(x \leq l) = 1 - (1 - e^{-\lambda l}) = e^{-\lambda l} = e^{-\frac{l}{l_0}} = e^{-10} \approx 4.54 \cdot 10^{-5}$$

Cioè è estremamente poco probabile che una particella percorra una distanza maggiore a 10 volte il cammino libero medio prima di interagire. \square

3.5 DISTRIBUZIONE DI GAUSS

3.5.1 Premessa: la funzione di Gauss

Come premessa matematica allo studio della distribuzione di Gauss elenchiamo di seguito alcune proprietà rilevanti della funzione di Gauss:

$$f(x) = e^{-x^2} \quad (73)$$

che ci saranno utili in seguito.

(1) $f(x)$ è definita su tutta la retta reale, cioè per $x \in (-\infty, \infty)$.

(2) $f(x)$ è pari, il che si esprime formalmente scrivendo:

$$f(x) = f(-x)$$

Graficamente questo significa che la funzione è simmetrica per riflessione rispetto all'asse $x = 0$.

(3) $f(x)$ tende a zero per grandi valori di x :

$$\lim_{x \rightarrow \pm\infty} f(x) = 0$$

(4) Il punto $x = 0$ è un punto di massimo, nel senso che:

$$f(x) \leq f(0) = 1 \quad \forall x \in (-\infty, \infty)$$

(5) La derivata prima della funzione è:

$$\frac{df(x)}{dx} = -2xe^{-x^2}$$

che vale o solamente per $x = 0$, per cui non ci sono altri punti di massimo o minimo. Inoltre:

$$\frac{df(x)}{dx} \begin{cases} > 0 & \forall x < 0 \\ < 0 & \forall x > 0 \end{cases}$$

quindi $f(x)$ è crescente per $x < 0$ e decrescente per $x > 0$.

(6) $f(x)$ ha la tipica forma di una distribuzione a campana. La semilarghezza a metà altezza si calcola attraverso il formalismo introdotto nel paragrafo 2.5.6. Imponendo:

$$f(x_{\pm}) = \frac{1}{2}f(0) = \frac{1}{2}$$

possiamo scrivere:

$$e^{-x_{\pm}^2} = \frac{1}{2}$$

da cui:

$$x_{\pm}^2 = \pm \ln 2$$

ed infine:

$$\text{HWHM} = \frac{x_+ - x_-}{2} = \sqrt{\ln 2} \approx 0.83 \quad (74)$$

(7) $f(x)$ decresce in modo estremamente rapido al crescere di x ; ad esempio:

$$f(3) = e^{-9} \approx 1.234 \cdot 10^{-4}$$

(8) L'integrale su tutto l'asse reale della funzione vale:

$$I = \int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi} \quad (75)$$

Infatti:


$$I^2 = \int_{-\infty}^{\infty} e^{-x^2} dx \cdot \int_{-\infty}^{\infty} e^{-y^2} dy = \int_{-\infty}^{\infty} dx \int_{-\infty}^{\infty} e^{-(x^2+y^2)} dy$$

che è un integrale doppio su tutto il piano xy . Passando in coordinate polari l'elemento di superficie diviene

$$ds = r dr d\phi$$

e l'integrale si scrive come:

$$I^2 = \int_0^{2\pi} d\phi \int_0^{\infty} r e^{-r^2} dr = 2\pi \left[-\frac{e^{-r^2}}{2} \right]_0^{\infty} = \pi$$

 La funzione di Gauss, nella forma (73) non è una funzione di distribuzione in quanto, come si evince dalla (75), essa non è normalizzata. Vedremo tra breve l'espressione completa (cioè correttamente normalizzata) per la distribuzione di Gauss.

3.5.2 La distribuzione di Gauss

La funzione di distribuzione di Gauss si scrive come:

$$g(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad (76)$$

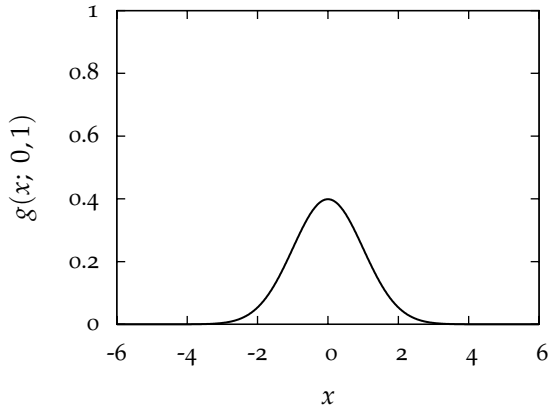
dove μ e σ sono due costanti di cui si chiarirà in seguito il significato.

Proprio come la distribuzione uniforme e quella esponenziale, si tratta di una funzione di distribuzione continua, cioè la variabile casuale è una variabile continua e $g(x; \mu, \sigma)$ rappresenta la densità di probabilità. Assomiglia molto alla funzione e^{-x^2} (studiata nel paragrafo precedente), cui è legata dal cambiamento di variabile:

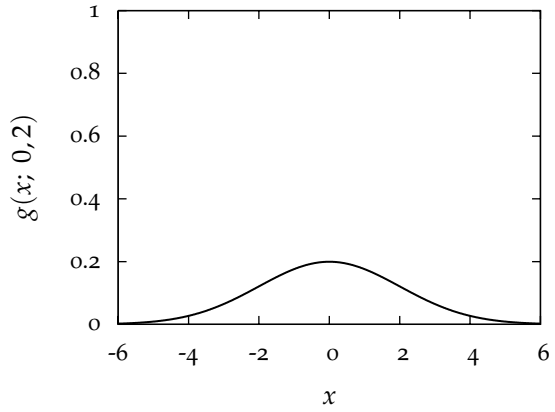
$$t = \frac{1}{\sqrt{2}} \frac{x - \mu}{\sigma} \quad (77)$$

attraverso il quale la (76) diviene appunto:

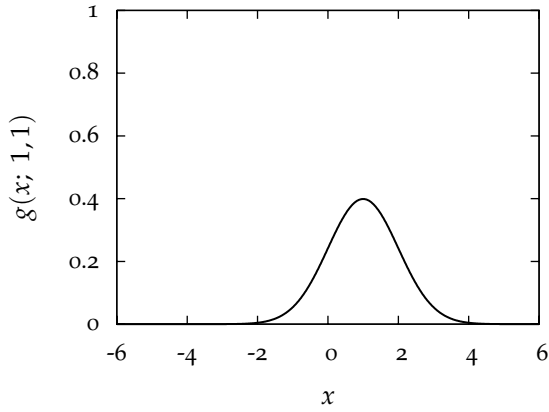
$$g(t) \propto e^{-t^2}$$



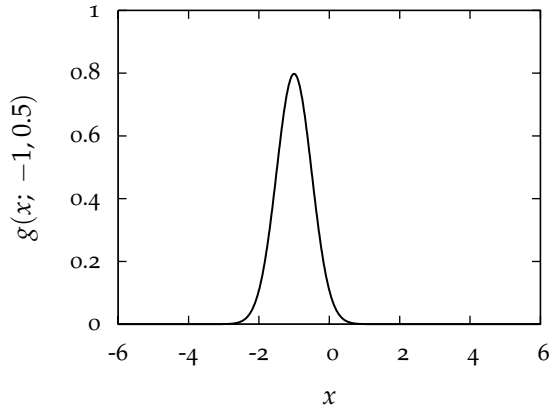
(17.1) Grafico della distribuzione di Gauss per $\mu = 0$ e $\sigma = 1$.



(17.2) Grafico della distribuzione di Gauss per $\mu = 0$ e $\sigma = 2$.



(17.3) Grafico della distribuzione di Gauss per $\mu = 1$ e $\sigma = 1$.



(17.4) Grafico della distribuzione di Gauss per $\mu = -1$ e $\sigma = 0.5$.

FIGURA 17.. Grafici della funzione $g(x; \mu, \sigma)$ per diversi valori di σ e μ . Dovendo valere la condizione di normalizzazione, il valore massimo della funzione di distribuzione diminuisce all'aumentare della deviazione standard.

Per calcolare la semilarghezza a metà altezza possiamo allora utilizzare (quasi) direttamente la (74). Nella variabile t si ha, con il linguaggio consueto:

$$t_{\pm} = \pm \sqrt{\ln 2}$$

e, passando alla x :

$$x_{\pm} = \mu + \sqrt{2} \sigma t_{\pm} = \mu \pm \sqrt{2 \ln 2} \sigma$$

da cui infine:

$$\text{WHM} = \frac{x_+ - x_-}{2} = \sqrt{2 \ln 2} \sigma \approx 1.2\sigma$$

Ne segue che, come vedremo più in dettaglio tra breve, μ rappresenta il centro della campana e σ dà una misura della larghezza.

3.5.3 Normalizzazione, media e varianza

Come tutte le funzioni di distribuzione la (76) è correttamente normalizzata:

$$\int_{-\infty}^{\infty} g(x; \mu, \sigma) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

Con il solito cambiamento di variabile (77) l'integrale diviene:

$$\int_{-\infty}^{\infty} g(x; \mu, \sigma) dx = \frac{1}{\sigma\sqrt{2\pi}} \cdot \sigma\sqrt{2} \int_{-\infty}^{\infty} e^{-t^2} dt = 1$$

dove, nell'ultimo passaggio, si è usata la (75).

La media della distribuzione è data da:

$$E[x] = \int_{-\infty}^{\infty} x \cdot g(x; \mu, \sigma) dx = \mu$$

come si può dimostrare calcolando direttamente l'integrale o, più semplicemente, osservando che la funzione $g(x; \mu, \sigma)$ è simmetrica rispetto ad $x = \mu$. Ne segue che la costante che compare nell'esponente della funzione di Gauss (ed è indicata con μ) è proprio la media della distribuzione nel senso della (34).

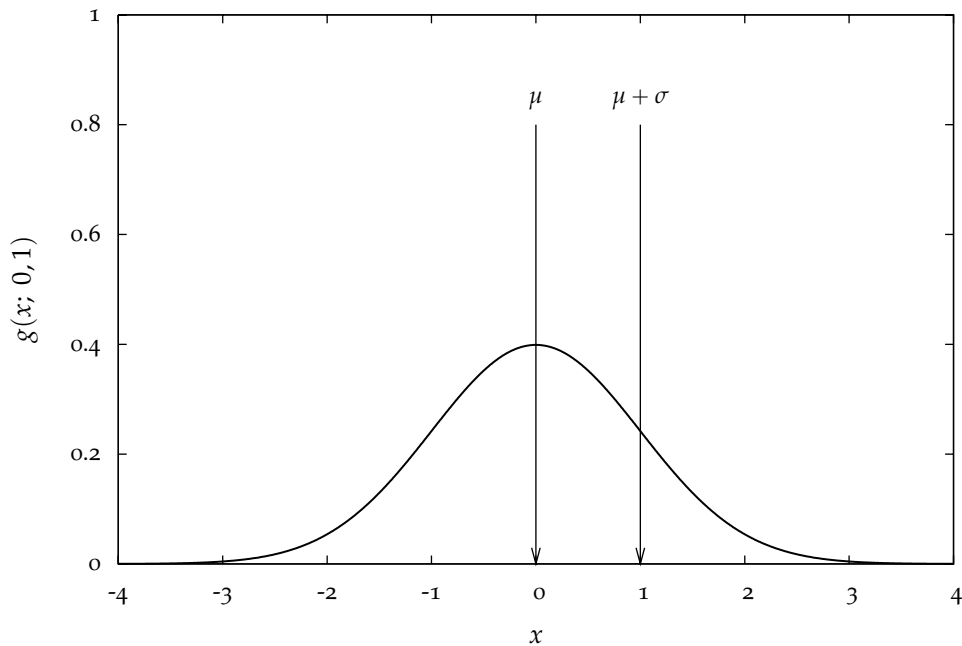


FIGURA 18.. Grafico della distribuzione di Gauss per $\mu = 0$ e $\sigma = 1$. Per completezza i punti corrispondenti a $x = \mu$ e $x = \mu + \sigma$ sono esplicitamente indicati sul grafico.

La varianza è:

$$E[(x - \mu)^2] = \int_{-\infty}^{\infty} (x - \mu)^2 \cdot g(x; \mu, \sigma) dx$$

Effettuando la sostituzione (77) si ha:

$$\begin{aligned} E[(x - \mu)^2] &= \int_{-\infty}^{\infty} 2t^2 \sigma^2 \cdot \frac{e^{-t^2}}{\sigma\sqrt{2\pi}} \cdot \sigma\sqrt{2} dt = \frac{\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} t \cdot 2te^{-t^2} dt = \\ &= \frac{\sigma^2}{\sqrt{\pi}} \left(\left[-te^{-t^2} \right]_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-t^2} dt \right) = \sigma^2 \end{aligned}$$

Cioè la costante indicata con σ nella (76) è proprio la deviazione standard nel senso della (37).

3.5.4 Integrazione della distribuzione di Gauss

Come sappiamo, in generale, la probabilità che una generica variabile gaussiana x sia compresa in nell'intervallo $[0, a]$ è:

$$P(0 \leq x \leq a) = \int_0^a g(x; \mu, \sigma) dx \quad (78)$$

Sfortunatamente questo integrale *non ha espressione analitica*²; il che non significa, ovviamente, che non possa essere valutato numericamente con il grado di accuratezza desiderato. Nelle appendici A.2 e A.3 i valori della (78) si trovano tabulati in corrispondenza di una serie di valori utili. Le tavole sono date per una variabile Gaussiana in forma standard (con media 0 e varianza 1):

$$z = \frac{x - \mu}{\sigma} \quad (79)$$

per cui la funzione di distribuzione diventa:

$$g(z; 0, 1) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}$$

In generale prima di usare le tavole è necessario passare alla variabile in forma standard; l'esempio seguente chiarisce il procedimento da seguire.

► **ESEMPIO 3.10.** Supponiamo di avere una variabile casuale gaussiana x , con media $\mu = 20$ e deviazione standard $\sigma = 4$. Qual è la probabilità che x sia compreso tra $x_1 = 22$ e $x_2 = 24$?

Passando alla variabile standard:

$$z = \frac{x - 20}{4}$$

i trasformati degli estremi x_1 e x_2 divengono:

$$\begin{aligned} z_1 &= \frac{22 - 20}{4} = 0.5 \\ z_2 &= \frac{24 - 20}{4} = 1 \end{aligned}$$

² L'esatto significato di questa affermazione viene solitamente chiarito nei corsi di Analisi matematica, ma essenzialmente l'integrale della funzione di distribuzione di Gauss su un generico intervallo non può essere espresso in forma chiusa in termini di funzioni elementari.

Per cui:

$$\begin{aligned} P(22 \leq x \leq 24) &= P(0.5 \leq z \leq 1) = \int_{0.5}^1 g(z; 0, 1) dz = \\ &= \int_0^1 g(z; 0, 1) dz - \int_0^{0.5} g(z; 0, 1) dz \approx 0.3413 - 0.1915 = 0.1498 \end{aligned}$$

in cui i valori numerici degli integrali si leggono direttamente in appendice A.2. \square

► **ESEMPIO 3.11.** Con riferimento alla variabile x definita nell'esempio precedente: qual è la probabilità che x sia compreso tra 15 e 25?

Esattamente come prima:

$$\begin{aligned} z_1 &= \frac{15 - 20}{4} = -1.25 \\ z_2 &= \frac{25 - 20}{4} = 1.25 \end{aligned}$$

$$\begin{aligned} P(15 \leq x \leq 25) &= P(-1.25 \leq z \leq 1.25) = \int_{-1.25}^{1.25} g(z; 0, 1) dz = \\ &= 2 \cdot \int_0^{1.25} g(z; 0, 1) dz \approx 2 \cdot 0.39435 = 0.7870 \end{aligned}$$

\square

► **ESEMPIO 3.12.** Sempre con riferimento alla variabile x dell'esempio precedente: qual è la probabilità che x sia maggiore di 30?

Scriviamo, al solito:

$$\begin{aligned} z_1 &= \frac{30 - 20}{4} = 2.5 \\ z_2 &= \frac{\infty - 20}{4} = \infty \end{aligned}$$

e di conseguenza:

$$\begin{aligned} P(x \geq 30) &= P(z \geq 2.5) = \int_{2.5}^{\infty} g(z; 0, 1) dz = \\ &= \int_0^{\infty} g(z; 0, 1) dz - \int_0^{2.5} g(z; 0, 1) dz \approx 0.5 - 0.4946 = 0.0054 \end{aligned}$$

\square

3.5.5 Relazione con altre distribuzioni

La distribuzione di Gauss si ottiene formalmente come caso limite della binomiale quando $n \rightarrow \infty$ e p resta costante. In pratica quando sia np che $n(1 - p)$ sono ≥ 5 la

gaussiana è già una buona approssimazione della binomiale [5]. Poiché la binomiale ha media np e deviazione standard $\sqrt{np(1-p)}$ la corrispondente gaussiana è:

$$g(x; n, p) = \frac{1}{\sqrt{2\pi np(1-p)}} e^{-\frac{1}{2} \frac{(x-np)^2}{np(1-p)}} \quad (80)$$

► ESEMPIO 3.13. Si calcoli la probabilità che, lanciando 10 monete, il numero k di teste sia compreso tra 3 e 6.

Come già sappiamo il modo corretto di impostare il problema è quello di utilizzare la statistica binomiale; con il consueto significato dei termini si ha $n = 10$ e $p = \frac{1}{2}$:

$$\mathcal{B}(k; 10, \frac{1}{2}) = \binom{10}{k} \cdot \left(\frac{1}{2}\right)^k \cdot \left(\frac{1}{2}\right)^{10-k}$$

si ottiene:

$$\begin{aligned} \mathcal{B}(3; 10, \frac{1}{2}) &= \binom{10}{3} \cdot \left(\frac{1}{2}\right)^3 \cdot \left(\frac{1}{2}\right)^7 = \frac{15}{128} \\ \mathcal{B}(4; 10, \frac{1}{2}) &= \binom{10}{4} \cdot \left(\frac{1}{2}\right)^4 \cdot \left(\frac{1}{2}\right)^6 = \frac{105}{512} \\ \mathcal{B}(5; 10, \frac{1}{2}) &= \binom{10}{5} \cdot \left(\frac{1}{2}\right)^5 \cdot \left(\frac{1}{2}\right)^5 = \frac{63}{256} \\ \mathcal{B}(6; 10, \frac{1}{2}) &= \binom{10}{6} \cdot \left(\frac{1}{2}\right)^6 \cdot \left(\frac{1}{2}\right)^4 = \frac{105}{512} \end{aligned}$$

per cui:

$$P(3 \leq k \leq 6) = \mathcal{B}(3; 10, \frac{1}{2}) + \mathcal{B}(4; 10, \frac{1}{2}) + \mathcal{B}(5; 10, \frac{1}{2}) + \mathcal{B}(6; 10, \frac{1}{2}) \approx 0.7734$$

Per usare l'approssimazione gaussiana, k deve essere considerata formalmente come una variabile continua (e non discreta, come essa è in realtà). In questo schema, dunque, ciò che si deve calcolare è $P(2.5 \leq x \leq 6.5)$ per una variabile gaussiana con media

$$\mu = np = 10 \cdot \frac{1}{2} = 5$$

e deviazione standard

$$\sigma = \sqrt{np(1-p)} = \sqrt{5 \cdot \frac{1}{2}} \approx 1.58$$

Passando alla variabile standard gli estremi di integrazione divengono:

$$\begin{aligned} z_1 &= \frac{2.5 - 5}{1.58} \approx -1.58 \\ z_2 &= \frac{6.5 - 5}{1.58} \approx 0.95 \end{aligned}$$

ed infine:

$$P(2.5 \leq k \leq 6.5) = 0.4429 + 0.3289 = 0.7718$$

Che è molto vicino al valore 0.7734 trovato prima.

In questo caso il lavoro per calcolare i due valori è circa lo stesso, ma se, ad esempio, $n = 100$ e si cerca la probabilità che k sia compreso tra 30 e 60, non disponendo di un computer o di molta pazienza, è di gran lunga preferibile fare uso dell'approssimazione gaussiana, che, tra l'altro, diventa più precisa al crescere di n . \square

Si può dimostrare che la distribuzione di Gauss si ottiene anche come limite della distribuzione di Poisson per grandi valori della media μ (vedi figura 19). In questo caso la deviazione standard, come sappiamo, è data da $\sigma = \sqrt{\mu}$ per cui la gaussiana corrispondente sarà:

$$g(x; \mu) = \frac{1}{\sqrt{2\pi\mu}} e^{-\frac{1}{2} \frac{(x-\mu)^2}{\mu}} \quad (81)$$

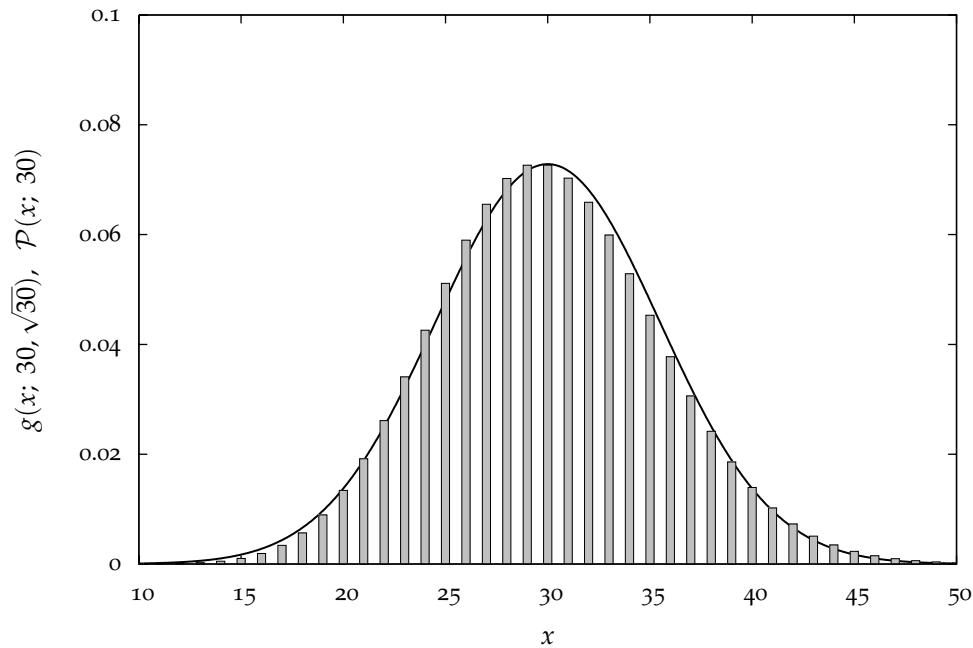


FIGURA 19.. Confronto tra una Poissoniana con media $\mu = 30$ ed una Gaussiana con media $\mu = 30$ e deviazione standard $\sigma = \sqrt{30}$.

► **ESEMPIO 3.14.** Supponiamo che in un dato campione di materiale radioattivo vi siano in media 30 decadimenti al secondo. Ci chiediamo quale sia la probabilità che in un certo intervallo di 1 s il numero di decadimenti sia compreso tra 30 e 35.

Se l'ipotesi di stazionarietà è verificata il processo segue la statistica di Poisson e la probabilità di avere k decadimenti è:

$$\mathcal{P}(k; 30) = \frac{30^k}{k!} e^{-30}$$

In particolare:

$$\begin{aligned}
 \mathcal{P}(30; 30) &= \frac{30^{30}}{30!} e^{-30} \approx 0.0726 \\
 \mathcal{P}(31; 30) &= \frac{30^{31}}{31!} e^{-30} \approx 0.0703 \\
 \mathcal{P}(32; 30) &= \frac{30^{32}}{32!} e^{-30} \approx 0.0659 \\
 \mathcal{P}(33; 30) &= \frac{30^{33}}{33!} e^{-30} \approx 0.0599 \\
 \mathcal{P}(34; 30) &= \frac{30^{34}}{34!} e^{-30} \approx 0.0529 \\
 \mathcal{P}(35; 30) &= \frac{30^{35}}{35!} e^{-30} \approx 0.0453
 \end{aligned}$$

per cui la probabilità cercata è:

$$P(30 \leq k \leq 35) = \sum_{k=30}^{35} \mathcal{P}(k; 30) \approx 0.3745$$

Proviamo adesso ad applicare l'approssimazione gaussiana. Analogamente all'esempio precedente dobbiamo calcolare $P(29.5 \leq x \leq 35.5)$ per una variabile gaussiana con media

$$\mu = 30$$

e deviazione standard

$$\sigma = \sqrt{\mu} \approx 5.48$$

Passando alla variabile standard gli estremi di integrazione divengono:

$$\begin{aligned}
 z_1 &= \frac{29.5 - 30}{5.48} \approx -0.09 \\
 z_2 &= \frac{35.5 - 30}{5.48} \approx 1.00
 \end{aligned}$$

ed infine:

$$P(29.5 \leq k \leq 35.5) = 0.0359 + 0.3413 = 0.3772$$

che di nuovo è molto vicino al valore trovato con il calcolo esatto. Notiamo anche che se ci fossimo chiesti, tanto per fare un esempio, la probabilità di avere meno di 30 conteggi, il calcolo nello schema di Poisson sarebbe stato estremamente più laborioso. \square

È importante sottolineare che l'utilità della distribuzione di Gauss non è limitata al fatto che essa sia il limite di una binomiale o di una Poissoniana. In realtà in tutti i casi in cui le fluttuazioni di una variabile casuale continua sono dovute alla somma di tanti piccoli contributi indipendenti, la gaussiana sembra descrivere bene la distribuzione di questa variabile. In fisica questa situazione si presenta molto spesso, per cui in pratica quando non ci sono errori sistematici o cause di errore fortemente correlate tra di loro, cioè non indipendenti, la distribuzione dei risultati di una misura è molto ben descritta da una gaussiana.

3.6 DISTRIBUZIONE DEL χ^2 3.6.1 Premessa: la funzione Γ di Eulero

Definiamo la funzione Γ di Eulero che, come vedremo, compare nella costante di normalizzazione della distribuzione del χ^2 , attraverso la relazione integrale:

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt$$

con x complesso e t reale. Si può dimostrare facilmente, integrando per parti, che:

$$\Gamma(x+1) = x \cdot \Gamma(x)$$

Allora, notando che $\Gamma(1) = 1$, è semplice verificare che, per ogni numero naturale, la funzione Γ gode della notevole proprietà:

$$\Gamma(n+1) = n!$$

per cui può essere considerata come una sorta di generalizzazione al campo complesso della funzione fattoriale.

3.6.2 La distribuzione del χ^2

Sia x la somma dei quadrati di n variabili gaussiane z_i in forma standard ($\mu = 0$, $\sigma = 1$) indipendenti:

$$x = \sum_{i=1}^n z_i^2$$

Questa quantità si chiama χ^2 (*chi quadro*) ad n gradi di libertà, è definito nell'intervallo $[0, \infty)$ e la sua funzione di distribuzione è:

$$\wp(x; n) = C_n x^{\frac{n-2}{2}} e^{-\frac{x}{2}} \quad (82)$$

dove C_n è una costante di normalizzazione che vale:

$$C_n = \frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} \quad (83)$$

Per grandi valori di n ($n \geq 30$), la funzione di distribuzione della variabile

$$z = \sqrt{2\chi^2} - \sqrt{2n-1}$$

tende ad una distribuzione normale con media zero e deviazione standard uno.

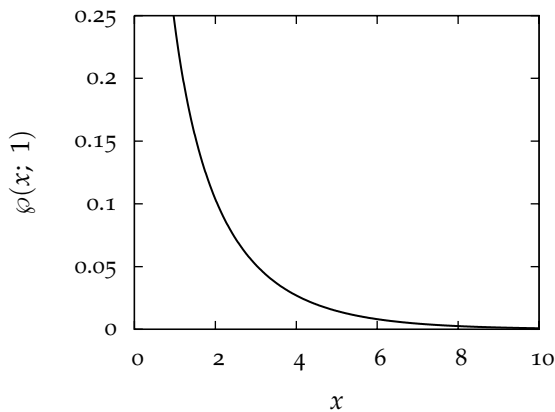
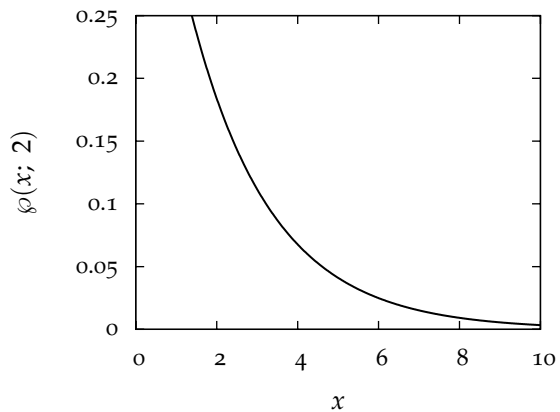
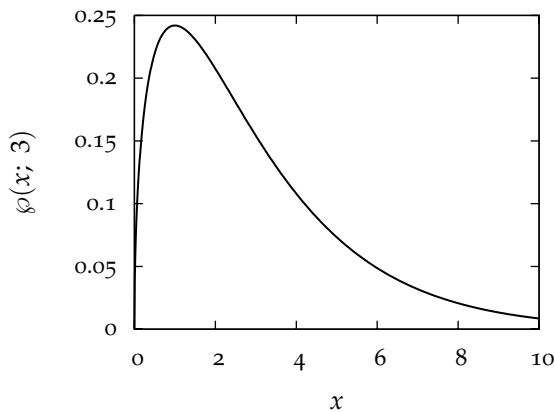
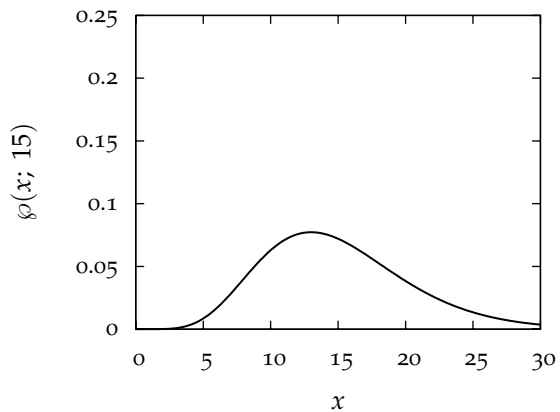
(20.1) Grafico della distribuzione del χ^2 ad 1 grado di libertà.(20.2) Grafico della distribuzione del χ^2 a 2 gradi di libertà.(20.3) Grafico della distribuzione del χ^2 a 3 gradi di libertà.(20.4) Grafico della distribuzione del χ^2 a 15 gradi di libertà.

FIGURA 20.. Grafici della distribuzione del χ^2 per diversi gradi di libertà. Notiamo che, all'aumentare del numero n di gradi di libertà, la distribuzione tende ad assumere la caratteristica forma a campana. Dovendo valere la condizione di normalizzazione, il valore massimo della funzione decresce al crescere di n .

3.6.3 Normalizzazione, media e varianza

Essendo il calcolo della media e quello della varianza un po' laboriosi, ci limitiamo qui a riportare i risultati, che utilizzeremo nel seguito. Il valor medio della distribuzione del χ^2 è uguale al numero di gradi di libertà

$$\mu = n \quad (84)$$

e la varianza è uguale a due volte il numero di gradi di libertà:

$$\sigma^2 = 2n \quad (85)$$

per cui la deviazione standard è data da:

$$\sigma = \sqrt{2n} \quad (86)$$

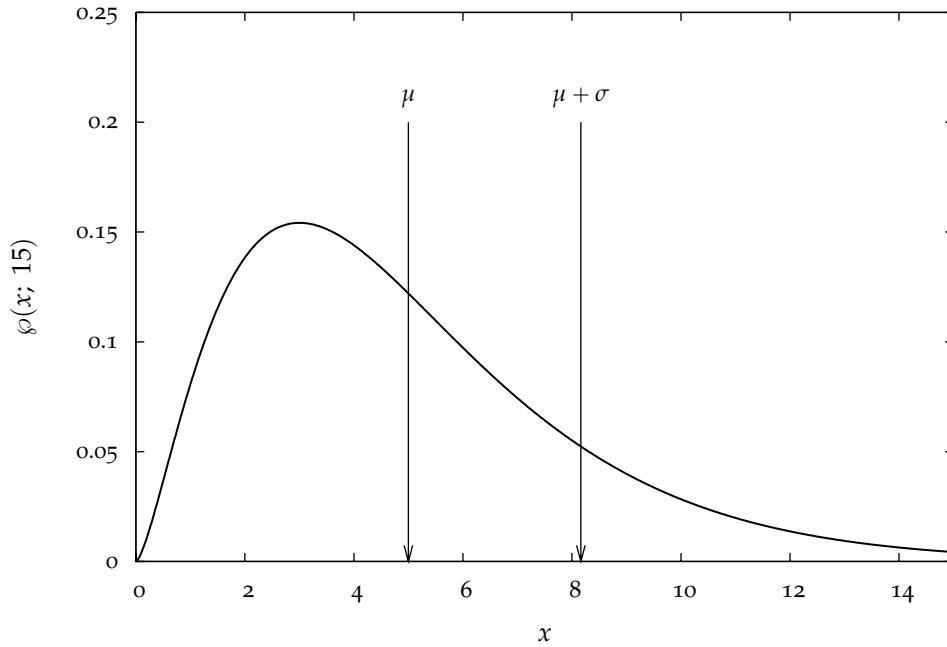


FIGURA 21.. Grafico della distribuzione del χ^2 a 5 gradi di libertà. Per completezza i punti corrispondenti a $x = \mu$ e $x = \mu + \sigma$ sono esplicitamente indicati sul grafico.

Anche per la distribuzione del χ^2 si trovano delle tavole che danno sia i valori della densità $\phi(x; n)$ sia i valori della probabilità

$$P_n(x) = \int_0^x \phi(\xi; n) d\xi$$

Rimandiamo in particolare alle tavole nelle appendici A.4 e A.5.

3.7 DISTRIBUZIONE DI CAUCHY

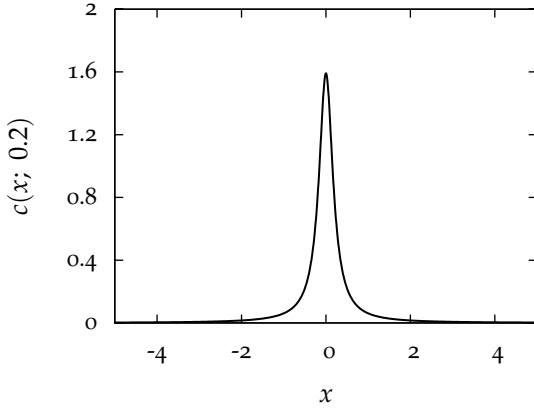
La distribuzione di Cauchy, che scriviamo nella forma:

$$c(x; a) = \frac{a}{\pi} \frac{1}{(a^2 + x^2)} \quad (87)$$

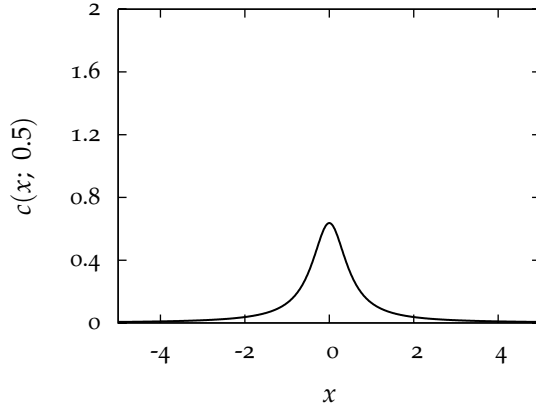
possiede alcune proprietà che la rendono particolarmente interessante. In effetti essa è ampiamente utilizzata, in fisica atomica ed in fisica delle particelle, in una forma leggermente modificata

$$\mathcal{W}(x; \mu_{\frac{1}{2}}, \Gamma) = \frac{\Gamma}{2\pi} \frac{1}{\left[\left(\frac{\Gamma}{2} \right)^2 + \left(x - \mu_{\frac{1}{2}} \right)^2 \right]} \quad (88)$$

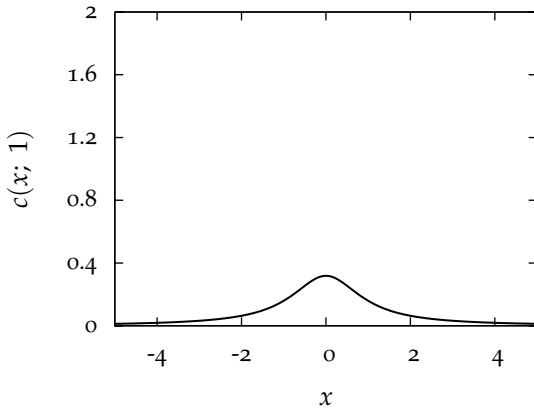
che è generalmente nota come distribuzione di Lorentz o distribuzione di Breit-Wigner. La (88) è essenzialmente identica alla (87), con la differenza che essa non è centrata in $x = 0$, ma in $x = \mu_{\frac{1}{2}}$. Nel seguito ci riferiremo alla (87) senza sostanziale perdita di generalità.



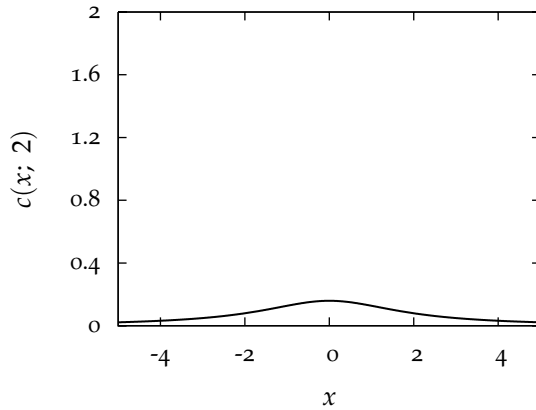
(22.1) Grafico della distribuzione di Cauchy per $a = 0.2$.



(22.2) Grafico della distribuzione di Cauchy per $a = 0.5$.



(22.3) Grafico della distribuzione di Cauchy per $a = 1$.



(22.4) Grafico della distribuzione di Cauchy per $a = 2$.

FIGURA 22.. Grafici della distribuzione di Cauchy per diversi valori di a . Dovendo valere la condizione di normalizzazione, il valore massimo della funzione di distribuzione decresce al crescere di a .

3.7.1 Normalizzazione, media e varianza

Verifichiamo prima di tutto che la distribuzione è normalizzata:

$$\int_{-\infty}^{\infty} c(x; a) dx = \int_{-\infty}^{\infty} \frac{a}{\pi} \frac{1}{(a^2 + x^2)} dx$$

Con il cambiamento di variabile $\frac{x}{a} = t$ l'integrale diviene:

$$\int_{-\infty}^{\infty} \frac{a}{\pi} \frac{1}{(a^2 + x^2)} dx = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{1}{(1 + t^2)} dt = \frac{1}{\pi} [\arctan(t)]_{-\infty}^{\infty} = \frac{1}{\pi} \cdot \left(\frac{\pi}{2} + \frac{\pi}{2} \right) = 1$$

La mediana della distribuzione esiste e, essendo la distribuzione stessa simmetrica rispetto all'asse $x = 0$, si ha $\mu_{\frac{1}{2}} = 0$.

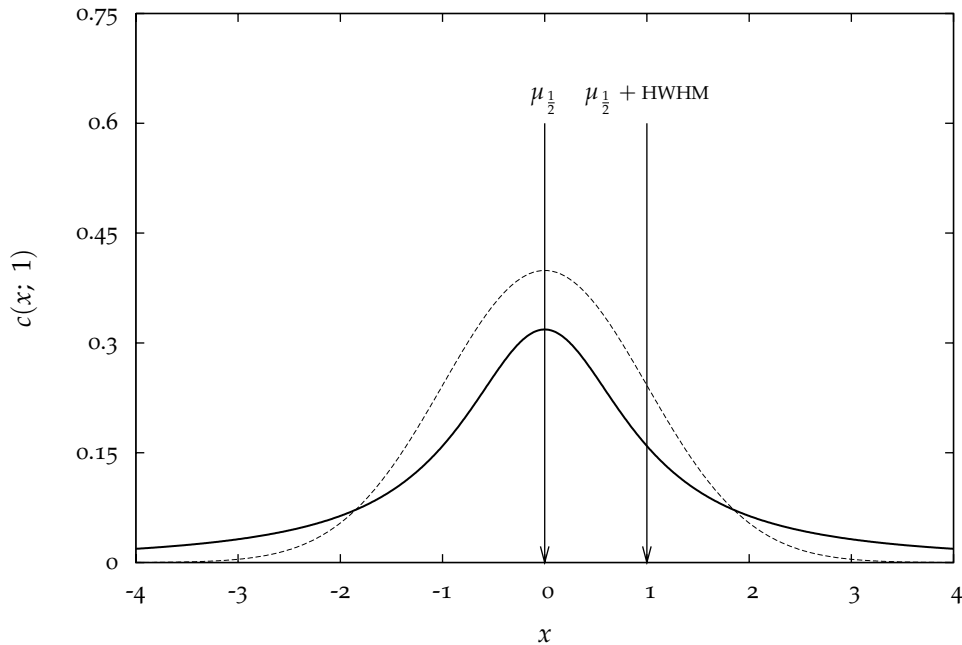


FIGURA 23.. Grafico della distribuzione di Cauchy per $a = 1$. Per completezza i punti corrispondenti a $x = \mu_{\frac{1}{2}}$ e $x = \mu_{\frac{1}{2}} + \text{HWHM}$ sono esplicitamente indicati sul grafico. La linea tratteggiata rappresenta una distribuzione di Gauss con media $\mu = 0$ e deviazione standard $\sigma = 1$; le code della gaussiana decrescono per grandi x molto più velocemente di quelle della distribuzione di Cauchy.

Formalmente la media si scrive, al solito, come:

$$\mu = E[x] = \int_{-\infty}^{\infty} x \cdot c(x; a) dx = \int_{-\infty}^{\infty} \frac{a}{\pi} \frac{x}{(a^2 + x^2)} dx$$

Purtroppo questo integrale *non è definito*. Una discussione completa dell'argomento esula dagli scopi di queste dispense; qui ci limitiamo a notare che l'integrale indefinito

$$\int \frac{a}{\pi} \frac{x}{(a^2 + x^2)} dx = \frac{1}{2} \log(a^2 + x^2)$$

quando viene valutato tra $-\infty$ e $+\infty$, porta ad una forma indeterminata del tipo $\infty - \infty$. Notiamo altresì che l'integrale definito

$$\int_{-c}^c \frac{a}{\pi} \frac{x}{(a^2 + x^2)} dx$$

esiste ed è nullo, come si può verificare per calcolo diretto oppure sfruttando il fatto che l'integrando è dispari ed il dominio di integrazione è finito e simmetrico rispetto all'origine. Si potrebbe allora essere tentati di definire la media come:

$$\mu = \lim_{c \rightarrow \infty} \int_{-c}^c \frac{a}{\pi} \frac{x}{(a^2 + x^2)} dx$$

Questo limite (si tratta di ciò che è chiamato di solito *parte principale* dell'integrale) esiste ed è zero; purtroppo questa definizione di media sarebbe causa di ulteriori problemi in relazione alla validità di alcuni risultati significativi della teoria delle probabilità. In altre parole la media della distribuzione di Cauchy non è definita.

Ne consegue che nemmeno la varianza è definita (dato che la media è necessaria alla sua definizione). Vale la pena notare che i momenti di ordine k centrati attorno all'origine divergono per $k > 1$:

$$E[x^k] = \frac{a}{\pi} \int_{-\infty}^{\infty} \frac{x^k}{(a^2 + x^2)} dx = \infty$$

Non essendo definita la varianza, in questo caso risulta particolarmente utile il concetto di semilarghezza a metà altezza; per la distribuzione di Cauchy si tratta, in un certo senso, dell'unica scelta possibile. La semilarghezza a metà altezza (HWHM) si trova calcolando i valori x_{\pm} di x per cui:

$$c(x_{\pm}; a) = \frac{1}{2} c(0; a)$$

da cui:

$$x_{\pm} = \pm a$$

ed infine:

$$\text{HWHM} = a$$

Questo significa che, a patto di scegliere a abbastanza piccolo, la distribuzione di Cauchy può essere resa *stretta* a piacere (figura 22). Il che è sorprendente se pensiamo al fatto che il momento di ordine due intorno allo 0 è infinito, indipendentemente da a .

► **ESEMPIO 3.15.** Supponiamo di avere una sorgente radioattiva, posta a distanza a da uno schermo³ infinitamente esteso, che emette particelle con distribuzione isotropa (cioè in tutte le direzioni con uguale probabilità). Siamo interessati alla distribuzione dei punti in cui queste particelle incidono sullo schermo; tratteremo il problema come se fosse bidimensionale, considerando solo le particelle emesse in un piano ortogonale allo schermo e contenente la sorgente—e quindi restringendoci alla retta identificata dall'intersezione di questo piano con lo schermo stesso.

³ Se lo schermo è, ad esempio, ricoperto con emulsione fotografica o equipaggiato opportunamente con un rivelatore, è possibile identificare con notevole precisione la posizione in cui una particella incide sullo schermo stesso.

Se indichiamo con θ l'angolo di emissione di una generica particella, misurato a partire dalla perpendicolare allo schermo passante per la sorgente, la densità di probabilità in questa variabile sarà uniforme:

$$p_{\theta}(\theta) = \begin{cases} \frac{1}{\pi} & -\frac{\pi}{2} \leq x \leq \frac{\pi}{2} \\ 0 & x < -\frac{\pi}{2}; x > \frac{\pi}{2} \end{cases}$$

(notare che abbiamo considerato solamente l'intervallo $[-\frac{\pi}{2}, \frac{\pi}{2}]$ poiché al di fuori di esso le particelle non raggiungono lo schermo). Il punto di impatto sullo schermo sarà dato da:

$$x = a \cdot \tan \theta \quad (89)$$

ed il problema consiste nel calcolare la densità di probabilità nella nuova variabile x . Formalmente si tratta di valutare la densità di probabilità di una funzione di una variabile casuale di cui sia nota la funzione di distribuzione. Tratteremo diffusamente questo problema in sezione 7.3 (cfr. in particolare l'esempio 7.8). Qui ci limitiamo ad enunciare il risultato:

$$p_x(x) = \frac{a}{\pi} \frac{1}{(a^2 + x^2)}$$

cioè la variabile x è distribuita secondo la distribuzione di Cauchy. □

3.8 SOMMARIO

Riportiamo nella tabella seguente le proprietà principali delle distribuzioni sin qui studiate.

DISTRIBUZIONE	ESPRESSIONE	MEDIA	VARIANZA
Binomiale	$\mathcal{B}(k; n, p) = \binom{n}{k} p^k q^{n-k}$	np	npq
Poissoniana	$\mathcal{P}(k; \mu) = \frac{\mu^k}{k!} e^{-\mu}$	μ	μ
Uniforme	$u(x; a, b) = \begin{cases} \frac{1}{(b-a)} & a \leq x \leq b \\ 0 & x < a; x > b \end{cases}$	$\frac{b+a}{2}$	$\frac{(b-a)^2}{12}$
Esponenziale	$\varepsilon(x; \lambda) = \begin{cases} \lambda e^{-\lambda x} & 0 \leq x \leq \infty \\ 0 & x < 0 \end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Gaussiana	$g(x; \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$	μ	σ^2
χ^2	$\wp(x; n) = \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} x^{\frac{n-2}{2}} e^{-\frac{x}{2}}$	n	$2n$
Cauchy	$c(x; a) = \frac{a}{\pi} \frac{1}{(a^2 + x^2)}$	-	-

3.9 ESERCIZI

► **ESERCIZIO 3.1.** Costruire sperimentalmente la distribuzione $\mathcal{B}(k; 4, \frac{1}{6})$ operando in questo modo:

- (1) Fare N prove, ogni prova sia un lancio di 4 dadi.
- (2) Contare quanti 5 escono in ciascuna prova.
- (3) Contare quante prove contengono k 5.
- (4) Dividere per il numero di prove, in modo da ottenere la frequenza relativa, che è una stima della probabilità.
- (5) Rappresentare le frequenze relative e confrontarle con quelle teoriche.

► **ESERCIZIO 3.2.** Sia k una variabile casuale con funzione di distribuzione binomiale. Si dimostri che vale la seguente relazione di ricorrenza:

$$E[k^{m+1}] = p(1-p) \cdot \frac{d}{dp} E[k^m] + np \cdot E[k^m]$$

Suggerimento: si scriva esplicitamente la derivata rispetto a p di $E[k^m]$:

$$\frac{d}{dp} E[k^m] = \frac{d}{dp} \sum_{k=0}^n k^m \cdot \mathcal{B}(k; n, p) = \frac{d}{dp} \sum_{k=0}^n k^m \cdot \binom{n}{k} p^k (1-p)^{n-k}$$

Si porti dunque la derivata dentro il segno di sommatoria e si esegua la derivata stessa termine per termine. A questo punto non dovrebbe essere troppo complicato provare la tesi.

► **ESERCIZIO 3.3.** Sia k una variabile casuale con funzione di distribuzione binomiale. Si utilizzi il risultato dell'esercizio 3.2 per calcolare $E[k^2]$, sapendo che $E[k] = np$.

Risposta: $[np(np - p + 1)]$

► **ESERCIZIO 3.4.** Sia k una variabile casuale con funzione di distribuzione binomiale. Si utilizzino i risultati degli esercizi 3.2 e 3.3 per calcolare $E[k^3]$.

Risposta: $[np(n^2p^2 + 2p^2 - 3np^2 + 3np - 3p + 1)]$

► **ESERCIZIO 3.5.** Calcolare il coefficiente di asimmetria γ_1 per la distribuzione binomiale. *Suggerimento: Si usino i risultati degli esercizi 3.4 e 2.23, oltre ovviamente ai fatti noti sulla distribuzione binomiale.*

Risposta: $\left[\frac{1-2p}{\sqrt{np(1-p)}} \right]$

► **ESERCIZIO 3.6.** Calcolare il limite per grande n del coefficiente di asimmetria γ_1 della distribuzione binomiale. Commentare.

Risposta: $[0]$

► ESERCIZIO 3.7. Sia k una variabile casuale con funzione di distribuzione Poissoniana. Si dimostri che vale la seguente relazione di ricorrenza:

$$E[k^{m+1}] = \mu \cdot \left(\frac{d}{d\mu} E[k^m] + E[k^m] \right)$$

Suggerimento: si scriva esplicitamente la derivata rispetto al valor medio μ di $E[k^m]$:

$$\frac{d}{d\mu} E[k^m] = \frac{d}{d\mu} \sum_{k=0}^{\infty} k^m \cdot \mathcal{P}(k; \mu) = \frac{d}{d\mu} \sum_{k=0}^{\infty} k^m \cdot \frac{\mu^k}{k!} e^{-\mu}$$

Si porti dunque la derivata dentro il segno di sommatoria e si esegua la derivata stessa termine per termine. A questo punto non dovrebbe essere troppo complicato provare la tesi.

► ESERCIZIO 3.8. Sia k una variabile casuale con funzione di distribuzione Poissoniana. Si utilizzi il risultato dell'esercizio 3.7 per calcolare $E[k^2]$, sapendo che $E[k] = \mu$.

$$\text{Risposta: } [\mu + \mu^2]$$

► ESERCIZIO 3.9. Sia k una variabile casuale con funzione di distribuzione Poissoniana. Si utilizzino i risultati degli esercizi 3.7 e 3.8 per calcolare $E[k^3]$.

$$\text{Risposta: } [\mu + 3\mu^2 - \mu^3]$$

► ESERCIZIO 3.10. Calcolare il coefficiente di asimmetria γ_1 per la distribuzione di Poisson.

Suggerimento: Si usino i risultati degli esercizi 3.9 e 2.23, oltre ovviamente ai fatti noti sulla distribuzione di Poisson.

$$\text{Risposta: } \left[\frac{1}{\sqrt{\mu}} \right]$$

► ESERCIZIO 3.11. Calcolare il limite per grande μ del coefficiente di asimmetria γ_1 della distribuzione di Poisson. Commentare.

$$\text{Risposta: } [0]$$

► ESERCIZIO 3.12. Calcolare la semilarghezza a metà altezza per una funzione di distribuzione uniforme scritta nella forma (61).

$$\text{Risposta: } \left[\frac{(b-a)}{2} \right]$$

► ESERCIZIO 3.13. Calcolare la mediana $\mu_{\frac{1}{2}}$ per una distribuzione esponenziale nella forma (65).

$$\text{Risposta: } \left[\frac{\ln(2)}{\lambda} \right]$$

► ESERCIZIO 3.14. Calcolare la semilarghezza a metà altezza per una distribuzione esponenziale nella forma (65).

$$\text{Risposta: } \left[\frac{\ln(2)}{2\lambda} \right]$$

► ESERCIZIO 3.15. Sia x una variabile casuale continua con funzione di distribuzione esponenziale, nella forma (65). Si calcoli $E[x^3]$.

Risposta: $\left[\frac{6}{\lambda^3} \right]$

► ESERCIZIO 3.16. Si calcoli il valore del coefficiente di asimmetria (45) per la distribuzione esponenziale nella forma (65).

Suggerimento: si usi il risultato dell'esercizio precedente, insieme a quello dell'esercizio 2.23.

Risposta: $[2]$

► ESERCIZIO 3.17. Si calcoli la moda della distribuzione del χ^2 in funzione del numero di gradi di libertà (si calcoli, cioè, il valore di x per cui la funzione di distribuzione $\wp(x; n)$ assume il suo valore massimo).

Suggerimento: si derivi $\wp(x; n)$ rispetto ad x e si ponga la derivata uguale a zero.

Risposta: $[n - 2]$

► ESERCIZIO 3.18. Si dimostri che per $n = 2$ la distribuzione del χ^2 $\wp(x; n)$ si riduce ad una distribuzione esponenziale.

► ESERCIZIO 3.19. Sfruttando i risultati degli esercizi precedenti ed i fatti noti dal corpo del capitolo, si compili una tabella in cui, per le principali distribuzioni continue note, siano riportate la deviazione standard σ , la semilarghezza a metà altezza HWHM ed il fattore di proporzionalità ε tra le due, dato da:

$$\text{HWHM} = \varepsilon \sigma$$

Si verifichi in particolare che, per ciascuna delle distribuzioni, ε è dell'ordine dell'unità.

 PROPAGAZIONE DEGLI ERRORI. PARTE II

 INDICE

4.1	Media e varianza di funzioni di variabili casuali	99
4.1.1	Funzioni di una sola variabile	100
4.1.2	Funzioni di due variabili	102
4.1.3	La covarianza	104
4.1.4	Caso generale	105
4.2	Media e varianza campione	107
4.3	Il teorema del limite centrale	109
4.3.1	Premessa	109
4.3.2	Il teorema del limite centrale	110
4.4	Intervalli di confidenza	114

In questo capitolo si riprende la propagazione degli errori/incertezze introdotta al Capitolo 1, dove si era trattato soltanto l'incertezza massima che si commette nell'eseguire operazioni con grandezze a, b, c, \dots affette da incertezze $\Delta a, \Delta b, \Delta c, \dots$

I risultati di quel capitolo si applicano quando le grandezze sono misurate indipendentemente e le incertezze $\Delta a, \Delta b, \Delta c, \dots$ non sono state determinate con metodi statistici. Tale situazione si incontra quando gli strumenti usati non sono molto sensibili, non si hanno fluttuazioni casuali e $\Delta a, \Delta b, \Delta c, \dots$ sono le risoluzioni degli strumenti stessi. Quando invece si presentano fluttuazioni casuali, l'incertezza è valutata tramite la stima della deviazione standard. Il problema della propagazione delle incertezze è allora il seguente: data una grandezza funzione di altre, come si possono combinare le deviazioni standard per le grandezze individuali per stimare l'incertezza sul risultato.

4.1 MEDIA E VARIANZA DI FUNZIONI DI VARIABILI CASUALI

Il problema della propagazione delle incertezze si riconduce al problema di determinare come e sotto quali ipotesi la media e la varianza di una funzione di variabili casuali sono legate alla media e varianza di tali variabili.

4.1.1 Funzioni di una sola variabile

Consideriamo dapprima il caso semplice di una grandezza G funzione di una sola variabile x :

$$G = f(x)$$

Sia μ_x il valor medio della variabile x e σ_x^2 la sua varianza:

$$\begin{aligned}\mu_x &= E[x] \\ \sigma_x^2 &= E[(x - \mu_x)^2]\end{aligned}$$

Anche la variabile G è una variabile casuale con una sua funzione di distribuzione. È pertanto possibile definirne il valor medio μ_G e la varianza σ_G^2 :

$$\begin{aligned}\mu_G &= E[G] \\ \sigma_G^2 &= E[(G - \mu_G)^2]\end{aligned}$$

Utilizzando lo sviluppo in serie di Taylor attorno al valor medio di x possiamo approssimare G come:

$$G = f(\mu_x) + \frac{df}{dx}(\mu_x) \cdot (x - \mu_x) + \frac{1}{2} \cdot \frac{d^2f}{dx^2}(\mu_x) \cdot (x - \mu_x)^2 + \mathcal{O}(3)$$

ed il suo valore di aspettazione (ricordiamo che si tratta di un operatore lineare) come:

$$\mu_G = E[G] \simeq E[f(\mu_x)] + E\left[\frac{df}{dx}(\mu_x) \cdot (x - \mu_x)\right] + E\left[\frac{1}{2} \cdot \frac{d^2f}{dx^2}(\mu_x) \cdot (x - \mu_x)^2\right]$$

in cui i termini di ordine 3 o superiore sono stati esplicitamente trascurati. Notiamo ancora una volta che le derivate sono *calcolate* nel punto attorno al quale si sviluppa (nel caso particolare il valor medio di x) per cui non sono altro che *numeri* (leggi: costanti indipendenti da x) e possono essere *portate fuori* dal segno di valore di aspettazione—questo vale anche per il valore $f(\mu_x)$ della funzione calcolata in μ_x :

$$\mu_G \simeq f(\mu_x) + \frac{df}{dx}(\mu_x) \cdot E[(x - \mu_x)] + \frac{1}{2} \cdot \frac{d^2f}{dx^2}(\mu_x) \cdot E[(x - \mu_x)^2]$$

È molto semplice convincersi che il primo ordine dello sviluppo è nullo:

$$E[(x - \mu_x)] = E[x] - E[\mu_x] = \mu_x - \mu_x = 0$$

per cui alla fine, mettendo tutto insieme, otteniamo:

$$\mu_G \simeq f(\mu_x) + \frac{1}{2} \cdot \frac{d^2f}{dx^2}(\mu_x) \cdot E[(x - \mu_x)^2]$$

Troncando lo sviluppo al second'ordine si deduce l'importante risultato:

$$\mu_G \simeq f(\mu_x)$$

(90)

valido *esattamente* quando la funzione $f(x)$ è lineare e *approssimativamente* nei casi in cui si possano trascurare i termini del second'ordine o successivi nello sviluppo della funzione stessa.

Data la (90) si ha:

$$\sigma_G^2 = E \left[(G - \mu_G)^2 \right] \simeq E \left[(f(x) - f(\mu_x))^2 \right]$$

Mettendo dentro la nostra espressione per lo sviluppo di $f(x)$ abbiamo:

$$f(x) - f(\mu_x) = \frac{df}{dx}(\mu_x) \cdot (x - \mu_x) + \mathcal{O}(2)$$

Trascurando di nuovo i termini al secondo ordine nello sviluppo in serie di $f(x)$ ne segue:

$$\sigma_G^2 \simeq E \left[\left(\frac{df}{dx}(\mu_x) \right)^2 \cdot (x - \mu_x)^2 \right] = \left(\frac{df}{dx}(\mu_x) \right)^2 \cdot E \left[(x - \mu_x)^2 \right] = \left(\frac{df}{dx}(\mu_x) \right)^2 \cdot \sigma_x^2$$

ed infine:

$$\sigma_G \simeq \left| \frac{df}{dx}(\mu_x) \right| \cdot \sigma_x \quad (91)$$

Notiamo che questo risultato è simile a quello che si ottiene per la propagazione dell'errore massimo (14):

$$\Delta G = \left| \frac{df}{da}(a_0) \right| \cdot \Delta a$$

con la deviazione standard al posto dell'errore massimo.

► **ESEMPIO 4.1.** Consideriamo una variabile casuale x distribuita uniformemente nell'intervallo $[9.9, 10.1]$:

$$p(x) = \begin{cases} 5 & 9.9 \leq x \leq 10.1 \\ 0 & x < 9.9; x > 10.1 \end{cases}$$

Sappiamo che la media e la deviazione standard di x sono, secondo la (62) e la (64):

$$\begin{aligned} \mu_x &= \frac{10.1 + 9.9}{2} = 10 \\ \sigma_x &= \frac{10.1 - 9.9}{\sqrt{12}} \approx 0.0577 \end{aligned}$$

Consideriamo adesso la funzione $G(x) = x^2$. In questo semplice caso possiamo valutare esplicitamente μ_G e σ_G :

$$\begin{aligned} \mu_G &= E[G] = \int_{-\infty}^{\infty} x^2 \cdot p(x) dx = \left[5 \frac{x^3}{3} \right]_{9.9}^{10.1} \approx 100.003 \\ \sigma_G^2 &= E[G^2] - \mu_G^2 = \int_{-\infty}^{\infty} x^4 \cdot p(x) dx - \mu_G^2 = \left[5 \frac{x^5}{5} \right]_{9.9}^{10.1} - \mu_G^2 \approx 1.333 \\ \sigma_G &= \sqrt{\sigma_G^2} \approx 1.155 \end{aligned}$$

Le equazioni (90) e (91) (che nel caso non lineare valgono solo approssimativamente) prevedono:

$$\begin{aligned}\mu_G &\simeq \mu_x^2 \approx 100 \\ \sigma_G &\simeq 2\mu_x\sigma_x \approx 1.154\end{aligned}$$

e sono in accordo eccellente con il risultato esatto. □

4.1.2 Funzioni di due variabili

Consideriamo adesso il caso di una grandezza G , funzione di due variabili casuali x e y :

$$G = f(x, y)$$

Lo sviluppo in serie di Taylor attorno al punto (μ_x, μ_y) si scrive adesso, trascurando sin dall'inizio i termini del second'ordine o superiori, come:

$$G \simeq f(\mu_x, \mu_y) + \frac{\partial f}{\partial x}(\mu_x, \mu_y) \cdot (x - \mu_x) + \frac{\partial f}{\partial y}(\mu_x, \mu_y) \cdot (y - \mu_y) \quad (92)$$

in cui le derivate, al solito, sono calcolate nel punto attorno a cui si sviluppa. Notiamo che nello sviluppo al prim'ordine compaiono le derivate parziali rispetto ad entrambe le variabili.

D'ora in avanti, per il resto di questo capitolo, ometteremo di indicare esplicitamente il punto in cui sono calcolate le derivate ovunque questo sia evidente dal contesto, nello spirito di ottenere i risultati che ci interessano in una forma più *compatta*. Riscriviamo dunque la (92) come:

$$G \simeq f(\mu_x, \mu_y) + \frac{\partial f}{\partial x} \cdot (x - \mu_x) + \frac{\partial f}{\partial y} \cdot (y - \mu_y) \quad (93)$$

e ancora, come prima nel caso di una sola variabile:

$$\begin{aligned}\mu_G &= E[G] \simeq E[f(\mu_x, \mu_y)] + E\left[\frac{\partial f}{\partial x} \cdot (x - \mu_x)\right] + E\left[\frac{\partial f}{\partial y} \cdot (y - \mu_y)\right] \\ &= f(\mu_x, \mu_y) + \frac{\partial f}{\partial x} \cdot E[(x - \mu_x)] + \frac{\partial f}{\partial y} \cdot E[(y - \mu_y)]\end{aligned}$$

Sfruttando il fatto che

$$\begin{aligned}E[(x - \mu_x)] &= 0 \\ E[(y - \mu_y)] &= 0\end{aligned}$$

abbiamo:

$$\boxed{\mu_G \simeq f(\mu_x, \mu_y)} \quad (94)$$

che, vale la pena ricordarlo ancora una volta, è valida nel limite in cui si possano trascurare i termini di ordine superiore al primo nello sviluppo in serie di Taylor di G .

Per ciò che riguarda la varianza si ha:

$$\begin{aligned}\sigma_G^2 &= E[(G - \mu_G)^2] \simeq E[(f(x, y) - f(\mu_x, \mu_y))^2] \simeq \\ &\simeq E\left[\left(\frac{\partial f}{\partial x} \cdot (x - \mu_x) + \frac{\partial f}{\partial y} \cdot (y - \mu_y)\right)^2\right] = \\ &= E\left[\left(\frac{\partial f}{\partial x}\right)^2 \cdot (x - \mu_x)^2 + \left(\frac{\partial f}{\partial y}\right)^2 \cdot (y - \mu_y)^2 + 2 \cdot \frac{\partial f}{\partial x} \cdot \frac{\partial f}{\partial y} \cdot (x - \mu_x) \cdot (y - \mu_y)\right]\end{aligned}$$

Come prima le derivate sono semplici costanti che si possono tirar fuori dai valori di aspettazione:

$$\begin{aligned}\sigma_G^2 &= \left(\frac{\partial f}{\partial x}\right)^2 \cdot E[(x - \mu_x)^2] + \left(\frac{\partial f}{\partial y}\right)^2 \cdot E[(y - \mu_y)^2] + \\ &+ 2 \cdot \frac{\partial f}{\partial x} \cdot \frac{\partial f}{\partial y} \cdot E[(x - \mu_x) \cdot (y - \mu_y)] = \\ &= \left(\frac{\partial f}{\partial x}\right)^2 \cdot \sigma_x^2 + \left(\frac{\partial f}{\partial y}\right)^2 \cdot \sigma_y^2 + 2 \cdot \frac{\partial f}{\partial x} \cdot \frac{\partial f}{\partial y} \cdot \sigma_{xy}\end{aligned}$$

Infine:

$$\sigma_G = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 \cdot \sigma_x^2 + \left(\frac{\partial f}{\partial y}\right)^2 \cdot \sigma_y^2 + 2 \cdot \frac{\partial f}{\partial x} \cdot \frac{\partial f}{\partial y} \cdot \sigma_{xy}} \quad (95)$$

dove, a costo di essere pedanti, ricordiamo ancora che le derivate si intendono calcolate nel punto (μ_x, μ_y) ; la quantità σ_{xy} , che abbiamo appena introdotto prende il nome di *covarianza* ed è definita da:

$$\sigma_{xy} = E[(x - \mu_x) \cdot (y - \mu_y)] \quad (96)$$

Ne parleremo più in dettaglio nella prossima sezione.

Se le grandezze x ed y sono statisticamente indipendenti, che è un caso piuttosto comune in laboratorio, la (95) diviene:

$$\sigma_G = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 \cdot \sigma_x^2 + \left(\frac{\partial f}{\partial y}\right)^2 \cdot \sigma_y^2} \quad (97)$$

► **ESEMPIO 4.2.** Supponiamo di lanciare due dadi a sei facce. Sia x l'uscita del primo dado, y l'uscita del secondo e $G = x + y$ la somma della uscite. Sappiamo già la funzione di distribu-

zione di x (e quindi di y) e ne abbiamo già calcolato media, varianza e deviazione standard (vedi esempi 2.26 e 2.33):

$$\begin{aligned}\mu_x &= 3.5 \\ \sigma_x^2 &= \frac{35}{12} \approx 2.917\end{aligned}$$

Analogamente per G vale (vedi esempi 2.27 e 2.34):

$$\begin{aligned}\mu_G &= 7 \\ \sigma_G^2 &= \frac{70}{12} \approx 5.833\end{aligned}$$

Le equazioni (94) e (97) prevedono:

$$\begin{aligned}\mu_G &= \mu_x + \mu_y = 2\mu_x = 7 \\ \sigma_G^2 &= \sigma_x^2 + \sigma_y^2 = 2\sigma_x^2 = \frac{70}{12}\end{aligned}$$

ed esse sono verificate esattamente e non solo approssimativamente, perché la legge che lega G a x e y è lineare. \square

4.1.3 La covarianza

Abbiamo definito la covarianza nella sezione precedente come:

$$\sigma_{xy} = E[(x - \mu_x) \cdot (y - \mu_y)]$$

Con un poco di algebra elementare, si può ricavare la relazione utile:

$$\sigma_{xy} = E[xy - x\mu_y - \mu_x y + \mu_x \mu_y] = E[xy] - \mu_x \mu_y - \mu_x \mu_y + \mu_x \mu_y = E[xy] - \mu_x \mu_y$$

La covarianza è una misura di quanto le fluttuazioni di x e y sono legate (o meglio *correlate*, come vedremo nel paragrafo 7.1) tra loro. In particolare se x e y sono statisticamente indipendenti, allora $\sigma_{xy} = 0$. È interessante notare come, in generale, il viceversa non sia sempre vero, nel senso che se la covarianza di due variabili è nulla non è detto che esse siano statisticamente indipendenti.

► **ESEMPIO 4.3.** Consideriamo una variabile casuale x e la variabile casuale x^2 . Esse non sono statisticamente indipendenti (x^2 è univocamente determinata da x). Supponiamo che la funzione di distribuzione di x abbia definiti i momenti almeno fino all'ordine tre e sia pari (cioè simmetrica rispetto all'asse $x = 0$). Allora, usando la (96):

$$\sigma_{xx^2} = E[x \cdot x^2] - \mu_x \cdot \mu_{x^2} = \mu_{x^3} - \mu_x \cdot \mu_{x^2} = 0$$

poiché se la funzione di distribuzione è pari sia μ_{x^3} che μ_x sono nulli. Dunque la covarianza è nulla. \square

Per completezza ricordiamo qui che si può dimostrare che per qualsiasi coppia di variabili vale la seguente relazione:

$$|\sigma_{xy}| \leq \sigma_x \sigma_y \quad (98)$$

4.1.4 Caso generale

I risultati appena ottenuti si possono facilmente generalizzare ad un numero arbitrario di variabili x_i ($i = 1 \dots n$):

$$\sigma_G = \sqrt{\sum_{i=1}^n \left(\frac{\partial f}{\partial x_i} \right)^2 \cdot \sigma_{x_i}^2 + 2 \sum_{i=1}^n \sum_{j \neq i}^n \left(\frac{\partial f}{\partial x_i} \right) \left(\frac{\partial f}{\partial x_j} \right) \cdot \sigma_{x_i x_j}} \quad (99)$$

dove tutte le derivate, questa volta, si intendono calcolate nel punto $(\mu_{x_1} \dots \mu_{x_n})$, ossia in corrispondenza dei valori medi di *tutte* le variabili x_i . Nell'ipotesi in cui le x_i siano statisticamente indipendenti tra di loro, la (99) diviene:

$$\sigma_G = \sqrt{\sum_{i=1}^n \left(\frac{\partial f}{\partial x_i} \right)^2 \cdot \sigma_{x_i}^2} \quad (100)$$

È di fondamentale importanza il fatto che tutti i risultati esposti rimangono validi se al posto della media e della varianza della *parent distribution* abbiamo le loro stime. Nel prossimo paragrafo risolveremo il problema, fondamentale in fisica, di stimare la media e la varianza di una distribuzione a partire da un numero finito di misure sperimentali. Per il momento riportiamo, per completezza, la forma generale della propagazione degli errori nel caso di grandezze indipendenti, che è una relazione usatissima in laboratorio e sostituisce la formula dell'errore massimo (17) derivata nel capitolo 1*. Utilizzando la notazione consueta, sia G una grandezza funzione di altre grandezze $a, b, c \dots$. Siano $a_0, b_0, c_0 \dots$ i valori misurati sperimentalmente e $\Delta a, \Delta b, \Delta c \dots$ le incertezze associate alle misure. Se interpretiamo $a_0, b_0, c_0 \dots$ come la nostra migliore stima per le medie delle distribuzioni generatrici di $a, b, c \dots$ e $\Delta a, \Delta b, \Delta c \dots$ come la migliore stima per le deviazioni standard della media, la formula di propagazione dell'errore su G si legge direttamente dalla (100):

$$\Delta G = \sqrt{\left(\frac{\partial f}{\partial a} \cdot \Delta a \right)^2 + \left(\frac{\partial f}{\partial b} \cdot \Delta b \right)^2 + \left(\frac{\partial f}{\partial c} \cdot \Delta c \right)^2 + \dots} \quad (101)$$

* Vale la pena sottolineare di nuovo come, benché formalmente simili, la (17) e la (101) differiscano profondamente in quanto a contenuto fisico (la prima ci dice che una certa grandezza è *sicuramente* contenuta entro un certo intervallo, mentre la seconda ci dice che un certo intervallo ha una ben definita probabilità di contenere il valore cercato).

in cui le derivate si intendono calcolate nei punti *misurati* ($a_0, b_0, c_0 \dots$). Notiamo ancora che, utilizzando la (98) nel caso di due variabili si ottiene:

$$\begin{aligned}\sigma_G^2 &= \left(\frac{\partial f}{\partial x}\right)^2 \cdot \sigma_x^2 + \left(\frac{\partial f}{\partial y}\right)^2 \cdot \sigma_y^2 + 2 \cdot \left(\frac{\partial f}{\partial x}\right) \left(\frac{\partial f}{\partial y}\right) \cdot \sigma_{xy} \leq \\ &\leq \left(\frac{\partial f}{\partial x}\right)^2 \cdot \sigma_x^2 + \left(\frac{\partial f}{\partial y}\right)^2 \cdot \sigma_y^2 + 2 \cdot \left(\frac{\partial f}{\partial x}\right) \left(\frac{\partial f}{\partial y}\right) \cdot \sigma_x \sigma_y = \\ &= \left(\frac{\partial f}{\partial x} \cdot \sigma_x + \frac{\partial f}{\partial y} \cdot \sigma_y\right)^2\end{aligned}$$

e dunque:

$$\sigma_G \leq \left| \frac{\partial f}{\partial x} \right| \cdot \sigma_x + \left| \frac{\partial f}{\partial y} \right| \cdot \sigma_y \quad (102)$$

che, per inciso, è generalizzabile ad un numero arbitrario di variabili. Il contenuto essenziale di questa relazione consiste nel fatto che quando, in un'operazione tra due grandezze x ed y , combiniamo le incertezze come abbiamo imparato nel capitolo 1 otteniamo sempre un limite superiore per l'incertezza valutata nel modo *statisticamente corretto* (per questo parlavamo di *incertezza massima* o *errore massimo*).

► **ESEMPIO 4.4.** Supponiamo di misurare indipendentemente il periodo T e la lunghezza l di un pendolo semplice ottenendo:

$$\begin{aligned}T &= 0.854 \pm 0.003 \text{ s} \\ l &= 1.140 \pm 0.005 \text{ m}\end{aligned}$$

Supponiamo anche che la valutazione delle incertezze associate alle misure sia stata eseguita in modo statisticamente corretto. Vogliamo ricavare una stima dell'accelerazione di gravità g utilizzando la relazione:

$$T = 2\pi\sqrt{\frac{l}{g}}$$

Invertendo l'equazione appena scritta troviamo:

$$g(l, T) = \frac{4\pi^2 l}{T^2} \approx 9.82$$

da cui:

$$\Delta g = \sqrt{\left(\frac{\partial g}{\partial l}(T_0, l_0) \cdot \Delta l\right)^2 + \left(\frac{\partial g}{\partial T}(T_0, l_0) \cdot \Delta T\right)^2} = \frac{4\pi^2}{T_0^2} \sqrt{(\Delta l)^2 + \left(\frac{2l_0 \Delta T}{T_0}\right)^2} \approx 0.08$$

Se avessimo valutato l'incertezza su g secondo la formula dell'errore massimo avremmo ottenuto:

$$\Delta g = \left| \frac{\partial g}{\partial l}(T_0, l_0) \right| \cdot \Delta l + \left| \frac{\partial g}{\partial T}(T_0, l_0) \right| \cdot \Delta T = \frac{4\pi^2}{T_0^2} \left(\Delta l + \frac{2l_0 \Delta T}{T_0} \right) \approx 0.11$$

che, come atteso, è più grande. □

C'è da osservare che anche nel caso in cui si prende come incertezza la risoluzione dello strumento, nello stimare una grandezza che è funzione di molte altre, conviene usare la (99) o la (100) perché è pessimistico pensare che tutti gli errori contribuiscano nello stesso verso; l'errore statistico è più significativo.

4.2 MEDIA E VARIANZA CAMPIONE

È rimasto il problema di vedere come trattare i risultati delle misure quando questi non sono costanti. Riprendiamo perciò alcuni dei concetti introdotti precedentemente. Era stato osservato che l'insieme delle misure sperimentali è una piccola parte di tutte le misure effettuabili ed è quindi un campione di tutte le possibili misure, che sono a loro volta distribuite secondo la cosiddetta distribuzione generatrice (o *parent distribution*). Conoscere *esattamente* la distribuzione generatrice richiederebbe un numero infinito di misure. Ci poniamo dunque il problema di valutare, a partire da un numero finito di misure, le migliori stime che possiamo dare della media e della varianza della distribuzione generatrice stessa.

Supponiamo dunque di eseguire in laboratorio n misure di una certa grandezza x e siano x_i ($i = 1 \dots n$) i risultati di tali misure. Come miglior stima della media μ si prende la media aritmetica m^+ :

$$m = \frac{1}{n} \sum_{i=1}^n x_i \quad (103)$$

m è la *media campione*, cioè la media della distribuzione campione e tende a μ quando n tende all'infinito o, meglio, converge in probabilità a μ , cioè

$$\forall \epsilon > 0 \quad P(|m - \mu| \geq \epsilon) \rightarrow 0 \quad \text{per } n \rightarrow \infty.$$

Come stima della varianza della distribuzione campione, tenuto conto del significato di varianza, potremmo essere tentati di prendere la media degli scarti al quadrato, cioè la quantità

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - m)^2$$

Questo andrebbe benissimo se le quantità $(x_i - m)$ fossero indipendenti tra di loro, ma in realtà non lo sono. Infatti:

$$\begin{aligned} \sum_{i=1}^n (x_i - m) &= \sum_{i=1}^n x_i - \sum_{i=1}^n m = \sum_{i=1}^n x_i - n \cdot m = \sum_{i=1}^n x_i - n \cdot \frac{1}{n} \sum_{i=1}^n x_i = \\ &= \sum_{i=1}^n x_i - \sum_{i=1}^n x_i = 0 \end{aligned}$$

† La notazione comunemente adottata è di indicare con m ed s^2 (lettere dell'alfabeto romano) i valori *approssimati*, ottenuti da una serie finita di misurazioni e calcolati mediante la (103) e la (104) della media μ e della varianza σ^2 (lettere dell'alfabeto greco) della distribuzione generatrice.

Fisicamente questo dipende dal fatto che il valore di m è stato ricavato proprio dagli x_i , per cui, conoscendo m e conoscendo $n - 1$ tra gli x_i , possiamo ricavare l' n -esimo x_i . Dunque le quantità indipendenti sono $n - 1$. Per dirla in altre parole, in linea di principio dovremmo prendere come stima della varianza la quantità

$$s^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2$$

ma il problema è che noi *non* conosciamo la media μ della distribuzione generatrice. Tutto ciò che conosciamo è la sua stima m e quando sostituiamo m a μ nella somma, i termini indipendenti cessano di essere n e divengono $n - 1$. Per farla breve la miglior stima della varianza della *parent distribution* (a partire da un numero finito di misure sperimentali) è data da:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - m)^2 \quad (104)$$

e, di conseguenza, la miglior stima della deviazione standard è:

$$s = \sqrt{\frac{1}{(n-1)} \sum_{i=1}^n (x_i - m)^2} \quad (105)$$

Si può dimostrare che s^2 converge in probabilità a σ^2 per n che tende all'infinito.

Vale la pena di notare, qui, che anche la media campione m è una variabile casuale che ha una sua funzione di distribuzione (torneremo su questo in seguito). Questo deriva dal fatto che, eseguendo diverse volte n misure di una certa grandezza, non otterremo sempre lo stesso valore di m . Molto spesso ciò che interessa è proprio la varianza della media campione. La stima della varianza della distribuzione della media campione s_m^2 si può calcolare a partire dalla (103):

$$s_m^2 = \sum_{i=1}^n \left(\frac{\partial m}{\partial x_i} (x_1 \dots x_n) \right)^2 \cdot \sigma_{x_i}^2$$

Ma ogni elemento del campione appartiene alla stessa *parent distribution* di varianza σ^2 stimata da s^2 ; inoltre per la (103):

$$\frac{\partial m}{\partial x_i} (x_1 \dots x_n) = \frac{1}{n}$$

per cui, mettendo tutto insieme:

$$s_m^2 = \sum_{i=1}^n \frac{1}{n^2} \cdot \sigma^2 \simeq \sum_{i=1}^n \frac{1}{n^2} \cdot s^2$$

Ma nella somma sia s che n non dipendono dall'indice i su cui la somma stessa è eseguita, per cui possono essere portati fuori:

$$s_m^2 = \frac{s^2}{n^2} \cdot \sum_{i=1}^n 1 = \frac{s^2}{n^2} \cdot n = \frac{s^2}{n}$$

La stima della deviazione standard (talvolta chiamata scarto quadratico medio della media) è dunque:

$$s_m = \sqrt{\frac{1}{n(n-1)} \sum_{i=1}^n (x_i - m)^2} \quad (106)$$

che si può scrivere anche come:

$$s_m = \frac{s}{\sqrt{n}} \quad (107)$$

► **ESEMPIO 4.5.** Supponiamo di misurare n volte una certa grandezza x e supponiamo che i valori misurati x_i fluttuino in modo casuale. Il risultato della misura si può scrivere come:

$$x = m \pm s_m$$

oppure anche

$$x = m \pm 2s_m$$

dove m e s_m sono dati dalle (103) e (106). La questione è di fondamentale importanza, in quanto questa è la situazione tipica che il fisico sperimentale incontra ogni giorno in laboratorio. Torneremo più avanti a discutere il significato di queste due scritte con maggior dettaglio. □

4.3 IL TEOREMA DEL LIMITE CENTRALE

4.3.1 Premessa

Consideriamo n variabili casuali indipendenti $x_1, x_2 \dots x_n$, con media $\mu_1, \mu_2 \dots \mu_n$ e varianza $\sigma_1^2, \sigma_2^2 \dots \sigma_n^2$ e sia S (che pure è una variabile casuale) la loro somma

$$S = x_1 + x_2 + \dots + x_n = \sum_{i=1}^n x_i$$

Poiché S è una funzione lineare delle x_i si può facilmente mostrare che la media μ_S e la varianza σ_S^2 di S sono date esattamente da:

$$\mu_S = \mu_1 + \mu_2 + \dots + \mu_n = \sum_{i=1}^n \mu_i \quad (108)$$

$$\sigma_S^2 = \sigma_1^2 + \sigma_2^2 + \dots + \sigma_n^2 = \sum_{i=1}^n \sigma_i^2 \quad (109)$$

La dimostrazione segue sostanzialmente dalla 100 ‡:

$$\sigma_S^2 = \sum_{i=0}^n \left(\frac{\partial S}{\partial x_i} \right) \cdot \sigma_i^2$$

‡ Ricordiamo ancora che, nel caso lineare, la 100 vale esattamente e non solo come approssimazione al prim'ordine.

notando che:

$$\frac{\partial S}{\partial x_i} = 1 \quad \forall i = 1 \dots n$$

ed abbiamo già visto un esempio nel caso di due variabili (cfr. esempio 4.2).

4.3.2 Il teorema del limite centrale

Il contenuto fondamentale del teorema del limite centrale [6] è racchiuso nel seguente enunciato:

TEOREMA 4.1 (DEL LIMITE CENTRALE). Siano date n variabili casuali indipendenti x_i ; sotto condizioni molto deboli, e indipendentemente dalle funzioni di distribuzione delle singole x_i , la somma $S = x_1 + x_2 + \dots + x_n$ è asintoticamente normale nel limite $n \rightarrow \infty$.

Più precisamente per n grande la funzione di distribuzione di S tende ad una distribuzione gaussiana con media e varianza date dalla (108) e dalla (109), rispettivamente.

In particolare se le variabili indipendenti x_i sono tutte distribuite secondo la stessa funzione di distribuzione, con media μ e varianza σ^2 , allora la somma S è asintoticamente normale con media:

$$\mu_S = \sum_{i=1}^n \mu = n\mu \quad (110)$$

e varianza:

$$\sigma_S^2 = \sum_{i=1}^n \sigma^2 = n\sigma^2 \quad (111)$$

► **ESEMPIO 4.6.** Sia la variabile casuale x il numero di teste che escono nel lancio di n monete. Formalmente x è la somma di n variabili casuali x_i distribuite secondo la stessa funzione di distribuzione elementare:

x_i	0	1
$P(x_i)$	$\frac{1}{2}$	$\frac{1}{2}$

che regola il numero di teste (può essere solo 0 o 1) che escono nel lancio di una moneta. È facile convincersi che la media e la varianza di tutte le x_i è:

$$\begin{aligned} \mu_i &= 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{1}{2} \\ \sigma_i^2 &= \left(0 - \frac{1}{2}\right)^2 \cdot \frac{1}{2} + \left(1 - \frac{1}{2}\right)^2 \cdot \frac{1}{2} = \frac{1}{4} \end{aligned}$$

Esaminiamo per prima cosa il caso semplice $n = 2$. Con due monete si hanno quattro possibili uscite (TT, TC, CT, CC), i valori che la variabile numero di teste può assumere sono 0, 1 e 2 e la funzione di distribuzione si calcola banalmente come rapporto tra i casi favorevoli e quelli possibili:

x_i	0	1	2
$P(x_i)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

In questo caso la media e la varianza sono rispettivamente:

$$\begin{aligned}\mu &= 0 \cdot \frac{1}{4} + 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} = 1 \\ \sigma^2 &= (0-1)^2 \cdot \frac{1}{4} + (1-1)^2 \cdot \frac{1}{2} + (2-1)^2 \cdot \frac{1}{4} = \frac{1}{2}\end{aligned}$$

ed è facile verificare che, in accordo con le (110) e (111), si ha:

$$\begin{aligned}\mu &= 2\mu_i \\ \sigma^2 &= 2\sigma_i^2\end{aligned}$$

Con $n = 3$ monete il ragionamento si replica esattamente:

x_i	0	1	2	3
$P(x_i)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

e, analogamente a prima:

$$\begin{aligned}\mu &= 0 \cdot \frac{1}{8} + 1 \cdot \frac{3}{8} + 2 \cdot \frac{3}{8} + 3 \cdot \frac{1}{8} = \frac{12}{8} = \frac{3}{2} = 3\mu_i \\ \sigma^2 &= \left(0 - \frac{3}{2}\right)^2 \cdot \frac{1}{8} + \left(1 - \frac{3}{2}\right)^2 \cdot \frac{3}{8} + \left(2 - \frac{3}{2}\right)^2 \cdot \frac{3}{8} + \left(3 - \frac{3}{2}\right)^2 \cdot \frac{1}{8} = \frac{3}{4} = 3\sigma_i^2\end{aligned}$$

In generale è facile convincersi che la probabilità di avere k teste lanciando n monete è data dalla distribuzione binomiale $\mathcal{B}(k; n, \frac{1}{2})$:

$$P(x = k) = \binom{n}{k} \cdot \left(\frac{1}{2}\right)^n$$

La media e la varianza sono:

$$\begin{aligned}\mu &= np = n \cdot \frac{1}{2} = n\mu_i \\ \sigma^2 &= npq = n \cdot \frac{1}{2} \cdot \frac{1}{2} = n \cdot \frac{1}{4} = n\sigma_i^2\end{aligned}$$

esattamente come previsto dalle (110) e (111). Al crescere di n , la funzione di distribuzione tende, come affermato dal teorema del limite centrale, alla classica forma a campana caratteristica della gaussiana (cfr. anche figura 8). \square

Una conseguenza fondamentale del teorema del limite centrale è che la media aritmetica di n variabili casuali indipendenti con la stessa funzione di distribuzione (di media μ e varianza σ^2):

$$m = \frac{1}{n} \sum_{i=1}^n x_i$$

è asintoticamente normale con media, varianza e deviazione standard date rispettivamente da:

$$\mu_m = \mu \quad (112)$$

$$\sigma_m^2 = \frac{\sigma^2}{n} \quad (113)$$

$$\sigma_m = \frac{\sigma}{\sqrt{n}} \quad (114)$$

Abbiamo già ricavato una relazione analoga che lega tra loro la stima della deviazione standard della *parent distribution* e quella della deviazione standard della media (cfr. 107). Il teorema del limite centrale ci dice anche che la media segue una distribuzione che è ben approssimata da una gaussiana al crescere delle dimensioni del campione. L'importanza del teorema sta inoltre nel fatto che la convergenza è rapida; già per la media di 6 misure la distribuzione è ragionevolmente gaussiana. Quando il campione è di 20 o più misure la distribuzione è praticamente indistinguibile da una gaussiana in quasi tutti i casi sperimentati.

► ESEMPIO 4.7. Torniamo per un attimo all'esempio 4.5; alla luce del teorema del limite centrale (che ci dice che la media è distribuita gaussianamente) possiamo chiarire il significato di ciò che avevamo scritto. Poiché per una gaussiana, come si può facilmente leggere dalle tavole, si ha:

$$P(|x - \mu| \leq \sigma) = 68.26\%$$

la scrittura:

$$x = m \pm s_m$$

significa che l'intervallo $[m - s_m, m + s_m]$ ha il 68.26% di probabilità di contenere il valor medio della distribuzione generatrice della nostra variabile casuale x (che essenzialmente è il numero che ci interessa). Se raddoppiamo l'intervallo, prendendo:

$$x = m \pm 2s_m,$$

dato che per una gaussiana:

$$P(|x - \mu| \leq 2\sigma) = 95.44\%$$

questa probabilità diviene del 95.44%. □

► ESEMPIO 4.8. Si misura preliminarmente per molte volte una certa grandezza x , ottenendo una serie di risultati x_i , da cui possiamo stimare la media e la deviazione standard della distribuzione generatrice del campione secondo le equazioni (103) e (105). Per fissare le idee supponiamo che, in una qualche unità di misura:

$$\begin{aligned} m &= 10.56 \\ s &= 0.53 \end{aligned}$$

Qual è la probabilità che, su una singola misura, otteniamo un valore di x maggiore di 11?

Senza ulteriori informazioni non esiste una risposta univoca a questa domanda, in quanto la probabilità cercata dipende dalla forma analitica della distribuzione generatrice (di cui noi conosciamo solamente una stima della media e della deviazione standard). Se assumiamo una particolare distribuzione generatrice $p(x)$ (per esempio sulla base dei dati sperimentali o di input della teoria), allora il calcolo si effettua banalmente secondo la formula consueta:

$$P(x \geq 11) = \int_{11}^{+\infty} p(x) dx$$

Tanto per fissare le idee, supponiamo che la distribuzione generatrice sia gaussiana, che è un caso piuttosto tipico in laboratorio. Passando alla variabile standard e consultando le tavole si ha:

$$P(x \geq 11) = P(z \geq 0.83) \simeq 0.2033$$

□

► **ESEMPIO 4.9.** Con riferimento alla situazione dell'esempio precedente: qual è la probabilità che, misurando la grandezza x per 30 volte, la media m delle nostre misure sia più grande di 11?

Questa volta il problema è ben posto in quanto il teorema del limite centrale ci assicura che la media di 30 misure è, con ottima approssimazione, distribuita gaussianamente. Le stime per la media e la deviazione standard sono:

$$\begin{aligned} m_m &= m = 10.56 \\ s_m &= \frac{s}{\sqrt{n}} = 0.10 \end{aligned}$$

Passando alla variabile standard e consultando le tavole in appendice:

$$P(m \geq 11) = P(z \geq 4.40) \simeq 5.41 \cdot 10^{-6}$$

Vale la pena sottolineare che la probabilità è molto più bassa rispetto al valore calcolato nell'esempio precedente. □

È chiaro che questi discorsi sono validi solo al tendere all'infinito delle dimensioni del campione. Se il campione è piccolo, per qualunque probabilità assegnata, bisogna introdurre un fattore correttivo che dipende dal numero di gradi di libertà ν ^{††}:

$$\nu = n - 1$$

cioè dal numero delle grandezze indipendenti [6, 7]. Torneremo sull'argomento con dettagli maggiori nel paragrafo 7.2; per il momento ci limitiamo a menzionare il fatto che esistono tavole (vedi appendice A.8) per i fattori correttivi t_p in funzione della probabilità p e del numero di gradi di libertà ν . L'esempio seguente chiarisce il loro uso.

^{††} Torneremo sul concetto di *numero di gradi di libertà* nel paragrafo 5.5. Nel caso in questione $\nu = n - 1$ poiché nella nostra stima della varianza s^2 utilizziamo la migliore stima di m anziché μ (che è ignota). Confronta anche la (104) e (105).

► **ESEMPIO 4.10.** Supponiamo di aver misurato una determinata grandezza per $n = 5$ volte e di aver valutato, a partire dalle nostre misure, le migliori stime m e s_m per la media e la deviazione standard della media. Ci chiediamo quale sia l'intervallo intorno a m in cui abbiamo il 95% di probabilità di trovare μ (cfr. paragrafo 4.4).

Se il nostro campione fosse molto grande m sarebbe distribuita gaussianamente intorno a μ con deviazione standard s_m e l'intervallo cercato sarebbe:

$$m \pm 1.96 \cdot s_m$$

(poco meno di 2 deviazioni standard, valore per cui la probabilità è, come sappiamo dall'esempio 4.7, il 95.44%). Con 5 misure il numero di gradi di libertà è invece:

$$\nu = 5 - 1 = 4$$

ed il fattore correttivo $t_{0.95}$ per 4 gradi di libertà si legge dalla tavola in appendice A.8 (terza riga, seconda colonna). Il nostro intervallo è dunque:

$$m \pm t_{0.95} \cdot s_m = m \pm 2.776 \cdot s_m$$

□

4.4 INTERVALLI DI CONFIDENZA

Approfondiamo qui alcuni dei concetti esposti nel paragrafo precedente. Il teorema del limite centrale ci assicura che, se abbiamo un campione costituito da un numero n molto grande di misure, ognuna estratta da una distribuzione generatrice di media μ e deviazione standard σ , la distribuzione della media m è normale con media $\mu_m = \mu$ e deviazione standard $\sigma_m = \sigma/\sqrt{n}$. Consideriamo la variabile:

$$z = \frac{m - \mu}{\sigma_m}$$

che ha valor medio 0 e varianza 1, come si può immediatamente verificare. La probabilità che z assuma valori compresi tra z_1 e z_2 è data da :

$$P(z_1 \leq z \leq z_2) = \int_{z_1}^{z_2} g(z) dz = \frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} e^{-\frac{z^2}{2}} dz$$

In particolare, tanto per fissare le idee, per $z_1 = -1$ e $z_2 = 1$ questa probabilità è, come sappiamo, circa il 68.26%:

$$P(-1 \leq z \leq 1) \approx 0.6826 \quad (115)$$

Notiamo esplicitamente che la disuguaglianza:

$$-1 \leq z \leq 1$$

si legge come:

$$-1 \leq \frac{m - \mu}{\sigma_m} \leq 1$$

e, dopo poche semplici manipolazioni:

$$m - \sigma_m \leq \mu \leq m + \sigma_m$$

La (115) si può dunque riscrivere come:

$$P(m - \sigma_m \leq \mu \leq m + \sigma_m) \approx 0.6826 \quad (116)$$

In questo caso l'intervallo $[m - \sigma_m, m + \sigma_m]$ viene chiamato *intervallo di confidenza* per il parametro μ con *coefficiente di confidenza* del 68%; $m - \sigma_m$ e $m + \sigma_m$ sono i *limiti di confidenza* inferiore e superiore rispettivamente.

Per $z_1 = -1.96$ e $z_2 = 1.96$ si costruisce l'intervallo di confidenza con coefficiente di confidenza del 95% $[m - 1.96 \cdot \sigma_m, m + 1.96 \cdot \sigma_m]$ intendendo che:

$$P(m - 1.96 \cdot \sigma_m \leq \mu \leq m + 1.96 \cdot \sigma_m) = 0.95$$

In generale, fissato il coefficiente di confidenza (cioè il valore di probabilità P_0) rimangono fissati gli estremi dell'intervallo di confidenza $[m - z_1 \cdot \sigma_m, m + z_2 \cdot \sigma_m]$. Data la simmetria della funzione di distribuzione è consuetudine prendere $z_1 = z_2 = z_0$. Scriviamo pertanto:

$$P(m - z_0 \cdot \sigma_m \leq \mu \leq m + z_0 \cdot \sigma_m) = P_0 \quad (117)$$

La quantità $1 - P_0$ è chiamata *livello di confidenza*.

Vale la pena notare che μ non è una variabile casuale e quindi non ci chiediamo quale sia la probabilità che il suo valore sia compreso in un generico intervallo $[m_1, m_2]$. Ma supponiamo di avere fissato il coefficiente di confidenza P_0 e ripetiamo varie volte le n prove considerate. Si otterranno in generale, secondo quanto detto sino ad ora, varie coppie di m_1, m_2 (cambia il valor medio del campione in ogni ripetizione, mentre z_1 e z_2 rimangono fissi) e ha senso chiedersi quale sia la probabilità che la variabile casuale m_1 assuma un valore minore di μ mentre la variabile casuale m_2 assuma un valore maggiore di μ . In altre parole ci chiediamo quale è la probabilità che l'intervallo $[m_1, m_2]$, variabile casualmente, contenga il punto μ . Questa probabilità, che è la stessa in ciascuna delle n prove, è P_0 . Questo significa che se uno usa ripetutamente l'intervallo di confidenza così costruito e ogni volta afferma che l'intervallo contiene μ , si può aspettare di fare una affermazione corretta $P_0 \cdot 100$ volte su 100 affermazioni.

Se σ (come tipicamente avviene) non è nota, l'intervallo di confidenza (117) ha estremi non noti. Dobbiamo allora sostituire alla varianza σ^2 una sua stima, e il contenuto probabilistico dell'intervallo non è più lo stesso, specialmente per piccoli campioni. Introduciamo la nuova variabile

$$t = \frac{m - \mu}{s_m} \quad (118)$$

dove s_m è la stima della deviazione standard della media (106). Si può dimostrare (cfr. paragrafo 7.2) che la variabile t ha la distribuzione di Student a $n - 1$ gradi di libertà, pertanto viene di solito indicata come t_{n-1} . Allora il coefficiente di confidenza rimane definito tramite la funzione di distribuzione di Student a $n - 1$ gradi di libertà che indichiamo S_{n-1} :

$$P(t_1, t_2) = \int_{t_1}^{t_2} S_{n-1}(t) dt$$

Data la simmetria della funzione di distribuzione S_{n-1} intorno allo zero, è conveniente, di nuovo, prendere $t_2 = -t_1 = t_0$; in analogia al caso precedente scriveremo l'intervallo di confidenza come $[m - t_0 \cdot s_m, m + t_0 \cdot s_m]$, intendendo che:

$$P(m - t_0 \cdot s_m \leq \mu \leq m + t_0 \cdot s_m) = P_0 \quad (119)$$

ed il valore di t_0 a P_0 fissato si leggono direttamente nella tavola in appendice A.8, come mostrato nell'esempio 4.10. Quando si scrive $m \pm s_m$ oppure $m \pm 1.96 s_m$ ci si riferisce, come prima, ad un intervallo di confidenza $[m - s_m, m + s_m]$ o $[m - 1.96 s_m, m + 1.96 s_m]$, rispettivamente. In questo caso, però, la probabilità corrispondente non è del 68% o del 95% ma, come già detto, dipende dal numero di misure utilizzate per la stima di s ed è prossimo a questi valori solo per grandi campioni (si vedano le tavole relative alla distribuzione di Student).

 METODI DI FIT

 INDICE

5.1	Problemi di fit	117
5.2	Metodo dei minimi quadrati	118
5.2.1	Fit dei minimi quadrati nel caso di una funzione costante	119
5.2.2	Fit dei minimi quadrati nel caso di una funzione lineare	120
5.3	Metodo del minimo χ^2	123
5.3.1	Fit del minimo χ^2 per una funzione costante: la media pesata	123
5.3.2	Fit del minimo χ^2 nel caso lineare	125
5.4	Fit di tipo generale	128
5.5	Test del χ^2	130
5.5.1	Test del χ^2 per una serie di misure	131
5.5.2	Test del χ^2 per una distribuzione	132
5.6	Esercizi	135

Questo capitolo costituisce un'introduzione elementare al problema della stima dei parametri ed alla definizione di criteri quantitativi che permettano di definire il grado di accordo (o di disaccordo) tra un dato modello teorico ed una serie di dati sperimentali.

5.1 PROBLEMI DI FIT

Dato un insieme di dati sperimentali, un problema di *fit* o *best fit* o *fitting* consiste nel cercare una curva *semplice*¹ che si adatti nel migliore dei modi ai punti misurati. Più precisamente, date n misure di due generiche grandezze x ed y , con le rispettive incertezze:

$$(x_i \pm \Delta x_i, y_i \pm \Delta y_i) \quad i = 1 \dots n$$

¹ In generale, dati n punti sperimentali, è sempre possibile trovare un polinomio di grado $n - 1$ che passi *esattamente* per tutti i punti. Ciò non significa che questa sia la scelta migliore e, anzi, in generale non lo è affatto poiché risulterebbe difficoltoso dare una interpretazione fisica agli n coefficienti del polinomio risultante. Spesso, inoltre, ci sono delle ragioni fisiche per preferire una certa famiglia di funzioni piuttosto che un'altra: previsioni della teoria, oppure analogia con altre situazioni simili...

ed una funzione (eventualmente dipendente da un certo numero m di parametri $p_1 \dots p_m$) che pensiamo possa rappresentare bene la dipendenza di y da x :

$$y = f(x; p_1 \dots p_m)$$

si tratta di scegliere i parametri $p_1 \dots p_m$ da cui questa funzione dipende in modo che essa sia la *migliore possibile*.

► **ESEMPIO 5.1.** Supponiamo di misurare, a certi istanti di tempo, la posizione x di un oggetto che si muove con velocità uniforme lungo una linea retta. Se rappresentiamo su di un grafico i punti (t_i, x_i) , eventualmente con i loro errori, il coefficiente angolare della retta di best fit costituisce in questo caso una stima della velocità dell'oggetto stesso. \square

Esistono vari criteri per dire che una curva è la migliore possibile, e tutti si giustificano mediante il *principio di massima verosimiglianza* [3, 4] fondato sull'ipotesi che ciascun punto misurato sia una variabile gaussiana centrata sul misurando e con semilarghezza a metà altezza data dalla incertezza di misura. Vedremo nel seguito alcuni tra i metodi di fit più comunemente usati.

5.2 METODO DEI MINIMI QUADRATI

Il metodo di fit più semplice è il cosiddetto *metodo dei minimi quadrati*. Da un punto di vista operativo definiamo, per ogni punto sperimentale, la differenza tra il valore misurato y_i ed il valore della funzione $f(x; p_1 \dots p_m)$ in corrispondenza del punto x_i , per una generica scelta dei parametri $p_1 \dots p_m$:

$$D_i = y_i - f(x; p_1 \dots p_m)$$

Notiamo che i singoli D_i possono essere maggiori, minori od uguali a zero. Costruiamo quindi la somma:

$$S = \sum_{i=1}^n D_i^2 = \sum_{i=1}^n [y_i - f(x; p_1 \dots p_m)]^2 \quad (120)$$

S costituisce, in un certo qual modo, la *distanza* dall'insieme dei punti sperimentali della funzione con cui eseguiamo il fit. Scegliamo dunque i parametri della funzione stessa in modo da minimizzare S :

$$\begin{cases} \frac{\partial S}{\partial p_1} = \frac{\partial}{\partial p_1} \sum_{i=1}^n [y_i - f(x; p_1 \dots p_m)]^2 = 0 \\ \frac{\partial S}{\partial p_2} = \frac{\partial}{\partial p_2} \sum_{i=1}^n [y_i - f(x; p_1 \dots p_m)]^2 = 0 \\ \vdots \\ \frac{\partial S}{\partial p_m} = \frac{\partial}{\partial p_m} \sum_{i=1}^n [y_i - f(x; p_1 \dots p_m)]^2 = 0 \end{cases} \quad (121)$$

Il principio che sta alla base del metodo dei minimi quadrati tratta tutte le misure y_i sullo stesso piano, cosa che è accettabile solo se le misure stesse *hanno tutte lo stesso errore*:

$$\Delta y_i = \Delta y \quad \forall i = 1 \dots n \quad (122)$$

Inoltre, poiché nella (120) il valore della funzione $f(x; p_1 \dots p_m)$ è calcolato esattamente in corrispondenza dei valori misurati x_i , cioè senza alcuna traccia delle incertezze Δx_i ad essi associate, è pure necessario che queste incertezze siano, in un senso che precisiamo tra un attimo, *trascurabili*. Più precisamente dobbiamo richiedere che per ognuno dei punti misurati l'errore Δx_i , *propagato dalla funzione $f(x)$ sulla variabile dipendente y , sia molto minore dell'errore di misura Δy_i* . In formule:

$$\left| \frac{df}{dx}(x_i) \cdot \Delta x_i \right| \ll \Delta y_i \quad \forall i = 1 \dots n \quad (123)$$

Come risulterà chiaro dalle sezioni seguenti, il metodo dei minimi quadrati dà risultati semplici in tutti i casi in cui la funzione $f(x)$ dipende in modo lineare dai parametri: infatti in questi casi le condizioni di minimo conducono a sistemi di equazioni lineari nei parametri.

5.2.1 Fit dei minimi quadrati nel caso di una funzione costante

Applichiamo il metodo dei minimi quadrati al caso costante:

$$y = f(x; a) = a$$

Notiamo esplicitamente che qui la condizione (123) è automaticamente soddisfatta in quanto

$$\frac{df}{dx}(x) = 0$$

per cui l'unica condizione che dobbiamo richiedere è che gli errori Δy_i siano tutti uguali. Per trovare il miglior valore di a bisogna minimizzare la quantità:

$$S = \sum_{i=1}^n (y_i - a)^2$$

in cui n , lo ricordiamo, è il numero totale di misure effettuate. Dobbiamo quindi derivare rispetto all'unico parametro:

$$\frac{dS}{da} = -2 \sum_{i=1}^n (y_i - a) = 0$$

il che fornisce:

$$\sum_{i=1}^n y_i = \sum_{i=1}^n a = na$$

ed infine:

$$a = \frac{1}{n} \sum_{i=1}^n y_i \quad (124)$$

Non dovrebbe sorprendere che, nel caso in cui le misure possano essere trattate alla pari (abbiano, cioè, la stessa incertezza associata), la media aritmetica rende minima la somma dei quadrati degli scarti (120). L'errore associato alla stima del parametro a si calcola, come di consueto:

$$(\Delta a)^2 = \sum_{i=1}^n \left(\frac{\partial a}{\partial y_i} \right)^2 \cdot (\Delta y)^2 = \sum_{i=1}^n \frac{1}{n^2} \cdot (\Delta y)^2 = \frac{1}{n^2} \cdot (\Delta y)^2 \cdot n = \frac{(\Delta y)^2}{n}$$

Da cui:

$$\Delta a = \frac{\Delta y}{\sqrt{n}} \quad (125)$$

5.2.2 Fit dei minimi quadrati nel caso di una funzione lineare

Applichiamo adesso il metodo dei minimi quadrati al caso lineare:

$$y = f(x; a, b) = a + bx$$

Se siamo nelle condizioni in cui si può applicare il metodo dei minimi quadrati, i migliori valori per a e b si trovano costruendo la somma:

$$S = \sum_{i=1}^n [y_i - f(x_i; a, b)]^2 = \sum_{i=1}^n (y_i - a - bx_i)^2$$

Bisogna cercare il minimo di S al variare di a e b :

$$\begin{cases} \frac{\partial S}{\partial a} = -2 \sum_{i=1}^n (y_i - a - bx_i) = 0 \\ \frac{\partial S}{\partial b} = -2 \sum_{i=1}^n (y_i - a - bx_i) \cdot x_i = 0 \end{cases}$$

Da cui:

$$\begin{cases} na + b \sum_{i=1}^n x_i = \sum_{i=1}^n y_i \\ a \sum_{i=1}^n x_i + b \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i y_i \end{cases}$$

ed infine:

$$\left\{ \begin{array}{l} a = \frac{\sum_{i=1}^n y_i \cdot \sum_{i=1}^n x_i^2 - \sum_{i=1}^n x_i \cdot \sum_{i=1}^n x_i y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \\ b = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \cdot \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \end{array} \right. \quad (126)$$

Gli errori si calcolano, al solito, come:

$$\left\{ \begin{array}{l} (\Delta a)^2 = \sum_{i=1}^n \left(\frac{\partial a}{\partial y_i} \right)^2 \cdot (\Delta y)^2 = \frac{(\Delta y)^2 \cdot \sum_{i=1}^n x_i^2}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \\ (\Delta b)^2 = \sum_{i=1}^n \left(\frac{\partial b}{\partial y_i} \right)^2 \cdot (\Delta y)^2 = \frac{n \cdot (\Delta y)^2}{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2} \end{array} \right.$$

da cui:

$$\left\{ \begin{array}{l} \Delta a = \frac{\Delta y \sqrt{\sum_{i=1}^n x_i^2}}{\sqrt{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}} \\ \Delta b = \frac{\Delta y \sqrt{n}}{\sqrt{n \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i \right)^2}} \end{array} \right. \quad (127)$$

► **ESEMPIO 5.2.** Consideriamo un caso concreto del problema generale esposto nell'esempio 5.1 (cfr. figura 24). I dati nella tabella sono stati simulati con un programma al calcolatore assumendo che l'oggetto si muova lungo l'asse x di moto rettilineo uniforme:

$$x(t) = x_0 + v_0 t$$

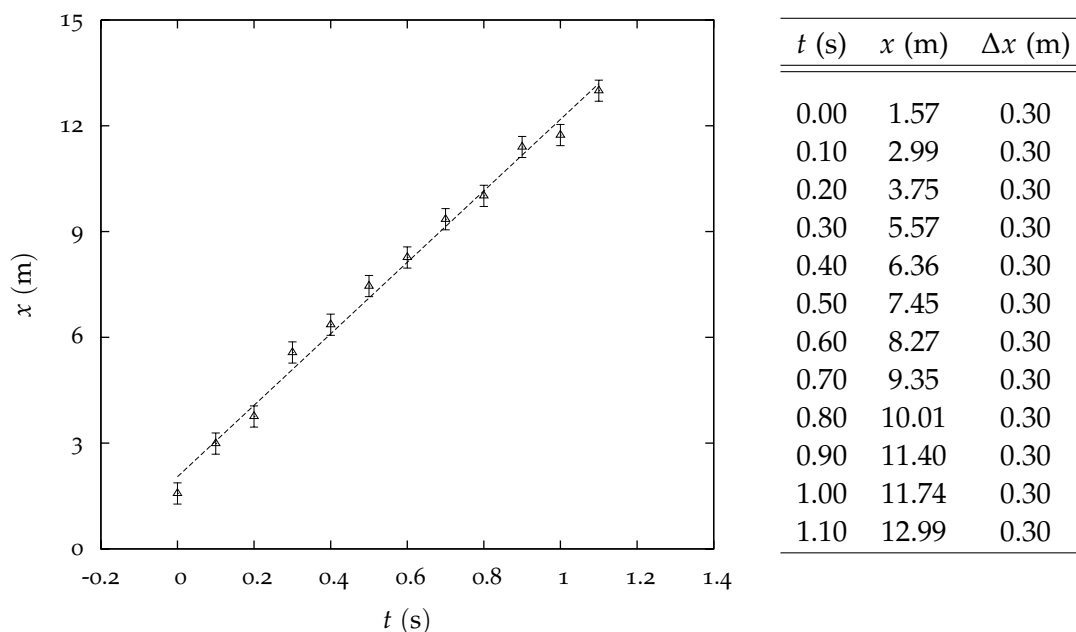


FIGURA 24.. Tabella dei dati relativi all'esempio 5.2 e grafico relativo ai dati stessi. La retta di fit restituita da un fit dei minimi quadrati è riportata sul grafico insieme ai punti sperimentali.

con

$$\begin{aligned}x_0 &= 2.3 \text{ m} \\v_0 &= 10 \text{ m/s}\end{aligned}$$

ed ancora assumendo un errore di misura gaussiano sulle lunghezze con deviazione standard pari a 0.3 m. Le misure di tempo possono essere considerate abbastanza accurate da far sì che la (123) sia verificata.

Notiamo esplicitamente che, avendo generato i dati al calcolatore, ci troviamo nella posizione estremamente privilegiata di conoscere il valore dei misurandi x_0 e v_0 ; cosa che, per ovvie ragioni, non è di norma vera in laboratorio. Questo ci permette di verificare in modo diretto la correttezza di un dato algoritmo di analisi ed è un procedimento usato molto comunemente nel progetto di esperimenti scientifici.

Applicando dunque i risultati appena derivati, un fit dei minimi quadrati (abbiamo già verificato che le ipotesi necessarie siano verificate) restituisce i seguenti valori:

$$\begin{aligned}x_0 &= 2.05 \pm 0.16 \text{ m} \\v_0 &= 10.14 \pm 0.25 \text{ m/s}\end{aligned}$$

che sono statisticamente compatibili (entro due σ) con i valori dei misurandi, noti a priori. \square

5.3 METODO DEL MINIMO χ^2

Supponiamo adesso che le incertezze Δy_i associate alle nostre misure siano, in generale, tutte diverse, mentre manteniamo, per il momento, l'ipotesi che l'errore sulle x_i sia molto piccolo, nel senso della (123). Il metodo dei minimi quadrati introdotto nel paragrafo precedente può essere generalizzato a questa situazione introducendo le quantità:

$$D_i = \frac{y_i - f(x; p_1 \dots p_m)}{\Delta y_i}$$

in cui ogni differenza viene *pesata* con l'inverso dell'errore sulle misure y_i (intuitivamente: maggiore è l'errore di misura, minore è il peso che la misura stessa ha nel fit). Si cercherà poi il minimo della funzione:

$$S = \sum_{i=1}^n D_i^2 = \sum_{i=1}^n \left[\frac{y_i - f(x; p_1 \dots p_m)}{\Delta y_i} \right]^2 \quad (128)$$

imponendo che le derivate rispetto a tutti i parametri siano nulle:

$$\begin{cases} \frac{\partial S}{\partial p_1} = \frac{\partial}{\partial p_1} \sum_{i=1}^n \left[\frac{y_i - f(x; p_1 \dots p_m)}{\Delta y_i} \right]^2 = 0 \\ \frac{\partial S}{\partial p_2} = \frac{\partial}{\partial p_2} \sum_{i=1}^n \left[\frac{y_i - f(x; p_1 \dots p_m)}{\Delta y_i} \right]^2 = 0 \\ \dots \\ \frac{\partial S}{\partial p_m} = \frac{\partial}{\partial p_m} \sum_{i=1}^n \left[\frac{y_i - f(x; p_1 \dots p_m)}{\Delta y_i} \right]^2 = 0 \end{cases} \quad (129)$$

Questo metodo viene indicato come *il metodo del minimo χ^2* . Infatti se si assume che le y_i abbiano una distribuzione gaussiana, le D_i sono variabili gaussiane con media zero e varianza uno ed S è una variabile che ha la distribuzione del χ^2 .

5.3.1 Fit del minimo χ^2 per una funzione costante: la media pesata

Supponiamo di eseguire un certo numero n di misure *indipendenti* di una stessa grandezza y ottenendo n valori y_i , in generali diversi tra loro, con relative incertezze Δy_i , che pure ammettiamo poter essere diverse tra di loro. Ci poniamo il problema di valutare, a partire dai nostri dati, la migliore stima per y e l'incertezza associata. Se potessimo trattare le misure tutte allo stesso modo (cioè se i Δy_i fossero tutti uguali), la migliore stima per y sarebbe costituita dalla media aritmetica delle y_i (cfr. sezione 5.2.1). Ma se questa ipotesi cade è ragionevole aspettarsi, intuitivamente, che le misure affette da incertezza maggiore debbano *contare* meno delle altre. Il problema si riduce, nel nostro linguaggio, ad un fit del minimo χ^2 con una funzione costante:

$$y = f(x; a) = \text{cost} = a$$

Notiamo, per inciso, che la condizione (123) è qui automaticamente soddisfatta in quanto

$$\frac{df}{dx}(x) = 0$$

Costruiamo allora la somma:

$$S = \sum_{i=1}^n \left(\frac{y_i - a}{\Delta y_i} \right)^2$$

ed imponiamo che la derivata rispetto all'unico parametro sia nulla:

$$\frac{\partial S}{\partial a} = -2 \sum_{i=1}^n \frac{1}{\Delta y_i} \cdot \left(\frac{y_i - a}{\Delta y_i} \right) = 0$$

Questo fornisce:

$$\sum_{i=1}^n \frac{y_i}{(\Delta y_i)^2} = \sum_{i=1}^n \frac{a}{(\Delta y_i)^2} = a \sum_{i=1}^n \frac{1}{(\Delta y_i)^2}$$

ed infine:

$$a = \frac{\sum_{i=1}^n \frac{y_i}{(\Delta y_i)^2}}{\sum_{i=1}^n \frac{1}{(\Delta y_i)^2}} \quad (130)$$

L'incertezza associata si valuta come di consueto:

$$(\Delta a)^2 = \sum_{i=1}^n \left(\frac{\partial a}{\partial y_i} \right)^2 (\Delta y_i)^2 = \frac{\sum_{i=1}^n \frac{1}{(\Delta y_i)^4} \cdot (\Delta y_i)^2}{\left[\sum_{i=1}^n \frac{1}{(\Delta y_i)^2} \right]^2} = \frac{\sum_{i=1}^n \frac{1}{(\Delta y_i)^2}}{\left[\sum_{i=1}^n \frac{1}{(\Delta y_i)^2} \right]^2}$$

per cui:

$$\Delta a = \frac{1}{\sqrt{\sum_{i=1}^n \frac{1}{(\Delta y_i)^2}}} \quad (131)$$

► ESEMPIO 5.3. Si eseguono 5 misure indipendenti dell'indice di rifrazione dell'acqua ottenendo i risultati in tabella².

n_i	Δn_i
1.325	0.012
1.3	0.1
1.33	0.01
1.331	0.005
1.335	0.006

² Dati realmente ottenuti negli anni accademici 2002-03, 2003-04, 2004-05.

Si vogliono combinare questi dati in modo da fornire una stima statisticamente corretta di n . Assumendo che le incertezze riportate in tabella siano state valutate correttamente e che gli effetti sistematici siano stati tenuti sotto controllo, il modo corretto di procedere è quello di eseguire una media pesata, secondo le equazioni (130) e (131):

$$n = 1.332 \pm 0.003$$

Notiamo esplicitamente che la media aritmetica degli n_i sarebbe 1.324, che è significativamente più bassa della media pesata (sostanzialmente essa è influenzata dal valore della seconda misura che, essendo affetta da un'incertezza relativa molto più grande degli altri, conta meno nella media pesata). \square

5.3.2 Fit del minimo χ^2 nel caso lineare

In analogia alla sezione 5.2.2 applichiamo il metodo del minimo χ^2 al caso lineare:

$$y = f(x; a, b) = a + bx$$

Cominciamo dal costruire la somma:

$$S = \sum_{i=1}^n \left(\frac{y_i - a - bx_i}{\Delta y_i} \right)^2$$

e deriviamo rispetto ai parametri:

$$\begin{cases} \frac{\partial S}{\partial a} = -2 \sum_{i=1}^n \frac{1}{\Delta y_i} \cdot \left(\frac{y_i - a - bx_i}{\Delta y_i} \right) = 0 \\ \frac{\partial S}{\partial b} = -2 \sum_{i=1}^n \frac{1}{\Delta y_i} \cdot \left(\frac{y_i - a - bx_i}{\Delta y_i} \right) x_i = 0 \end{cases}$$

da cui:

$$\begin{cases} a \sum_{i=1}^n \frac{1}{(\Delta y_i)^2} + b \sum_{i=1}^n \frac{x_i}{(\Delta y_i)^2} = \sum_{i=1}^n \frac{y_i}{(\Delta y_i)^2} \\ a \sum_{i=1}^n \frac{x_i}{(\Delta y_i)^2} + b \sum_{i=1}^n \frac{x_i^2}{(\Delta y_i)^2} = \sum_{i=1}^n \frac{x_i y_i}{(\Delta y_i)^2} \end{cases}$$

Si ha infine:

$$\left\{ \begin{array}{l} a = \frac{\sum_{i=1}^n \frac{y_i}{(\Delta y_i)^2} \cdot \sum_{i=1}^n \frac{x_i^2}{(\Delta y_i)^2} - \sum_{i=1}^n \frac{x_i}{(\Delta y_i)^2} \cdot \sum_{i=1}^n \frac{x_i y_i}{(\Delta y_i)^2}}{\sum_{i=1}^n \frac{1}{(\Delta y_i)^2} \cdot \sum_{i=1}^n \frac{x_i^2}{(\Delta y_i)^2} - \left(\sum_{i=1}^n \frac{x_i}{(\Delta y_i)^2} \right)^2} \\ b = \frac{\sum_{i=1}^n \frac{1}{(\Delta y_i)^2} \cdot \sum_{i=1}^n \frac{x_i y_i}{(\Delta y_i)^2} - \sum_{i=1}^n \frac{y_i}{(\Delta y_i)^2} \cdot \sum_{i=1}^n \frac{x_i}{(\Delta y_i)^2}}{\sum_{i=1}^n \frac{1}{(\Delta y_i)^2} \cdot \sum_{i=1}^n \frac{x_i^2}{(\Delta y_i)^2} - \left(\sum_{i=1}^n \frac{x_i}{(\Delta y_i)^2} \right)^2} \end{array} \right. \quad (132)$$

ed ancora:

$$\left\{ \begin{array}{l} (\Delta a)^2 = \sum_{i=1}^n \left(\frac{\partial a}{\partial y_i} \right)^2 \cdot (\Delta y_i)^2 = \frac{\sum_{i=1}^n \frac{x_i^2}{(\Delta y_i)^2}}{\sum_{i=1}^n \frac{1}{(\Delta y_i)^2} \cdot \sum_{i=1}^n \frac{x_i^2}{(\Delta y_i)^2} - \left(\sum_{i=1}^n \frac{x_i}{(\Delta y_i)^2} \right)^2} \\ (\Delta b)^2 = \sum_{i=1}^n \left(\frac{\partial b}{\partial y_i} \right)^2 \cdot (\Delta y_i)^2 = \frac{\sum_{i=1}^n \frac{1}{(\Delta y_i)^2}}{\sum_{i=1}^n \frac{1}{(\Delta y_i)^2} \cdot \sum_{i=1}^n \frac{x_i^2}{(\Delta y_i)^2} - \left(\sum_{i=1}^n \frac{x_i}{(\Delta y_i)^2} \right)^2} \end{array} \right.$$

da cui:

$$\left\{ \begin{array}{l} \Delta a = \frac{\sqrt{\sum_{i=1}^n \frac{x_i^2}{(\Delta y_i)^2}}}{\sqrt{\sum_{i=1}^n \frac{1}{(\Delta y_i)^2} \cdot \sum_{i=1}^n \frac{x_i^2}{(\Delta y_i)^2} - \left(\sum_{i=1}^n \frac{x_i}{(\Delta y_i)^2} \right)^2}} \\ \Delta b = \frac{\sqrt{\sum_{i=1}^n \frac{1}{(\Delta y_i)^2}}}{\sqrt{\sum_{i=1}^n \frac{1}{(\Delta y_i)^2} \cdot \sum_{i=1}^n \frac{x_i^2}{(\Delta y_i)^2} - \left(\sum_{i=1}^n \frac{x_i}{(\Delta y_i)^2} \right)^2}} \end{array} \right. \quad (133)$$

► **ESEMPIO 5.4.** I dati nella tabella sono stati simulati con lo stesso programma dell'esempio 5.2 con una piccola variante: questa volta gli errori di misura sulle lunghezze non sono costanti ma direttamente proporzionali alle lunghezze stesse, di modo che l'errore relativo è costante e pari al 5% (cfr. figura 25).

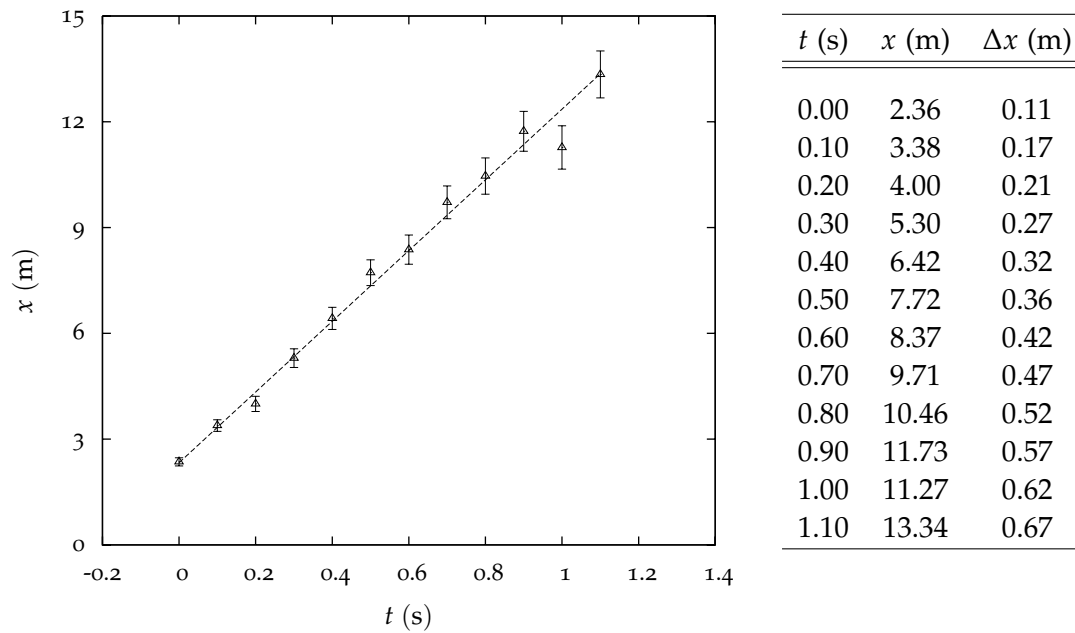


FIGURA 25.. Tabella dei dati relativi all'esempio 5.4 e grafico relativo ai dati stessi. La retta di fit restituita da un fit del minimo χ^2 è riportata sul grafico insieme ai punti sperimentali.

Ammettendo di nuovo che gli errori sulle misure di tempo siano trascurabili, secondo la (123), siamo nelle ipotesi del minimo χ^2 . Il fit del minimo χ^2 in effetti fornisce per i parametri i seguenti valori:

$$\begin{aligned}x_0 &= 2.33 \pm 0.09 \\v_0 &= 10.03 \pm 0.28\end{aligned}$$

che di nuovo sono compatibili con i misurandi.

Se avessimo applicato il metodo dei minimi quadrati—procedimento statisticamente non corretto, in questo caso, dato che gli errori di misura sono diversi tra loro—avremmo ottenuto:

$$\begin{aligned}x_0 &= 2.41 \pm 0.06 \text{ m} \\v_0 &= 9.87 \pm 0.10 \text{ m/s}\end{aligned}$$

In questo caso le stime dei parametri sono leggermente peggiori (come è logico aspettarsi), sia pure ancora ragionevoli; gli errori sui parametri stessi, viceversa, sono un artefatto del fit e non devono essere presi seriamente. \square

5.4 FIT DI TIPO GENERALE

Nei metodi finora discussi abbiamo sempre supposto che l'errore sulle x_i fosse trascurabile, e come tale l'abbiamo trattato. Finora ci siamo quindi occupati di casi relativamente semplici. Relativamente, s'intende, al caso che esamineremo adesso: vogliamo infatti studiare il problema di fitting nel caso più generale in cui non sono trascurabili né gli errori su y né quelli su x .

Il primo problema che si incontra è che in un'equazione del tipo $y = f(x; p_1 \dots p_m)$ non è chiaro quale sia il valore da usare per x in modo da calcolare il valore teorico di y e quindi cercare i migliori valori dei parametri che soddisfino la richiesta di rendere minimo il χ^2 . Non conoscendo x il problema è quindi concettualmente più complicato: bisogna determinare i valori di $(\tilde{x}_i, \tilde{y}_i)$ tali che sia minimo

$$\sum_{i=1}^n \left[\left(\frac{\tilde{x}_i - x_i}{\Delta x_i} \right)^2 + \left(\frac{\tilde{y}_i - y_i}{\Delta y_i} \right)^2 \right] \quad (134)$$

con la condizione che

$$\mathcal{F}(\tilde{x}_i, \tilde{y}_i; p_1 \dots p_m) = 0 \quad \forall i = 1 \dots n \quad (135)$$

dove $\mathcal{F}(x, y; p_1 \dots p_m)$ è la relazione che si vuole che legghi la x alla y scritta in forma implicita.

Si tratta di un problema di minimo condizionato: usando il metodo dei *moltiplicatori di Lagrange* ci si riduce al problema seguente: minimizzare

$$S = \sum_{i=1}^n \left[\left(\frac{\tilde{x}_i - x_i}{\Delta x_i} \right)^2 + \left(\frac{\tilde{y}_i - y_i}{\Delta y_i} \right)^2 + \lambda_i \mathcal{F}(\tilde{x}_i, \tilde{y}_i; p_1 \dots p_m) \right] \quad (136)$$

Le cose da determinare sono $\tilde{x}_i, \tilde{y}_i, \lambda_i, p_1 \dots p_m$: in tutto $3n + m$; ci sono quindi derivate parziali rispetto a $3n + m$ variabili.

Il problema si può risolvere nel caso più generale solo con metodi numerici approssimati. Diamo alcuni cenni sulla tecnica di risoluzione:

- (1) Il primo passo è di conoscere un valore approssimato dei parametri.
- (2) Il secondo passo è di usare invece di $\mathcal{F}(x, y; p_1 \dots p_m)$ la sua approssimazione al prim'ordine. Pertanto ad $\mathcal{F}(x, y; p_1 \dots p_m)$ viene sostituita una funzione lineare.
- (3) Quindi c'è una certa quantità di algebra, sostituzioni, ecc.
- (4) Il risultato è un sistema di equazioni lineari che dicono di quanto devono essere corrette le stime iniziali dei parametri.
- (5) A questo punto si procede con metodo iterativo, cioè si prendono i nuovi valori dei parametri e si ricomincia.
- (6) Se il processo converge alla fine si ottengono i valori richiesti per i parametri.

► **ESEMPIO 5.5.** Analizziamo in dettaglio la procedura generale di fit nel caso puramente lineare³:

$$y = f(x; p) = px$$

La forma implicita della funzione che lega la y alla x si scrive come:

$$\mathcal{F}(x, y; p) = y - px = 0$$

Costruiamo allora la somma (136):

$$S = \sum_{i=1}^n \left[\left(\frac{\tilde{x}_i - x_i}{\Delta x_i} \right)^2 + \left(\frac{\tilde{y}_i - y_i}{\Delta y_i} \right)^2 + \lambda_i (\tilde{y}_i - p \tilde{x}_i) \right]$$

Le condizioni di minimo sono:

$$\begin{cases} \frac{\partial S}{\partial \tilde{x}_i} = \frac{2(\tilde{x}_i - x_i)}{(\Delta x_i)^2} - \lambda_i p = 0 \\ \frac{\partial S}{\partial \tilde{y}_i} = \frac{2(\tilde{y}_i - y_i)}{(\Delta y_i)^2} - \lambda_i = 0 \\ \frac{\partial S}{\partial \lambda_i} = \tilde{y}_i - p \tilde{x}_i = 0 \\ \frac{\partial S}{\partial p} = - \sum_{i=1}^n \lambda_i \tilde{x}_i \approx - \sum_{i=1}^n \lambda_i x_i = 0 \end{cases}$$

Dove nell'ultima relazione si è linearizzata la \mathcal{F} . Si ricavano \tilde{x}_i ed \tilde{y}_i dalle prime due relazioni:

$$\begin{cases} \tilde{x}_i = \frac{\lambda_i p \cdot (\Delta x_i)^2}{2} + x_i \\ \tilde{y}_i = \frac{\lambda_i \cdot (\Delta y_i)^2}{2} + y_i \end{cases}$$

e si sostituiscono nella terza:

$$\frac{\lambda_i \cdot (\Delta y_i)^2}{2} + y_i - \frac{\lambda_i p^2 \cdot (\Delta x_i)^2}{2} - x_i = 0$$

da cui si ricava λ_i :

$$\lambda_i = \frac{2(y_i - px_i)}{p^2 \cdot (\Delta x_i)^2 + (\Delta y_i)^2}$$

che, sostituito nella quarta relazione dà:

$$p' = \frac{\sum_{i=1}^n \frac{x_i y_i}{p^2 \cdot (\Delta x_i)^2 + (\Delta y_i)^2}}{\sum_{i=1}^n \frac{x_i^2}{p^2 \cdot (\Delta x_i)^2 + (\Delta y_i)^2}} \quad (137)$$

³ Per le funzioni che, come questa, sono lineari nei parametri, potrebbero essere usati metodi più semplici (provate). Tuttavia la semplicità dell'esempio permette di seguire meglio la tecnica proposta.

a questo punto entra in funzione il metodo iterativo: si stima p (ad esempio dal grafico) e si usa questa stima nel secondo membro della (137) ottenendo un nuovo p (chiamato p'), si ripete il procedimento fino a quando la differenza fra i successivi valori di p è entro l'errore su p .

Per stimare gli errori su p si usano l'equazione (137) e le tecniche della propagazione degli errori ed il valore stimato di p :

$$(\Delta p)^2 = \sum_{i=1}^n \left[\left(\frac{\partial p}{\partial x} \cdot \Delta x_i \right)^2 + \left(\frac{\partial p}{\partial y} \cdot \Delta y_i \right)^2 \right] = \frac{1}{\sum_{i=1}^n \frac{x_i^2}{p^2 \cdot (\Delta x_i)^2 + (\Delta y_i)^2}}$$

□

5.5 TEST DEL χ^2

Richiamiamo la definizione della variabile χ^2 a n gradi di libertà: date n variabili x_i gaussiane con media μ_i e deviazione standard σ_i e costruite le variabili standard

$$z_i = \frac{x_i - \mu_i}{\sigma_i}$$

la quantità

$$S = \sum_{i=1}^n z_i^2 = \sum_{i=1}^n \left(\frac{x_i - \mu_i}{\sigma_i} \right)^2 \quad (138)$$

è distribuita secondo la funzione di distribuzione del χ^2 ad n gradi di libertà, che è data dalla (82). La media e la varianza valgono rispettivamente:

$$\begin{aligned} \mu_S &= n \\ \sigma_S^2 &= 2n \end{aligned}$$

Fissato il numero di gradi di libertà n , le tavole riportate nelle appendici A.4 e A.5 forniscono il valore γ per cui la probabilità che S sia minore di γ sia uguale ad un certo valore \wp :

$$P(S \leq \gamma) = \wp$$

► ESEMPIO 5.6. Sia $n = 10$; ci si domanda quale sia il valore γ tale che:

$$P(S \leq \gamma) = 95\%$$

Il risultato si legge direttamente nella tavola in appendice A.4 ed è:

$$\gamma = 18.308$$

□

5.5.1 Test del χ^2 per una serie di misure

Il χ^2 può fornire un criterio generale ed obiettivo per decidere se una certa equazione o una certa legge descriva bene oppure no i risultati sperimentali. Supponiamo, al solito, di aver misurato le quantità (x_i, y_i) legate tra di loro da una funzione $y = f(x; p_1 \dots p_m)$, la cui forma viene ipotizzata in base a considerazioni fisiche, ed i cui parametri $p_1 \dots p_m$ sono determinati con uno dei metodi di fit discussi prima, per esempio il metodo dei minimi quadrati o del minimo χ^2 . Il χ^2 in questo caso vale

$$S = \sum_{i=1}^n \left[\frac{y_i - f(x_i; p_1 \dots p_m)}{\Delta y_i} \right]^2$$

La quantità S viene spesso indicata con il nome di *somma dei residui*; intuitivamente è chiaro che, quanto più S è grande, tanto più la curva teorica si discosta dai punti sperimentali.

Supposto che gli errori siano stati valutati in modo corretto, andiamo a vedere qual è il significato da attribuire al numero S . Si cerca sulle tavole in corrispondenza al numero di gradi di libertà dove si trova il valore S . Quando si fa un fit il numero di gradi di libertà ν è il numero n delle misure meno il numero m di parametri della funzione $f(x; p_1 \dots p_m)$

$$\nu = n - m \quad (139)$$

► ESEMPIO 5.7. Supponiamo di aver ottenuto $S = 7.2$ con $\nu = 5$. Dalla tavola in appendice A.4 si ottiene:

$$P(\chi^2 \leq S) \approx 80\%$$

Questo significa che ripetendo le misure tante volte ci si aspetta che nell'80% dei casi si otterrà un valore minore di 7.2; cioè 7.2 è un valore non molto buono per il χ^2 , ma nemmeno assurdo, nel 20% dei casi può succedere che venga un risultato più alto.

Per inciso: $P(\chi^2 \geq S)$ viene solitamente chiamato livello di significatività ed è ovviamente:

$$P(\chi^2 \geq S) = 1 - P(\chi^2 \leq S)$$

□

► ESEMPIO 5.8. Nel caso dell'esempio 5.2 il valore di S restituito dal fit vale:

$$S = 12.47 \quad (\nu = 10)$$

Il livello di significatività del fit è dunque circa il 75%, come si legge in appendice A.4. □

► ESEMPIO 5.9. Nel caso dell'esempio 5.4 il valore di S restituito dal fit del minimo χ^2 vale:

$$S = 8.00 \quad (\nu = 10)$$

mentre quello restituito dal fit dei minimi quadrati:

$$S = 8.76 \quad (\nu = 10)$$

Il lettore non dovrebbe essere stupito dal fatto che S sia più piccolo nel caso del fit del minimo χ^2 (i nomi non sono in genere scelti a caso). \square

Convenzionalmente si accetta come limite $P(\chi^2 \leq S)$ per i risultati buoni quello del 95% da un lato, e quello del 5% dall'altro, perché, anche se a prima vista sembra che la cosa migliore sia ottenere dei χ^2 molto piccoli (o addirittura zero), in realtà questi casi vanno guardati con sospetto. Di solito χ^2 piccoli si ottengono perché gli errori sono sovrastimati, oppure i risultati sono truccati, cioè lo sperimentatore (consapevolmente o no) tende ad attribuire al risultato delle misure il valore che gli piacerebbe. Esistono tuttavia altri livelli di significatività a seconda dei particolari problemi studiati. È comunque buona regola indicare sempre quale è il valore del χ^2 ottenuto e quindi la probabilità del χ^2 associata alle ipotesi fatte ed ai parametri dedotti dalle misure.

5.5.2 Test del χ^2 per una distribuzione

Supponiamo di fare un esperimento che abbia come oggetto l'osservazione di n possibili eventi $E_1 \dots E_n$ e registriamo le occorrenze $o_1 \dots o_n$ di ciascuno di questi possibili eventi⁴. D'altra parte possiamo fare certe ipotesi sul fenomeno che stiamo studiando le quali ci permettono di determinare le occorrenze teoriche⁵ $e_1 \dots e_n$. Ci saranno differenze tra le occorrenze osservate e quelle teoriche. Per determinare se le differenze sono significative (e quindi eventualmente rifiutare le ipotesi che hanno portato a calcolare gli e_i) si usa la quantità

$$S = \sum_{i=1}^n \frac{(o_i - e_i)^2}{e_i} \quad (140)$$

che viene considerata ancora una variabile χ^2 . Formalmente ricorda un χ^2 perché si suppone che sia Poissoniana la distribuzione di ognuna delle occorrenze o_i : e_i rappresenta il valor medio della parent distribution e per una variabile Poissoniana la varianza coincide con la media, per cui $\sigma_{o_i}^2 = e_i$. È tuttavia importante ricordare che o_i ed e_i non sono variabili gaussiane; o meglio: lo sono soltanto in condizioni limite, cioè per o_i (e quindi e_i) grande. In queste condizioni la (140) tende ad una variabile χ^2 ; $o_i > 3$ è una condizione sufficiente per la maggior parte dei casi che si incontrano in pratica.

I gradi di libertà sono $n - m - 1$ dove n è il numero delle classi (cioè il numero degli eventi possibili) ed m è il numero dei parametri che servono per calcolare le occorrenze teoriche $e_1 \dots e_n$; inoltre deve essere soddisfatta la seguente relazione:

$$\sum_{i=1}^n o_i = N$$

dove N è il numero totale delle osservazioni, e questa relazione causa la perdita di un ulteriore grado di libertà. Nel limite in cui la quantità (140) può essere considerata un χ^2 ad essa si attribuisce lo stesso significato probabilistico trattato precedentemente.

⁴ Le occorrenze osservate vengono generalmente indicate con la lettera o dall'inglese *observed*, che significa appunto *osservato*.

⁵ Le occorrenze teoriche si indicano solitamente con la lettera e dall'inglese *expected* che significa *atteso*.

► **ESEMPIO 5.10.** Si supponga di avere un foglio di carta diviso in 64 quadrati numerati (da 1 a 64). Si prendono 64 dischi e si lasciano cadere sul foglio di carta uno alla volta, in modo che si distribuiscano pressoché uniformemente sul quadrato. Quindi si fa una tavola suddividendo i quadrati in funzione del numero dei dischi che essi contengono:

k	o_k	e_k
0	21	23.54
1	26	23.54
2	13	11.77
3	4	3.22
4	0	0.38

Per calcolare le frequenze teoriche, si è fatta l'ipotesi che la distribuzione della variabile k sia una Poissoniana con media 1 (64 dischi in 64 quadrati cioè in media un disco per quadrato):

$$P(k) = \frac{1^k}{k!} e^{-1} = \frac{1}{k!} \cdot \frac{1}{e}$$

Le frequenze teoriche si calcolano moltiplicando $P(k)$ per il numero totale $n = 64$ dei quadrati:

$$e_k = \frac{1}{k!} \cdot \frac{64}{e}$$

Calcoliamo il valore S del χ^2 :

$$\begin{aligned} S &= \sum_{i=0}^3 \frac{(o_i - e_i)^2}{e_i} = \\ &= \frac{(21 - 23.54)^2}{23.54} + \frac{(26 - 23.54)^2}{23.54} + \frac{(13 - 11.77)^2}{11.77} + \frac{(4 - 5.15)^2}{5.15} \approx 0.91 \end{aligned}$$

I gradi di libertà sono $4 - 0 - 1 = 3$ (qui nessun parametro è stato ricavato dalle misure: questa è una Poissoniana limite di una Binomiale in cui sia n che p sono noti). Utilizzando le tavole si trova:

$$P(\chi^2 \leq S) \approx 22\%.$$

□

► **ESEMPIO 5.11.** L'istogramma in figura 26 si riferisce ai numeri del gioco del lotto usciti nella ruota di Bari dal 1937 ad oggi. Per ognuna delle possibili uscite k ($k = 1 \dots 90$) è riportato il numero di volte in cui essa si è verificata.

I dati si riferiscono a 4223 estrazioni; ricordiamo che ogni volta vengono estratti 5 numeri (senza rimpiazzo) per cui ogni numero ha una probabilità di essere estratto pari a $\frac{5}{90}$. Se il gioco non è truccato ci aspettiamo che il numero medio e_k di uscite per il numero k in 4223 estrazioni sia uguale per tutti e pari a:

$$e_k = 4223 \cdot \frac{5}{90} \approx 234.6 \quad k = 1 \dots 90$$

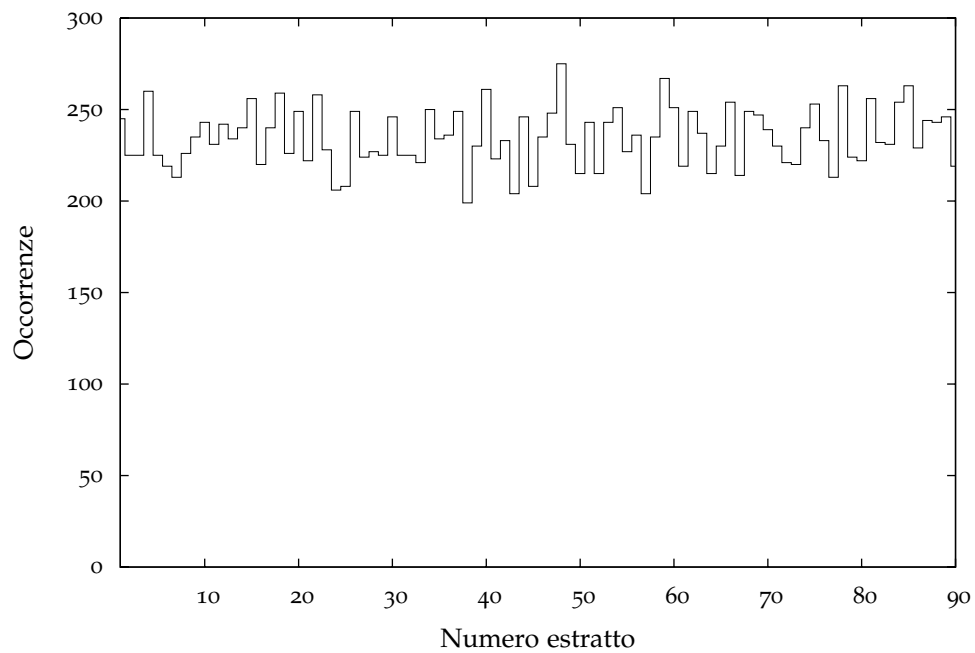


FIGURA 26.. Istogramma in frequenza delle uscite del gioco del lotto sulla ruota di Bari dal 1937 ad oggi.

Un test del χ^2 su questa distribuzione restituisce una somma dei residui pari a:

$$S = 98.313$$

Il numero di gradi di libertà è qui $\nu = 90 - 0 - 1$ per cui il livello di significatività è pari a circa il 75%. Con i dati in nostro possesso dobbiamo accettare l'ipotesi che il gioco non sia truccato (per lo meno nel senso del χ^2). \square

5.6 ESERCIZI

- ESERCIZIO 5.1. Seguendo le sezioni 5.2.1 e 5.2.2 ricavare con il metodo dei minimi quadrati il miglior valore di a per la funzione:

$$f(x, a) = ax$$

- ESERCIZIO 5.2. Seguendo le sezioni 5.2.1 e 5.2.2 ricavare con il metodo dei minimi quadrati il miglior valore di a per la funzione:

$$f(x, a) = ax^2$$

- ESERCIZIO 5.3. Seguendo le sezioni 5.2.1 e 5.2.2 ricavare con il metodo dei minimi quadrati il miglior valore di a per la funzione:

$$f(x) = ae^{-x}$$

- ESERCIZIO 5.4. Calcolare esplicitamente la media pesata (con relativa incertezza) e la media aritmetica dei valori dell'indice di rifrazione dell'acqua riportati nell'esempio 5.3 e verificare che i risultati dell'esempio stesso sono corretti.

- ESERCIZIO 5.5. Con riferimento all'esempio 5.2, eseguire un fit dei minimi quadrati utilizzando i risultati ottenuti nella sezione 5.2.2 e verificare che i risultati coincidano con quelli forniti alla fine dell'esercizio stesso—i dati sono riportati nella tabella in figura 24.

- ESERCIZIO 5.6. Con riferimento all'esempio 5.4, eseguire un fit del minimo χ^2 utilizzando i risultati ottenuti nella sezione 5.3.2 e verificare che i risultati coincidano con quelli forniti alla fine dell'esercizio stesso—i dati sono riportati nella tabella in figura 25.

- ESERCIZIO 5.7. Discutere se per i dati riportati nell'esercizio 2.8 è accettabile l'ipotesi di distribuzione Gaussiana (qual è il livello di significatività).

 FORMULE APPROSSIMATE

 INDICE

6.1	Derivata di una funzione data per punti	137
6.2	Integrale di una funzione data per punti	139
6.3	Interpolazione	140

Trattiamo in questo capitolo alcuni problemi connessi alla *manipolazione* (integrazione, derivazione, interpolazione) di funzioni tabulate—funzioni cioè di cui si conoscono i valori solamente in un determinato numero di punti discreti dell'intervallo di variabilità.

6.1 DERIVATA DI UNA FUNZIONE DATA PER PUNTI

Supponiamo di conoscere i valori di una generica funzione $f(x)$ in un certo numero di punti x_i egualmente spazati di una quantità h . Per fissare le idee supponiamo di avere una tabella con $2n + 1$ valori di $f(x)$, in corrispondenza di:

$$x_i = x_0 + i \cdot h \quad i = -n \dots n$$

Supponiamo ancora che la funzione sia *derivabile* su tutto l'intervallo delle x a cui siamo interessati, e che h sia abbastanza piccolo in modo che la funzione non compia forti oscillazione all'interno di ciascun intervallo $[x_i, x_{i+1}]$. In queste condizioni avrà senso usare le formule approssimate che ricaveremo.

Per semplicità indicheremo nel seguito con f_i il valore della funzione nel punto x_i :

$$f_i = f(x_i)$$

ed useremo la seguente notazione per le derivate:

$$\left\{ \begin{array}{l} f'_i = \frac{df}{dx}(x_i) \\ f''_i = \frac{d^2f}{dx^2}(x_i) \\ f'''_i = \frac{d^3f}{dx^3}(x_i) \\ f^{iv}_i = \frac{d^4f}{dx^4}(x_i) \end{array} \right.$$

Il valore f_{i+m} di $f(x)$ in un generico punto $x_i + mh$ si può esprimere in termini di $f(x)$ e delle sue derivate nel punto x_i per mezzo di uno sviluppo in serie di Taylor intorno ad x_i :

$$f_{i+m} = f(x_i + mh) = f_i + mh f'_i + \frac{(mh)^2}{2!} f''_i + \frac{(mh)^3}{3!} f'''_i + \dots$$

Quindi per $m = 1$ si ha:

$$f_{i+1} = f(x_i + h) = f_i + h f'_i + \frac{h^2}{2} f''_i + \frac{h^3}{6} f'''_i + \dots$$

da cui si può scrivere facilmente lo sviluppo completo per la derivata di $f(x)$ in x_i :

$$f'_i = \frac{f_{i+1} - f_i}{h} - \left(\frac{h}{2} f''_i + \frac{h^2}{6} f'''_i + \dots \right)$$

Trascurando i termini tra parentesi si ottiene un'espressione *utilizzabile* per f'_i :

$$\boxed{\begin{aligned} f'_i &\simeq \frac{f_{i+1} - f_i}{h} \\ \Delta f'_i &\simeq -\frac{h}{2} f''_i \end{aligned}} \quad (141)$$

dove per $\Delta f'_i$ si prende il primo dei termini tra parentesi, che solitamente costituisce la correzione dominante allo sviluppo.

Per $m = -1$ si ottiene:

$$f_{i-1} = f(x_i - h) = f_i - h f'_i + \frac{h^2}{2} f''_i - \frac{h^3}{6} f'''_i + \dots$$

Sottraendo f_{i-1} da f_{i+1} si ottiene un'altra espressione approssimata per la derivata prima:

$$\boxed{\begin{aligned} f'_i &\simeq \frac{f_{i+1} - f_{i-1}}{2h} \\ \Delta f'_i &\simeq -\frac{h^2}{6} f'''_i \end{aligned}} \quad (142)$$

Sommando f_{i+1} con f_{i-1} si ottiene infine un'espressione approssimata per f''_i :

$$\boxed{\begin{aligned} f''_i &\simeq \frac{f_{i+1} - 2f_i + f_{i-1}}{h^2} \\ \Delta f''_i &\simeq -\frac{h^2}{12} f^{iv}_i \end{aligned}} \quad (143)$$

Sempre mediante uno sviluppo in serie di Taylor si possono ottenere espressioni approssimate anche per derivate di ordine superiore con diversi gradi di approssimazione [11].

6.2 INTEGRALE DI UNA FUNZIONE DATA PER PUNTI

L'integrale di una funzione $f(x)$ si calcola integrando il suo sviluppo in serie di Taylor termine per termine. Indichiamo con I_i l'integrale della funzione $f(x)$ calcolato tra i due estremi x_i ed x_{i+1} :

$$I_i = \int_{x_i}^{x_{i+1}} f(x) dx = \int_{x_i}^{x_i+h} f(x) dx$$

Mediante il cambiamento di variabile $z = x - x_i$ possiamo riscrivere la relazione precedente come:

$$I_i = \int_0^h f(x_i + z) dz$$

e valutare l'integrale utilizzando lo sviluppo in serie:

$$f(x_i + z) = f_i + zf'_i + \frac{z^2}{2}f''_i + \frac{z^3}{6}f'''_i + \dots$$

Integrando termine per termine si ottiene:

$$\begin{aligned} \int_0^h f(x_i + z) dz &= \int_0^h f_i dz + \int_0^h zf'_i dz + \int_0^h \frac{z^2}{2}f''_i dz + \dots \\ &= hf_i + \frac{h^2}{2}f'_i + \frac{h^3}{6}f''_i + \dots \end{aligned}$$

Sostituiamo a f'_i la formula a due punti della derivata, data dalla 141:

$$I_i = hf_i + \frac{h^2}{2} \left(\frac{f_{i-1} - f_{i+1}}{2h} - \frac{h}{2}f''_i + \dots \right) + \frac{h^3}{6}f''_i + \dots$$

da cui, trascurando i termini in h^3 e superiori, si ottiene

$$\begin{aligned} I_i &\simeq hf_i + \frac{h^2}{2} \cdot \frac{f_{i+1} - f_{i-1}}{2h} \\ \Delta I_i &\simeq -\frac{h^3}{12}f''_i \end{aligned}$$

(144)

Questa è la nota regola del trapezio. Per avere l'integrale su n passi si sommano gli integrali fatti su ciascun passo.

L'integrale fatto in questo modo non converge molto rapidamente, cioè bisogna prendere molti intervallini per avere una buona approssimazione. Molto usata è invece la *formula di Simpson* che andiamo adesso a ricavare. Indichiamo con J_i l'integrale della funzione $f(x)$ calcolato tra i due estremi x_{i-1} ed x_{i+1} :

$$J_i = \int_{x_{i-1}}^{x_{i+1}} f(x) dx = \int_{x_i-h}^{x_i+h} f(x) dx$$

Facciamo il solito cambiamento di variabile $z = x - x_i$:

$$J_i = \int_{-h}^h f(x_i + z) dz$$

e, come prima, integriamo termine a termine:

$$\begin{aligned} \int_{-h}^h f(x_i + z) dz &= \int_{-h}^h f_i dz + \int_{-h}^h z f'_i dz + \int_{-h}^h \frac{z^2}{2} f''_i dz + \int_{-h}^h \frac{z^3}{6} f'''_i dz + \int_{-h}^h \frac{z^4}{24} f^{iv}_i dz + \dots \\ &= 2hf_i + 0 + 2 \cdot \frac{h^3}{6} f''_i + 0 + 2 \cdot \frac{h^5}{120} f^{iv}_i + \dots \end{aligned}$$

Facendo uso dell'espressione (143) per la derivata seconda possiamo allora scrivere:

$$\int_{-h}^h f(x_i + z) dz = 2hf_i + \frac{h^3}{3} \cdot \frac{1}{h^2} (f_{i+1} - 2f_i + f_{i-1}) - \left(\frac{h^3}{3} \cdot \frac{h^2}{12} f^{iv}_i - \frac{h^5}{60} f^{iv}_i \right) + \dots \quad (145)$$

da cui otteniamo infine:

$$\begin{aligned} J_i &\simeq \frac{h}{3} (f_{i-1} + 4f_i + f_{i+1}) \\ \Delta J_i &\simeq -\frac{h^5}{90} f^{iv}_i \end{aligned} \quad (146)$$

che, come anticipato, è proprio la formula di Simpson.

6.3 INTERPOLAZIONE

L'*interpolazione* consiste operativamente nella stima del valore della funzione $f(x)$ in un punto x compreso tra due punti successivi x_i ed x_{i+1} della tabella. La stima di $f(x)$ al di fuori dell'intervallo coperto dalla tabella prende invece il nome di *estrapolazione*.

Il metodo più semplice è quello di approssimare la curva ad una retta, scrivere l'equazione della retta che passa per x_i ed x_{i+1} e quindi calcolare $f(x)$ nel punto desiderato. Nel nostro linguaggio consideriamo lo sviluppo in serie di $f(x)$ intorno a x_i troncato al prim'ordine:

$$f(x) = f_i + (x - x_i) \cdot f'_i + \dots$$

Utilizziamo la relazione (141) per la derivata prima otteniamo la formula di *interpolazione lineare*:

$$f(x) \simeq f_i + (x - x_i) \cdot \frac{f_{i+1} - f_i}{h} \quad (147)$$

Se invece lo sviluppo in serie viene troncato al second'ordine e f''_i viene approssimata con la (143) si ottiene la formula di *interpolazione quadratica a tre punti*:

$$f(x) \simeq f_i + (x - x_i) \cdot \frac{f_{i+1} - f_{i-1}}{2h} + (x - x_i)^2 \cdot \frac{f_{i+1} - 2f_i + f_{i-1}}{2h^2} \quad (148)$$

 COMPLEMENTI

 INDICE

7.1	Correlazione	141
7.1.1	Coefficiente di correlazione	143
7.2	La distribuzione di Student	145
7.2.1	Variabile t di Student a n gradi di libertà	145
7.2.2	Variabile t di Student a $n-1$ gradi di libertà	146
7.2.3	t -test	147
7.3	Funzione di distribuzione per funzioni di variabile casuale	149
7.3.1	Variabili casuali discrete	149
7.3.2	Variabili casuali continue	150
7.4	Esercizi	157

7.1 CORRELAZIONE

Dato un insieme di dati (x_i, y_i) , non evidentemente legati da una legge fisica, se osserviamo che grandi valori della variabile x tendono ad associarsi con grandi valori della variabile y e piccoli valori di x con piccoli valori di y o viceversa, diciamo che vi è una *correlazione* tra x e y . Un diagramma (*scatter diagram*) in cui vengano riportate le coppie ordinate (x_i, y_i) indica *visivamente* se esiste una qualche correlazione tra questi dati, ma è necessario avere una misura quantitativa di tale fatto. Bisogna quindi sviluppare una statistica che possa misurare il grado di connessione o meglio di correlazione tra x ed y .

► **ESEMPIO 7.1.** Si misura la lunghezza l (con un calibro Palmer) e la massa m (con una bilancia di precisione) di ciascun elemento di un campione di 70 chiodini¹. La seguente tabella include i risultati ottenuti (tutte le lunghezze sono in cm e le masse in mg)

¹ Dati realmente ottenuti il 6 Agosto 2005.

LUNGHEZZA, MASSA (cm, g)				
(1.731, 276)	(1.728, 281)	(1.724, 277)	(1.679, 274)	(1.742, 278)
(1.681, 274)	(1.737, 280)	(1.724, 279)	(1.712, 274)	(1.661, 266)
(1.705, 274)	(1.722, 280)	(1.707, 274)	(1.706, 275)	(1.739, 281)
(1.696, 275)	(1.729, 279)	(1.756, 282)	(1.723, 282)	(1.657, 267)
(1.717, 278)	(1.706, 276)	(1.731, 280)	(1.685, 274)	(1.730, 278)
(1.724, 276)	(1.713, 278)	(1.726, 279)	(1.698, 275)	(1.714, 275)
(1.714, 284)	(1.689, 273)	(1.712, 280)	(1.663, 266)	(1.736, 280)
(1.699, 275)	(1.729, 275)	(1.695, 272)	(1.723, 278)	(1.729, 281)
(1.722, 281)	(1.732, 279)	(1.700, 277)	(1.697, 275)	(1.703, 278)
(1.743, 278)	(1.729, 280)	(1.694, 272)	(1.709, 276)	(1.701, 277)
(1.703, 276)	(1.703, 281)	(1.716, 276)	(1.695, 276)	(1.695, 277)
(1.703, 275)	(1.697, 272)	(1.697, 274)	(1.696, 274)	(1.747, 280)
(1.739, 277)	(1.693, 272)	(1.705, 274)	(1.708, 277)	(1.698, 274)
(1.740, 280)	(1.711, 276)	(1.721, 277)	(1.649, 264)	(1.707, 277)

È abbastanza immediato verificare, a partire dai numeri nella tabella, che a valori grandi (piccoli) di lunghezza tendono a corrispondere valori grandi (piccoli) della massa misurato. Questo risulta ancora più evidente dallo scatter plot dei dati mostrato in figura 27.

Il fatto che le due serie siano correlate non dovrebbe stupire affatto: in fondo è perfettamente normale che i chiodini più lunghi siano anche i più pesanti... Ma torneremo ancora su questo esempio nel seguito. □

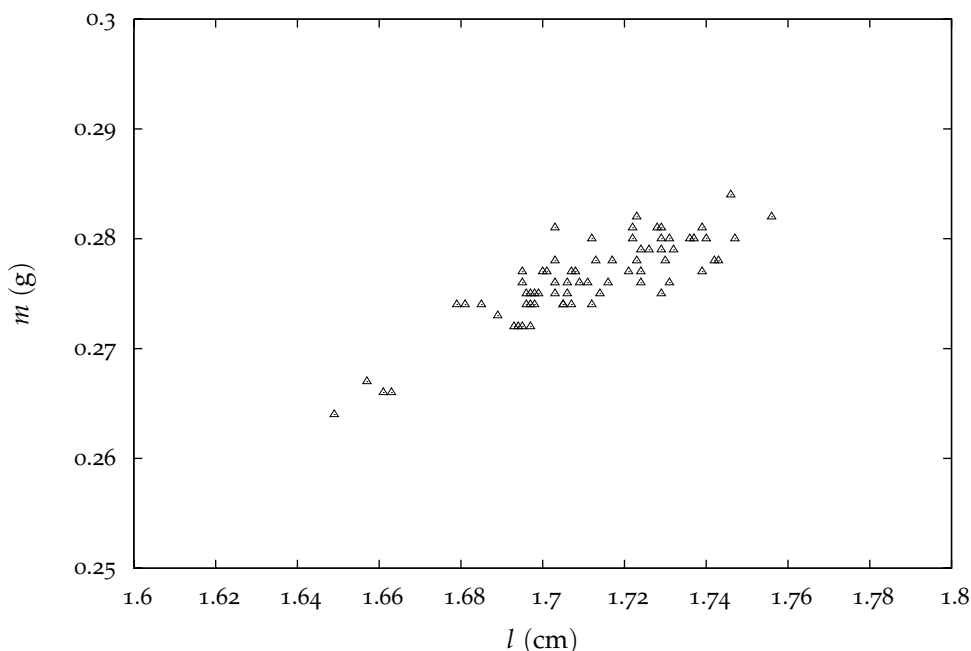


FIGURA 27.. Scatter plot dei dati relativi all'esempio 7.1 (misura di lunghezza e massa di un campione di 70 chiodini).

7.1.1 Coefficiente di correlazione

Si definisce coefficiente di correlazione r_{xy} tra due qualsiasi variabili casuali x ed y la quantità:

$$r_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} \quad (149)$$

dove σ_{xy} è la covarianza tra x ed y

$$\sigma_{xy} = E[(x - \mu_x) \cdot (y - \mu_y)]$$

e σ_x^2 e σ_y^2 sono le varianze di x ed y :

$$\begin{aligned} \sigma_x^2 &= E[(x - \mu_x)^2] = E[x^2] - \mu_x^2 \\ \sigma_y^2 &= E[(y - \mu_y)^2] = E[y^2] - \mu_y^2 \end{aligned}$$

(ricordiamo che μ_x e μ_y rappresentano il valor medio di x ed y , rispettivamente). Elenchiamo qui di seguito alcune osservazioni che ci saranno utili nel seguito.

- (1) r_{xy} è una quantità adimensionale.
- (2) Se le variabili x e y sono statisticamente indipendenti, allora $r_{xy} = 0$; infatti in tal caso $\sigma_{xy} = 0$. L'implicazione inversa non è necessariamente verificata: se la covarianza è uguale a zero non è detto che le variabili siano statisticamente indipendenti. In effetti la condizione $\sigma_{xy} = 0$ è necessaria ma non sufficiente per l'indipendenza statistica di due variabili.
- (3) Quando $r_{xy} > 0$ si dice che le variabili sono *positivamente* correlate; quando, al contrario, $r_{xy} < 0$ si dice le variabili sono *negativamente* correlate o anche *anticorrelate*.
- (4) Se le variabili casuali x ed y sono legate da una relazione funzionale del tipo $y = a + bx$ allora

$$r_{xy} = \pm 1$$

a seconda che b sia positivo o negativo. In tutti gli altri casi

$$-1 \leq r_{xy} \leq 1.$$

In generale quanto più $|r_{xy}|$ è piccolo, tanto più la correlazione è debole.

- (5) Date le variabili casuali

$$\begin{aligned} x' &= a + bx \\ y' &= c + dy \end{aligned}$$

si trova facilmente che:

$$r_{x'y'} = r_{xy}$$

purché b e d siano entrambe positive o entrambe negative. Questo ci dice che la correlazione non cambia quando i dati $\{x_i\}$ o $\{y_i\}$ sono traslati ciascuno della stessa

quantità o quando sono moltiplicati ciascuno per la stessa costante. Ne consegue, in particolare, che il coefficiente di correlazione non dipende dalle unità scelte per misurare i dati.

Per un campione costituito da un set di dati sperimentali (x_i, y_i) il coefficiente di correlazione si scrive :

$$\begin{aligned}
 r_{xy} &= \frac{\sum_{i=1}^n (x_i - m_x) \cdot (y_i - m_y)}{(n-1) \cdot s_x s_y} = \\
 &= \frac{\sum_{i=1}^n (x_i - m_x) \cdot (y_i - m_y)}{\sqrt{\sum_{i=1}^n (x_i - m_x)^2 \cdot \sum_{i=1}^n (y_i - m_y)^2}}.
 \end{aligned} \tag{150}$$

dove m_x , m_y , s_x ed s_y sono le medie e le varianze campione, rispettivamente. Anche in questo caso, ovviamente, valgono tutte le proprietà elencate sopra.

► **ESEMPIO 7.2.** Tornando all'esempio 7.1, adesso abbiamo gli strumenti per precisare ciò che abbiamo detto in precedenza. Il coefficiente di correlazione tra le due serie di dati si può facilmente calcolare utilizzando la (150) e risulta essere

$$r_{lm} \approx 0.849.$$

Dunque, in questo caso, la lunghezza e la massa dei chiodini del campione sono effettivamente due grandezze positivamente correlate, come del resto appariva già evidente, sia pure a livello qualitativo, dallo scatter diagram. Questa correlazione è inoltre piuttosto stretta, essendo r_{lm} prossimo ad 1. \square

Vale la pena notare che la correlazione misura l'associazione tra variabili casuali, non la causa. Quindi non bisogna fare deduzioni arbitrarie quando si applicano questi concetti a situazioni complesse (inerenti per esempio la natura umana, l'ambiente, la società...). La spiegazione di una associazione osservata può stare in fattori inespressi che magari coinvolgono entrambe le variabili in esame.

Ricordiamo che, in generale, l'incertezza associata ad una grandezza G che sia funzione di due quantità *non* statisticamente indipendenti

$$G = f(x, y)$$

si scrive come:

$$\sigma_G^2 = \left(\frac{\partial f}{\partial x}\right)^2 \cdot \sigma_x^2 + \left(\frac{\partial f}{\partial y}\right)^2 \cdot \sigma_y^2 + 2 \cdot \frac{\partial f}{\partial x} \cdot \frac{\partial f}{\partial y} \cdot \sigma_{xy}$$

che adesso, utilizzando la (149), diviene:

$$\sigma_G^2 = \left(\frac{\partial f}{\partial x}\right)^2 \cdot \sigma_x^2 + \left(\frac{\partial f}{\partial y}\right)^2 \cdot \sigma_y^2 + 2r_{xy} \cdot \frac{\partial f}{\partial x} \cdot \frac{\partial f}{\partial y} \cdot \sigma_x \sigma_y \tag{151}$$

Nel nostro lavoro di fisici un esempio importante di variabili correlate è costituito dall'insieme dei parametri forniti da un qualunque procedimento di fit. Infatti tali parametri sono funzioni delle stesse misure, per cui in generale sono tra di loro correlati. Ogni programma di fitting, solitamente, riporta tra i risultati anche la *matrice di correlazione* dei parametri (vedi per esempio le uscite di gnuplot dopo una procedura di fit). Gli elementi di matrice sono costituiti dai coefficienti di correlazione tra i parametri: quelli diagonali sono uguali ad 1 (ogni parametro è completamente correlato a sé stesso, il che si verifica immediatamente applicando la definizione di coefficiente di correlazione), gli elementi di matrice fuori diagonale danno il coefficiente di correlazione tra i parametri diversi, noto il quale è possibile determinarne la covarianza, che, come sappiamo, è un ingrediente necessario nella propagazione delle incertezze per funzioni di tali parametri.

► ESEMPIO 7.3. Supponiamo di avere fatto un fit lineare

$$y = a + bx$$

di un insieme di misure (x_i, y_i) e di dovere calcolare una quantità funzione di a e b

$$G = f(a, b),$$

come succede quando per esempio usiamo funzioni di taratura o quando ci interessa l'intercetta con l'asse $y = 0$; a e b sono in generale correlati (cambiando coefficiente angolare è necessario cambiare anche l'intercetta per far sì che la retta continui ad adattarsi ai dati), per cui il modo giusto di propagare le incertezze su G è dato dalla (151). Un programma di fit completo fornisce di solito anche r_{ab} , σ_a e σ_b , ed allora l'incertezza sulla grandezza G è subito calcolata. □

7.2 LA DISTRIBUZIONE DI STUDENT

Questa distribuzione entra in gioco quando abbiamo a che fare con campioni *piccoli* e vogliamo determinare intervallo di confidenza e coefficiente di confidenza, oppure vogliamo verificare alcune ipotesi statistiche (confronto di medie). In effetti l'abbiamo già incontrata nel paragrafo 4.4.

7.2.1 Variabile t di Student a n gradi di libertà

Sia data una variabile casuale x con distribuzione gaussiana di media μ e varianza σ^2 . Costruiamo la corrispondente variabile in forma standard (o scarto standardizzato):

$$z = \frac{x - \mu}{\sigma}$$

La sua funzione di distribuzione è ancora gaussiana ed ha valor medio 0 e varianza 1, come si verifica facilmente. Dati n campionamenti x_i ($i = 1 \dots n$) della variabile x (le x_i sono quantità statisticamente indipendenti provenienti dalla stessa distribuzione

normale con media μ e deviazione standard σ), consideriamo la variabile X definita come:

$$X = \sum_{i=1}^n z_i^2 = \sum_{i=1}^n \frac{(x_i - \mu)^2}{\sigma^2}$$

La variabile X è distribuita come un χ^2 a n gradi di libertà (cfr. sezione 3.6). Costruiamo adesso la nuova variabile

$$t = \frac{z}{\sqrt{\frac{X}{n}}}$$

che è in sostanza il rapporto tra una variabile normale standard e la radice quadrata di un χ^2 ridotto. Tale variabile è chiamata *variabile t di Student a n gradi di libertà* e si può dimostrare che la sua funzione di distribuzione è data da:

$$\tau(t; n) = \frac{C_n}{\left(1 + \frac{t^2}{n}\right)^{\frac{n+1}{2}}} \quad (152)$$

Tale funzione di distribuzione è simmetrica rispetto a $t = 0$ e per n grande tende ad una distribuzione normale standard (cioè con media 0 e varianza 1)². Si può anche dimostrare che la media della distribuzione t è:

$$\mu = E[t] = 0 \quad (153)$$

e la sua varianza (per $n > 2$):

$$\sigma^2 = E[t^2] = \frac{n}{n-2} \quad (154)$$

Da cui ancora:

$$\sigma = \sqrt{\frac{n}{n-2}} \quad (155)$$

7.2.2 Variabile t di Student a $n-1$ gradi di libertà

Dato un campione di n variabili x_i , provenienti da una stessa distribuzione gaussiana con media μ e varianza σ^2 , costruiamo la migliore stima della media e della varianza :

$$\begin{aligned} m &= \sum_i^n \frac{x_i}{n} \\ s^2 &= \sum_i^n \frac{(x_i - m)^2}{n-1} \end{aligned}$$

² Per dimostrarlo si sfrutta il limite notevole $\lim_{x \rightarrow \infty} \left(1 + \frac{k}{x}\right)^x = e^k$.

Si può dimostrare che la variabile

$$X' = (n-1) \frac{s^2}{\sigma^2}$$

ha la funzione di distribuzione di un χ^2 a $n-1$ gradi di libertà. Inoltre si verifica facilmente che la variabile

$$z' = \frac{m - \mu}{\frac{\sigma}{\sqrt{n}}}$$

è una variabile normale standard. Allora la variabile

$$t_{n-1} = \frac{z'}{\sqrt{\frac{X'}{n-1}}}$$

segue la distribuzione t di Student a $n-1$ gradi di libertà (per questo la indichiamo come t_{n-1}). Se esplicitiamo z' e X' si ottiene :

$$t_{n-1} = \frac{m - \mu}{\frac{\sigma}{\sqrt{n}}} \cdot \frac{\sigma}{s}$$

e infine si arriva alla interessante forma:

$$t_{n-1} = \frac{m - \mu}{s_m}$$

dove, al solito:

$$s_m = \frac{s}{\sqrt{n}}$$

Ed è proprio in questa forma che abbiamo utilizzato nel paragrafo 4.4 la variabile casuale t_{n-1} per costruire, nel caso di campioni *piccoli*, l'intervallo di confidenza per un assegnato coefficiente di confidenza.

7.2.3 t -test

La variabile t di Student viene utilizzata anche nel confronto di medie. Per esempio supponiamo di avere due campioni x ed y , non tanto grandi, inizialmente provenienti dalla stessa popolazione con media μ e deviazione standard σ . Supponiamo che uno dei due campioni subisca qualche variazione nelle condizioni ambientali o sia sottoposto a qualche trattamento particolare. Registriamo i valori medi per ciascun campione e la loro differenza. Ci possiamo domandare se tale differenza è significativa, se cioè indica che nel secondo campione è cambiato qualche cosa (se le mutate condizioni ambientali del secondo campione o il trattamento a cui è stato sottoposto lo hanno fatto evolvere verso stati diversi rispetto al primo, se hanno cambiato in un qualche modo le sue condizioni fisiche).

Per rispondere a questa domanda, facciamo l'ipotesi che non ci siano state alterazioni nei due campioni anche dopo il diverso trattamento. Si cerca quindi di costruire una

variabile statistica che permetta di trovare quale sia la probabilità che l'ipotesi fatta sia corretta. Consideriamo la variabile:

$$\Delta = m_x - m_y$$

dove:

$$m_x = \frac{1}{n_x} \sum_{i=1}^{n_x} x_i$$

$$m_y = \frac{1}{n_y} \sum_{i=1}^{n_y} y_i$$

e n_x ed n_y sono le dimensioni dei campioni. Dalla (100) segue che:

$$\sigma_{\Delta}^2 = \sigma_{m_x}^2 + \sigma_{m_y}^2$$

Pertanto, nell'ipotesi fatta che i due campioni appartengano sempre alla stessa distribuzione, si ha:

$$\sigma_{\Delta}^2 = \frac{\sigma^2}{n_x} + \frac{\sigma^2}{n_y} = \frac{n_x + n_y}{n_x n_y} \cdot \sigma^2$$

Quale stima di σ^2 si prende la quantità:

$$s^2 = \frac{\sum_{i=1}^{n_x} (x_i - m_x)^2 + \sum_{i=1}^{n_y} (y_i - m_y)^2}{n_x + n_y - 2}$$

Allora σ_{Δ}^2 è stimato da:

$$s_{\Delta}^2 = \frac{n_x + n_y}{n_x n_y} \cdot s^2$$

Costruiamo ora la variabile :

$$t = \frac{\Delta}{s_{\Delta}} \tag{156}$$

Questa variabile segue la distribuzione t di Student a $n_x + n_y - 2$ gradi di libertà. Infatti la variabile:

$$z'' = \frac{\Delta}{\sigma \sqrt{\frac{n_x + n_y}{n_x n_y}}}$$

è una variabile normale standard (provate a verificarlo). D'altra parte la variabile:

$$X'' = (n_x + n_y - 2) \frac{s^2}{\sigma^2}$$

è un χ^2 a $n_x + n_y - 2$ gradi di libertà. Allora la variabile

$$t = \frac{z''}{\sqrt{\frac{X''}{n_x + n_y - 2}}}$$

segue la distribuzione t di Student a $n_x + n_y - 2$ gradi di libertà. Se esplicitiamo z'' e X'' si ottiene, come si voleva, la (156).

A questo punto non rimane che calcolare le quantità Δ ed s_Δ dai dati registrati dei due campioni, e quindi il valore di t , che indichiamo con \tilde{t} . Allora

$$P(t < \tilde{t}) = \int_{-\infty}^{\tilde{t}} \tau(t; n_x + n_y - 2) dt$$

e il livello di significatività dell'ipotesi fatta è dato da $1 - P$.

7.3 FUNZIONE DI DISTRIBUZIONE PER FUNZIONI DI VARIABILE CASUALE

Supponiamo di avere una variabile casuale, che chiameremo x , di cui è nota la funzione di distribuzione. Consideriamo ora una funzione y della variabile x :

$$y = \phi(x)$$

y è pertanto una variabile casuale generata dalla funzione $\phi(x)$. Il problema che si pone è come trovare la funzione di distribuzione della variabile casuale y , nota la funzione di distribuzione della variabile casuale x .

7.3.1 Variabili casuali discrete

Supponiamo dapprima che x sia una variabile discreta. Pertanto la sua funzione di distribuzione $P(x)$ è essenzialmente la probabilità associata ad ogni valore che la variabile x può assumere.

Se facciamo l'ulteriore assunzione che la funzione $\phi(x)$ sia *biunivoca* nell'intervallo di variabilità di x (cioè che non ci siano due o più valori x_i della variabile x che corrispondano allo stesso valore di y), si ha semplicemente:

$$P(y = \phi(x_i)) = P(x = x_i) \quad (157)$$

► **ESEMPIO 7.4.** Consideriamo una variabile casuale discreta x che possa assumere valori interi compresi tra 0 e 3 (compresi) con la seguente funzione di distribuzione $P(x)$.

x_i	0	1	2	3
$P(x_i)$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

Supponiamo adesso di essere interessati alla variabile y definita da $y = x^2$. La funzione di distribuzione di y si scrive banalmente come:

y_i	0	1	4	9
$P(y_i)$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

□

Lasciamo adesso cadere l'ipotesi di biunivocità. È facile convincersi che in questo caso, preso un valore (ammesso) \tilde{y} della variabile y :

$$P(y = \tilde{y}) = \sum_{i=1}^k P(x = x_i) \quad (158)$$

dove la somma è estesa a tutti i valori (che qui supponiamo genericamente essere k) di x per cui:

$$\phi(x_i) = \tilde{y}$$

► **ESEMPIO 7.5.** Consideriamo una variabile casuale discreta x con la seguente funzione di distribuzione:

x_i	-2	-1	0	1	2	3
$P(x_i)$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{2}$

Se, come prima, siamo interessati alla variabile $y = x^2$, in questo caso dobbiamo utilizzare la 158, non essendo l'ipotesi di biunivocità verificata. Per fissare le idee, ad esempio:

$$P(y = 1) = P(x = -1) + P(x = 1) = \frac{1}{16} + \frac{1}{8} = \frac{3}{16}$$

La funzione di distribuzione di y si scrive esplicitamente come:

y_i	0	1	4	9
$P(y_i)$	$\frac{1}{8}$	$\frac{3}{16}$	$\frac{3}{16}$	$\frac{1}{2}$

□

7.3.2 Variabili casuali continue

Consideriamo ora il caso in cui x sia una variabile continua. Supponiamo dapprima che la funzione ϕ sia biunivoca nell'intervallo di variabilità di x . Supponiamo inoltre che $\phi(x)$ sia derivabile. La biunivocità della corrispondenza tra le due variabili assicura che se x è compresa in un intervallo infinitesimo centrato su un generico valore x_0 :

$$x \in [x_0 - dx, x_0 + dx]$$

allora (e solo allora) y è compresa in un intervallo (infinitesimo)

$$y \in [y_0 - dy, y_0 + dy]$$

centrato intorno al valore

$$y_0 = \phi(x_0)$$

e di semiampiezza³:

$$dy = \left| \frac{d\phi}{dx}(x) \right| dx$$

Essendo la corrispondenza tra le due variabili biunivoca, è possibile, in linea di principio, invertire la funzione ϕ e scrivere la variabile originaria x in funzione di quella trasformata y . Senza volere seguire una dimostrazione formale, questo fatto (si pensi anche al caso precedente) permette di affermare che la probabilità che la y sia in un intervallo infinitesimo generico di ampiezza dy è uguale alla probabilità che la x sia compresa nell'intervallo corrispondente, centrato nel valore

$$x = \phi^{-1}(y)$$

e di semiampiezza

$$dx = \left| \frac{d\phi^{-1}}{dy}(y) \right| dy$$

Pertanto, indicando con $p_x(x)$ la densità di probabilità per la variabile casuale x e con $p_y(y)$ la corrispondente densità di probabilità per la y , si può scrivere:

$$p_y(y) dy = p_x(x) dx = p_x(\phi^{-1}(y)) \left| \frac{d\phi^{-1}}{dy}(y) \right| dy$$

da cui, in definitiva:

$$\boxed{p_y(y) = p_x(\phi^{-1}(y)) \left| \frac{d\phi^{-1}}{dy}(y) \right|} \quad (159)$$

che costituisce sostanzialmente la relazione cercata.

► **ESEMPIO 7.6.** Consideriamo una variabile casuale x definita nell'intervallo:

$$x \in [0, 1]$$

e ivi distribuita uniformemente:

$$p_x(x) = 1$$

Vogliamo calcolare la funzione di distribuzione della variabile y definita da:

$$y = \phi(x) = 2x + 1$$

L'intervallo di variabilità della nuova variabile sarà banalmente:

$$y \in [1, 3]$$

³ Il modulo tiene conto del fatto che $\phi(x)$ può essere sempre crescente oppure sempre decrescente.

e per scrivere la funzione di distribuzione dovremo prima invertire ϕ :

$$x = \phi^{-1}(y) = \frac{1}{2}(y - 1)$$

e quindi calcolare:

$$\frac{d\phi^{-1}}{dy}(y) = \frac{1}{2}$$

Dalla (159) leggiamo direttamente:

$$p_y(y) = 1 \cdot \frac{1}{2} = \frac{1}{2}$$

per cui la nostra funzione di trasformazione $\phi(x)$, in questo caso, trasforma semplicemente un variabile distribuita uniformemente in $[0, 1]$ in una distribuita uniformemente in $[1, 3]$. \square

► ESEMPIO 7.7. Supponiamo di avere una variabile casuale x definita nell'intervallo:

$$x \in \left[0, \frac{\pi}{2}\right]$$

con funzione di distribuzione:

$$p_x(x) = \frac{4}{\pi} \left(1 - \frac{2x}{\pi}\right)$$

Vogliamo calcolare la funzione di distribuzione della variabile casuale y definita da:

$$y = \phi(x) = \sin x$$

Prima di tutto notiamo che la variabile y sarà definita nell'intervallo:

$$y \in [0, 1]$$

Come prima invertiamo la funzione di trasformazione $\phi(x)$:

$$x = \phi^{-1}(y) = \arcsin y$$

e calcoliamo la derivata:

$$\frac{d\phi^{-1}}{dy}(y) = \frac{1}{\sqrt{1-y^2}}$$

Tanto basta per scrivere immediatamente la funzione di distribuzione della variabile y , che sarà:

$$p_y(y) = \frac{4}{\pi} \left(1 - \frac{2}{\pi} \arcsin y\right) \frac{1}{\sqrt{1-y^2}}$$

come si legge dalla 159 \square

► ESEMPIO 7.8. Consideriamo la variabile casuale x definita nell'intervallo:

$$x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$$

e qui distribuita uniformemente:

$$p_x(x) = \frac{1}{\pi}$$

Supponiamo che la nostra legge di trasformazione sia:

$$\phi(x) = a \tan x$$

con a reale e positivo. Esattamente come prima:

$$x = \phi^{-1}(y) = \arctan\left(\frac{y}{a}\right)$$

e:

$$\frac{d\phi^{-1}}{dy}(y) = \frac{a}{a^2 + y^2}$$

Da cui:

$$p_y(y) = \frac{1}{\pi} \frac{a}{a^2 + y^2}$$

che è una distribuzione di Cauchy (cfr. esempio 3.15). □

Se abbandoniamo l'ipotesi della corrispondenza biunivoca dobbiamo, esattamente come nel caso di variabile discreta, considerare il fatto che ad uno stesso valore di y possono corrispondere più valori di x . Pertanto, analogamente al caso discreto, si avrà:

$$p_y(y) = \sum_{i=1}^k p_x(\phi_i^{-1}(y)) \left| \frac{d\phi_i^{-1}}{dy}(y) \right| \quad (160)$$

dove la somma è estesa a tutti i valori x_i di x (che supponiamo genericamente essere k) che sono immagine di y attraverso la funzione inversa di ϕ , in tutti i suoi possibili rami indipendenti⁴:

$$x_i = \phi_i^{-1}(y)$$

► ESEMPIO 7.9. Consideriamo una variabile gaussiana x in forma standard ($\mu = 0, \sigma = 1$):

$$p_x(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

e sia:

$$y = \phi(x) = x^2$$

È facile rendersi conto che la variabile y può assumere tutti i valori compresi nell'intervallo:

$$y \in [0, \infty)$$

⁴ Potremmo dire, altrettanto correttamente, che la somma è estesa a tutti i valori di x per cui:

$$y = \phi(x_i)$$

Ma il punto fondamentale è che la funzione di trasformazione ϕ non è biunivoca, nel senso che per ogni y esistono esattamente⁵ due valori, x_1 ed x_2 , di x che sono immagine di y attraverso ϕ^{-1} :

$$\begin{aligned}x_1 &= \phi_1^{-1}(y) = +\sqrt{y} \\x_2 &= \phi_2^{-1}(y) = -\sqrt{y}\end{aligned}$$

Notiamo che:

$$\left| \frac{d\phi_1^{-1}}{dy}(y) \right| = \left| \frac{d\phi_2^{-1}}{dy}(y) \right| = \frac{1}{2\sqrt{y}}$$

Dall'equazione (160) leggiamo allora:

$$p_y(y) = \frac{1}{\sqrt{2\pi}} e^{-y/2} \frac{1}{2\sqrt{y}} + \frac{1}{\sqrt{2\pi}} e^{-y/2} \frac{1}{2\sqrt{y}} = \frac{1}{\sqrt{2\pi y}} e^{-y/2}$$

che può essere confrontata direttamente con la distribuzione attesa nel caso di una variabile χ^2 ad un solo grado di libertà. \square

► **ESEMPIO 7.10.** Supponiamo di avere n variabili gaussiane x_i di media $\mu = 0$ e deviazione standard (uguale per tutte) σ . Consideriamo adesso la variabile:

$$y = \sum_{i=1}^n x_i^2$$

Le relazioni che abbiamo ricavato fino ad adesso non si possono applicare direttamente perché y non è funzione di una sola variabile casuale. Noi sappiamo però che se σ fosse uguale ad 1, y sarebbe la somma di n variabili gaussiane standard, per cui la sua funzione di distribuzione sarebbe quella di una variabile χ^2 ad n gradi di libertà. Proviamo allora a scrivere:

$$y = \sum_{i=1}^n \sigma^2 \left(\frac{x_i}{\sigma} \right)^2 = \sigma^2 \sum_{i=1}^n \left(\frac{x_i}{\sigma} \right)^2 = \sigma^2 x$$

dove abbiamo definito la nuova variabile x come:

$$x = \sum_{i=1}^n z_i^2 = \sum_{i=1}^n \left(\frac{x_i}{\sigma} \right)^2$$

È facile convincersi che x è distribuita proprio come una variabile χ^2 ad n gradi di libertà, poiché, se le x_i sono variabili gaussiane di media $\mu = 0$ e deviazione standard σ , le z_i date da:

$$z_i = \frac{x_i}{\sigma}$$

⁵ A rigore l'affermazione non è esatta per il punto $y = 0$, ma questo, come vedremo tra un attimo, non avrà nessuna influenza su quanto diremo.

sono variabili gaussiane in forma standard. Cerchiamo allora di riassumere il tutto nel linguaggio che abbiamo sviluppato nel resto del paragrafo. Abbiamo una variabile casuale x la cui funzione di distribuzione è data da (cfr. equazione (82)):

$$p_x(x) = \wp_n(x) = \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} x^{\frac{n-2}{2}} e^{-x/2}$$

ed una variabile y , funzione della x , data da:

$$y = \phi(x) = \sigma^2 x$$

Da cui:

$$x = \phi^{-1}(y) = \frac{y}{\sigma^2}$$

e

$$\frac{d\phi_1^{-1}}{dy}(y) = \frac{1}{\sigma^2}$$

Combinando tutto insieme si ha, banalmente:

$$p_y(y) = \wp_n\left(\frac{y}{\sigma^2}\right) \cdot \frac{1}{\sigma^2} = \frac{1}{2^{\frac{n}{2}} \Gamma\left(\frac{n}{2}\right)} \left(\frac{y}{\sigma^2}\right)^{\frac{n-2}{2}} e^{-y/2\sigma^2} \cdot \frac{1}{\sigma^2}$$

□

► ESEMPIO 7.11. Supponiamo di avere una variabile casuale x definita da:

$$x = \sum_{i=1}^n x_i^2$$

dove le x_i sono n variabili gaussiane di media $\mu = 0$ e deviazione standard (uguale per tutte) σ . Abbiamo trovato la funzione di distribuzione di x nell'esercizio precedente (in cui, tanto per essere chiari, la variabile in questione si chiamava y). Consideriamo adesso la variabile:

$$y = \phi(x) = \sqrt{x}$$

Abbiamo, al solito:

$$x = \phi^{-1}(y) = y^2$$

e

$$\frac{d\phi_1^{-1}}{dy}(y) = 2y$$

Da cui si ricava immediatamente:

$$p_y(y) = \wp_n\left(\frac{y}{\sigma^2}\right) \cdot \frac{2y}{\sigma^2}$$

□

► **ESEMPIO 7.12.** Si tratta di un'applicazione tutto sommato elementare dell'esercizio precedente. Consideriamo una data molecola di un gas omogeneo che si trova all'equilibrio termico ad una certa temperatura T . In generale i valori delle tre componenti della velocità (v_x , v_y e v_z) della molecola in questione, ad un istante di tempo fissato, sono variabili casuali distribuite gaussianamente con media $\mu = 0$ ⁶ e deviazione standard σ_v ⁷, uguale nelle tre direzioni. Ci chiediamo quale sia la funzione di distribuzione del modulo della velocità v , definito da:

$$v = \sqrt{v_x^2 + v_y^2 + v_z^2}$$

È chiaro che, a parte lievi differenze di notazione, siamo essenzialmente nel caso dell'esercizio precedente con $n = 3$ (adesso v_x , v_y e v_z rappresentano le x_i e v rappresenta la y). La funzione di distribuzione per una variabile χ^2 con tre gradi di libertà si scrive esplicitamente come:

$$\wp_3(x) = \frac{x^{\frac{1}{2}} e^{-\frac{x}{2}}}{2^{\frac{3}{2}} \Gamma\left(\frac{3}{2}\right)}$$

Possiamo esplicitare il valore della funzione Γ di Eulero seguendo la definizione (cfr. sezione 3.6) e sfruttando la relazione di ricorrenza:

$$\Gamma\left(\frac{3}{2}\right) = \Gamma\left(\frac{1}{2} + 1\right) = \frac{1}{2} \Gamma\left(\frac{1}{2}\right)$$

dove:

$$\Gamma\left(\frac{1}{2}\right) = \int_0^\infty t^{-\frac{1}{2}} e^{-t} dt$$

Tramite la sostituzione $s = \sqrt{t}$ otteniamo:

$$\Gamma\left(\frac{1}{2}\right) = \int_0^\infty e^{-s^2} ds = \sqrt{\pi}$$

da cui:

$$\Gamma\left(\frac{3}{2}\right) = \frac{\sqrt{\pi}}{2}$$

ed infine:

$$\wp_3(x) = \frac{x^{\frac{1}{2}} e^{-\frac{x}{2}}}{\sqrt{2\pi}}$$

La funzione di distribuzione di v si scrive come:

$$p_v(v) = \frac{1}{\sigma_v^3} \sqrt{\frac{2}{\pi}} v^2 e^{-\frac{v^2}{2\sigma_v^2}}$$

che è generalmente nota come distribuzione di Maxwell delle velocità. □

⁶ Il fatto che il valor medio di ognuna delle componenti della velocità sia nullo è connesso con il fatto che, in assenza di una direzione spaziale privilegiata, ad ogni istante la molecola ha la stessa probabilità di muoversi verso destra o verso sinistra.

⁷ Il valore di σ_v resta sostanzialmente fissato, come si può leggere in un qualsiasi testo di termodinamica o meccanica statistica, dalla temperatura T e dalla costante di Boltzmann k .

7.4 ESERCIZI

- ▶ ESERCIZIO 7.1. Dimostrare la proprietà 4 enunciata nel paragrafo 7.1.1.
- ▶ ESERCIZIO 7.2. Dimostrare la proprietà 5 enunciata nel paragrafo 7.1.1.
- ▶ ESERCIZIO 7.3. Calcolare esplicitamente il coefficiente di correlazione tra le due serie di dati dell'esempio 7.1 e confrontarlo con il risultato fornito nell'esempio 7.2 (è probabilmente consigliabile utilizzare una calcolatrice programmabile oppure un foglio elettronico).
- ▶ ESERCIZIO 7.4. Verificare che la distribuzione $p_x(x)$ dell'esempio 7.7 è correttamente normalizzata.
- ▶ ESERCIZIO 7.5. Verificare che la distribuzione $p_y(y)$ dell'esempio 7.7 è correttamente normalizzata.
- ▶ ESERCIZIO 7.6. Dimostrare che la distribuzione $p_y(y)$ dell'esempio 7.9 è esattamente la funzione di distribuzione di una variabile χ^2 (cfr. equazione (82)) nel caso di un grado di libertà.

Parte II

INTRODUZIONE AL CALCOLATORE

INTRODUZIONE

Non vi è dubbio che *condensare* in poche decine di pagine le informazioni necessarie ad un uso (anche solo vagamente) proficuo del calcolatore per l'analisi di dati scientifici sia una pretesa non da poco. Da una parte, infatti, la quantità di informazioni disponibile è virtualmente sconfinata, il che rende la selezione del materiale, di per sé, difficoltosa. Anche una volta scelto il materiale, si pone poi il problema di presentarlo in maniera comprensibile, organica e, per quanto possibile, originale.

L'idea di base attorno alla quale questa seconda parte è organizzata è sostanzialmente la definizione delle nozioni e delle competenze che lo studente deve acquisire perché il primo incontro con il calcolatore si riveli fruttuoso. Nozioni che si riassumono essenzialmente nella capacità utilizzare il *computer* per realizzare grafici, eseguire *fit* numerici e presentare degnamente i risultati dell'analisi. I pacchetti *software* che abbiamo selezionato a questo scopo sono gnuplot e L^AT_EX, supportati nella nostra scelta dalla loro ampia diffusione all'interno della comunità scientifica e dalla semplicità di utilizzo (oltre alla nostra predilezione per il *free software*).

Abbiamo anche ritenuto opportuno preporre alla descrizione di questi pacchetti alcune nozioni di base sull'uso del calcolatore; una sorta di manuale di sopravvivenza per utilizzatori (molto) inesperti del sistema operativo Linux.

I riferimenti bibliografici dovrebbero essere più che sufficienti per colmare le lacune delle pagine che seguono e soddisfare la curiosità dei più esigenti.

IL SISTEMA OPERATIVO LINUX

INDICE

8.1	Terminologia	164
8.2	Una breve storia di Linux	164
8.3	Concetti di base	165
8.3.1	Il <i>login</i>	166
8.3.2	La <i>shell</i> ed i comandi	166
8.3.3	Il <i>logout</i>	167
8.4	Il <i>filesystem</i> Linux	167
8.5	Navigare il <i>filesystem</i>	168
8.5.1	<code>pwd</code>	168
8.5.2	<code>ls</code>	168
8.5.3	<code>cd</code>	170
8.6	Modificare il <i>filesystem</i>	170
8.6.1	<code>mkdir</code>	171
8.6.2	<code>rmdir</code>	171
8.6.3	<code>rm</code>	172
8.6.4	<code>cp</code>	172
8.6.5	<code>mv</code>	172
8.6.6	<code>chmod</code>	173
8.7	Visualizzare i <i>file</i> di testo	173
8.7.1	<code>more</code>	173
8.7.2	<code>less</code>	173
8.7.3	<code>head</code>	174
8.7.4	<code>tail</code>	174
8.8	Modificare i <i>file</i> di testo	174
8.9	I processi	174
8.9.1	<code>ps</code>	175
8.9.2	<code>top</code>	175
8.9.3	<code>kill</code>	175
8.10	La stampa	176
8.10.1	<code>lpr</code>	176
8.10.2	<code>lpq</code>	176
8.10.3	<code>lprm</code>	176
8.11	Accesso alle periferiche	176
8.11.1	<code>mount</code>	177
8.11.2	<code>umount</code>	178

8.1 TERMINOLOGIA

Esattamente come per la prima parte, questa breve introduzione all'uso del calcolatore prende avvio con un elenco dei termini *tecnici* che saranno diffusamente utilizzati nel seguito. Ovviamente la lista non ha alcuna pretesa di essere esaustiva; il lettore più esigente perdonerà anche alcuni passaggi non particolarmente rigorosi.

HARDWARE. Termine generico per indicare i dispositivi (per la maggior parte elettronici) che costituiscono fisicamente il calcolatore.

HARD DISK (*disco rigido*). Dispositivo magnetico per l'archiviazione *permanente* dei dati.

SISTEMA OPERATIVO. Insieme di programmi che costituiscono l'interfaccia tra l'utente e l'*hardware*. Gestisce il corretto funzionamento dell'*hardware* stesso e l'esecuzione delle applicazioni.

FILE. È sostanzialmente un'unità di memoria logica persistente (nel senso che risiede fisicamente su un disco rigido o su una memoria di massa in genere), gestita dal sistema operativo, contenente un insieme di informazioni.

DIRECTORY. È una sorta di *contenitore* che può avere al suo interno *file* o altre *directory*. Nel seguito, con un lieve abuso di terminologia, useremo anche la parola *cartella* come sinonimo.

FILESYSTEM. Parte del sistema operativo che si occupa della gestione dei *file*.

SHELL. Interprete interattivo dei comandi, i quali vengono inviati tramite terminale al sistema operativo. La *shell* è un programma di sistema che agisce da interfaccia tra utente e sistema operativo stesso.

SOFTWARE. Termine generico per indicare i programmi per il calcolatore. Ogni programma è costituito da un insieme di istruzioni che possono essere eseguite dal calcolatore stesso.

FREEWARE. Letteralmente si traduce come *software* gratuito. Indica genericamente tutto quel *software* per il cui utilizzo non è necessario pagare una licenza.

OPEN SOURCE. Termine che indica i programmi il cui codice sorgente è pubblico e disponibile per la consultazione e la modifica. Notiamo esplicitamente che non tutti i programmi *freeware* sono anche *open source*.

8.2 UNA BREVE STORIA DI LINUX

Nel 1965 i laboratori della Bell Telephone (una divisione della AT&T) decisero di interrompere una collaborazione avviata con la General Electric ed il Project MAC del MIT per la scrittura di un sistema operativo che avrebbe dovuto chiamarsi *Multics*. Pur

tuttavia due programmatori, Ken Thompson e Dennis Ritchie, ritennero che fosse un peccato sprecare il lavoro fatto; i loro sforzi culminarono nella realizzazione di una prima versione del sistema che un altro ricercatore, Brian Kernighan (futuro *inventore* del linguaggio di programmazione C), decise di chiamare, in opposizione a Multics, Unix.

Nel 1973 Unix, originariamente scritto in *assembly language* venne interamente tradotto in C; questa riscrittura aumentò notevolmente la portabilità del codice e segnò l'inizio della sua diffusione. Alla fine degli anni '70 la AT&T fornì a basso costo le licenze del sistema operativo a diversi *college* ed università americane ed in pochi anni la popolarità di Unix riuscì a superare i confini ristretti delle istituzioni accademiche; ovviamente col tempo le cose sono andate avanti ed oggi esistono svariate versioni commerciali del sistema, per le diverse piattaforme *hardware*.

Linux è un sistema operativo Unix a tutti gli effetti; è una sorta di *clone*, riscritto da zero, di Unix. La cosa interessante è che nessuna istituzione commerciale è stata coinvolta nella sua redazione (ed in effetti è assolutamente *freeware*); esso è il risultato dello sforzo congiunto di migliaia di persone che, in tutto il mondo, hanno dedicato (e continuano a dedicare) tempo e fatica per perfezionare ciò che era iniziato come progetto di esplorazione del chip 80386. Può essere curioso ricordare che uno dei primi *prodotti* di Linus Torvalds (l'ideatore e padre di Linux) fu un programma che, tramite due processi funzionanti in *multitasking*, doveva stampare alternativamente AAAA e BBBB. Ne è stata fatta di strada. . .

8.3 CONCETTI DI BASE

Linux è un sistema operativo *multi-utente* e *multi-tasking*. Preciseremo, in parte, il senso di queste affermazione strada facendo; pur tuttavia spendere due parole al proposito è probabilmente un buon punto di partenza.

Multi-utenza significa che diversi utenti possono avere diversi dati, programmi ed impostazioni di sistema memorizzate separatamente sulla stessa macchina; la struttura del *filesystem* Linux è basata, come vedremo, sui concetti di proprietà e permesso.

Multi-tasking significa semplicemente che più processi possono essere eseguiti contemporaneamente. Anche di questo ci occuperemo più in dettaglio tra breve.

Come tutti i sistemi Unix, Linux è formato da un *kernel* e da programmi di sistema, oltre ovviamente alle *applicazioni* propriamente dette, con cui l'utente fa il lavoro vero e proprio. Il *kernel* costituisce il cuore del sistema operativo; avvia i processi e permette loro di operare contemporaneamente, assegna le risorse di sistema ai vari processi ed in definitiva fornisce tutti gli strumenti necessari ad implementare i servizi che normalmente si richiedono ad un sistema operativo; questa implementazione avviene attraverso chiamate di sistema. I programmi di sistema, ed a maggior ragione le applicazioni, girano *sopra* al *kernel*, in quella che si chiama modalità utente.

8.3.1 Il login

Essendo Linux, come abbiamo appena detto, un sistema operativo multi-utente, l'accesso al sistema è subordinato ad una procedura di identificazione che prende il nome di *login*. A tale scopo l'utente è tenuto a fornire il proprio *username* e la propria *password* (che presumibilmente sono stati forniti dall'amministratore di sistema) per potersi *loggar*e (cioè eseguire il *login*) alla macchina. Esistono diverse schermate di *login*, che possono essere grafiche o testuali a seconda delle impostazioni di sistema e della distribuzione di Linux utilizzata.



Linux (al contrario di Microsoft DOS e di Microsoft Windows, come probabilmente qualcuno sa) distingue maiuscole e minuscole (cioè è case sensitive) per cui il nome utente, la password ed i comandi in generale devono essere digitati esattamente come sono, con lettere maiuscole e minuscole al loro posto.

8.3.2 La shell ed i comandi

Una volta loggato sulla macchina, l'utente può aprire una *shell* (se il sistema non è configurato per farlo automaticamente). La *shell* è il luogo in cui si impartiscono i comandi Linux; è per certi aspetti l'equivalente del *prompt* di Microsoft DOS, per chi ha un po' di esperienza al proposito. Esistono svariati tipi di *shell*, che sostanzialmente differiscono tra di loro per quanto riguarda il linguaggio di *scripting* ed alcune altre funzionalità che il neofita tipicamente non utilizza.

L'uso della *shell* è semplicissimo e consiste essenzialmente nello scrivere un comando Linux valido e nel premere il tasto di <INVIO>. Ogni comando può avere delle opzioni (che sono precedute da un carattere "-" e ne determinano l'esito) e può accettare dei parametri; in sostanza la struttura tipica di un comando Linux è del tipo¹:

```
>comando -opzione argomento
```

Sarà tutto più chiaro tra poco, quando vedremo alcuni esempi espliciti. Vale la pena di notare che Linux ha un efficiente sistema di documentazione che può rivelarsi utilissimo quando non si ricorda con precisione la sintassi di un comando o l'effetto di un'opzione sul comando stesso. È sufficiente digitare dalla *shell*:

```
>man comando
```

per avere sul display tutta la documentazione disponibile (si torna alla *shell* semplicemente digitando "q").

Prima di andare avanti è utile ricordare che ci sono alcuni modi per riprendere il controllo della *shell* nel caso questo dovesse momentaneamente sfuggire (il che accade tipicamente quando si esegue un comando che si aspetta l'inserimento di dati da parte dell'utente oppure quando si avvia un programma la cui esecuzione impiega più tempo del previsto):

¹ Notiamo per chiarezza che, qui e nel seguito del capitolo, i comandi da digitare sono costituiti da tutto (e solo) ciò che segue il carattere ">", la cui unica funzione è quella di segnalare che ci troviamo nella *shell* Linux.

- (1) <CTRL + c>: solitamente elimina ogni programma in esecuzione e riporta alla linea di comando.
- (2) <CTRL + z>: sospende il processo corrente mentre è in esecuzione e riporta alla linea di comando.

8.3.3 Il logout

Il processo di *logout* chiude tutti i *file* che sono (per qualche ragione) aperti e termina ogni programma mandato in esecuzione dall'utente. È buona regola scollegarsi sempre, una volta che abbiamo finito di utilizzare il computer; questo permette agli altri utenti, eventualmente collegati da un terminale remoto, di utilizzare nel modo più efficiente possibile le risorse di sistema. Per fare il *logout* da una sessione testuale è sufficiente digitare dalla *shell*:

```
>logout
```

ed il sistema torna al *prompt* di *login*; un nuovo utente può collegarsi. Se la sessione è grafica si può avviare la procedura di *logout* (ed eventualmente quella di spegnimento della macchina) accedendo all'apposito menu con il mouse.

8.4 IL *filesystem* LINUX

A differenza di ciò che accade sotto Microsoft Windows (in cui le varie unità, interne o esterne, di memorizzazione permanente dei dati sono associate alle lettere A, B, C...), il *filesystem* di Linux ha un'unica *directory* radice (che si chiama *root* e che si indica con il simbolo "/"). La *directory root* può contenere *file* ed altre *directory* e le *directory* dei livelli sottostanti possono a loro volta fare altrettanto. La struttura fisica dell'*hardware* è in un certo senso *nascosta* all'utente: tutte le unità di archiviazione appaiono semplicemente come sotto-*directory* della *directory root*.

Ogni utente ha inoltre una *home directory*, che non ha niente di particolare a parte il fatto di essere quella in cui l'utente stesso si trova immediatamente dopo il *login*.

Sotto Linux i *file* sono caratterizzati da tre attributi fondamentali: "r", "w" ed "x" (che stanno per *read*, *write* ed *execute*); come dire che un *file* si può leggere, scrivere ed eseguire. O meglio che un determinato utente può avere (o non avere) i privilegi necessari per leggere, scrivere od eseguire un determinato *file*. Qualsiasi combinazione di questi tre attributi è ammissibile per il sistema operativo, anche se non tutte sono propriamente sensate.

Ogni *file* Linux ha un proprietario (che è un utente con un account valido sulla macchina) ed ogni utente appartiene ad uno o più gruppi di utenti. I permessi su ogni *file* sono settati su tre livelli: quello del proprietario del *file*, quello del gruppo a cui appartiene il proprietario e quello di tutti gli altri utenti del sistema; tanto per fare un esempio, è possibile che un *file* sia leggibile, scrivibile ed eseguibile dal proprietario, solamente leggibile dal gruppo del proprietario ed assolutamente inaccessibile a tutti gli altri. Vedremo tra poco come è possibile visualizzare (ed eventualmente cambiare) i permessi relativi ad un determinato *file* o ad una *directory*.

8.5 NAVIGARE IL *filesystem*

Esiste un certo numero di comandi dedicati alla navigazione nel *filesystem* e ne vedremo tra breve alcuni esempi; dovrebbero essere più o meno sufficienti per sopravvivere nel sistema, ma è sempre importante ricordare che ogni comando può avere un gran numero di opzioni che in generale è utile conoscere per soddisfare in pieno le proprie esigenze (ed il comando *man* è, in questo senso, insostituibile).

8.5.1 *pwd*

È l'acronimo di *present work directory* e restituisce sullo schermo il nome della *directory* corrente. È sufficiente digitare *pwd* per sapere dove siamo; ad esempio:

```
>pwd
/afs/pi.infn.it/user/lbaldini
>
```

8.5.2 *ls*

Elenca semplicemente i *file* contenuti in una certa *directory* (più o meno l'equivalente di *dir* in Microsoft DOS); se l'utente non passa alcun argomento al comando, elenca i *file* nella *directory* corrente. Ad esempio:

```
>ls
Ciao.doc
Temp
ClusterAxes.eps
>
```

Nel caso specifico la *directory* contiene tre oggetti. A questo livello non siamo in grado di dire se essi sono *file* o *directory*. Per far questo è necessario, come vedremo tra un attimo, invocare il comando con l'opzione "*-l*":

```
>ls -l
total 5
-rw-rw-r-- 1 lbaldini glast 36352 Dec 20 18:18 Ciao.doc
drwxrwxr-x 2 lbaldini glast 2048 Jul 7 2002 Temp
-rw-rw-r-- 1 lbaldini glast 21847 Jul 24 2002 Cluster.ps
>
```

La prima riga dice quanti sono gli oggetti nella *directory* (in questo caso cinque... capiremo tra un attimo il motivo per cui ne sono elencati solamente tre), mentre le linee successive visualizzano alcuni dettagli sugli oggetti in questione. Il primo campo indica se si tratta di un *file* ("*-*"), di una *directory* ("*d*") o di un link ("*l*"); di seguito compaiono i permessi per il proprietario, per il gruppo del proprietario e per l'utente generico ("*r*", "*w*" ed "*x*" indicano rispettivamente che l'utente possiede il permesso

di lettura, scrittura ed esecuzione, mentre il simbolo “-” indica che il corrispondente permesso è disabilitato). Nel caso in questione “Ciao.doc”, come indicato dalla stringa “-rw-rw-r-”, è un *file* che il proprietario ed il suo gruppo possono leggere e modificare, mentre tutti gli altri utenti possono solamente leggere. Invece “Temp” è una *directory* e così via. . .

Il campo successivo nella linea indica il numero di *hard link* all’oggetto in questione, ma per noi è poco importante. Seguono dunque il nome del proprietario (in questo caso “lbaldini”) e del gruppo (“glast”). Infine vediamo la dimensione dell’oggetto su disco, la data della sua ultima modifica ed il nome; da notare che, nel caso delle *directory*, la dimensione non corrisponde allo spazio su disco occupato dal contenuto, ma a quello occupato dall’unità logica che controlla la *directory* stessa.

In generale Linux non mostra all’utente (a meno che esso non lo richieda esplicitamente) tutti i *file* contenuti in una *directory*. In particolare il semplice comando `ls` nasconde i *file* e le *directory* che cominciano con il carattere “.” (generalmente si tratta di *file* di configurazione). Per visualizzare tutti i *file* si può utilizzare l’opzione “-a” di `ls`:

```
>ls -a
.
..
Ciao.doc
Temp
ClusterAxes.eps
>
```

Adesso compaiono anche gli oggetti “.” (che in Linux sta per la *directory* corrente) e “..” che viceversa indica la *directory* che nel *filesystem* si trova al livello immediatamente superiore rispetto a quella corrente (a volte detta anche *directory* genitrice).

Notiamo che l’oggetto “.” esiste in tutte le *directory* e che “..” esiste in tutte le *directory* fatta eccezione per la *directory* root (“/”), che si trova in assoluto al livello più alto nel *filesystem*.

Il comando `ls` può accettare un argomento, che indica la *directory* della quale si vogliono elencare gli oggetti:

```
>ls -a Temp
.
..
Luca.jpg
>
```

In questo caso ci dice che la *directory* “Temp” contiene (oltre ovviamente a “.” e “..”) un solo *file* che si chiama “Luca.jpg”.



Sotto Linux il tasto <TAB> funziona da auto completion, cioè completa automaticamente i nomi di comandi od oggetti del *filesystem*, a patto che l’utente ne abbia scritto una parte abbastanza lunga da rimuovere eventuali ambiguità. Una volta provato non si può più rinunciarevi!

8.5.3 cd

Esattamente come in Microsoft DOS, questo comando serve per cambiare *directory*; necessita, come argomento, del nome della *directory* di destinazione. Il percorso verso quest'ultima può essere quello assoluto (cioè partendo da *root*) oppure quello relativo (cioè partendo dalla *directory* corrente). L'esempio seguente, che unisce i pochi comandi visti sino ad adesso, dovrebbe chiarire le idee:

```
>pwd
/afs/pi.infn.it/user/lbaldini
>ls
Ciao.doc
Temp
ClusterAxes.eps
>cd Temp
>ls
Luca.jpg
>cd .
>ls
Luca.jpg
>cd ..
>ls
Ciao.doc
Temp
ClusterAxes.eps
>cd /afs/pi.infn.it/user/lbaldini/Temp
>ls
Luca.jpg
>
```

Notiamo che il comando "cd ." non fa niente (essendo "." la *directory* corrente), mentre "cd .." porta nella *directory* immediatamente superiore nel *filesystem*. La documentazione di Linux contiene un'esposizione esauriente di tutte le possibili opzioni del comando; per visualizzarla, al solito:

```
>man cd
```

(dopo di che barra spaziatrice o frecce per scorrere e "q" per tornare alla *shell*).

8.6 MODIFICARE IL *filesystem*

Ovviamente navigare non è l'unica cosa che si può fare col *filesystem*. . . Esistono comandi per creare, copiare, muovere *file* e così via. Al solito diamo una rapida occhiata a quelli indispensabili.

8.6.1 `mkdir`

Come si può intuire dal nome, crea una *directory* e richiede come argomento il nome della *directory* da creare. Al solito il percorso può essere assoluto o relativo; ovviamente se non si specifica niente, il nuovo oggetto viene creato nella *directory* corrente. Ad esempio:

```
>pwd
/afs/pi.infn.it/user/lbaldini
>ls
Ciao.doc
Temp
ClusterAxes.eps
>mkdir NewDirectory
>ls
Ciao.doc
Temp
ClusterAxes.eps
NewDirectory
>
```



Sotto Linux è buona norma evitare gli spazi all'interno dei nomi di file o directory. Lo spazio è un carattere molto particolare poiché costituisce il separatore tra comandi, opzioni ed argomenti e, quando si lavora dalla shell, gestire spazi nei nomi dei file può essere difficoltoso, se non si è preparati.

8.6.2 `rmdir`

È il contrario del comando precedente e semplicemente rimuove una *directory* esistente, a patto che sia vuota:


```
>ls
Ciao.doc
Temp
ClusterAxes.eps
NewDirectory
>rmdir NewDirectory
>ls
Ciao.doc
Temp
ClusterAxes.eps
>
```

Come vedremo tra un attimo, per cancellare una *directory* non vuota è necessario usare il comando `rm` con l'opzione `"-r"`.

8.6.3 rm

Rimuove il *file* che gli viene passato come argomento e la sintassi è analoga a quella di *rmdir*. Ha alcune opzioni che vale la pena di conoscere:

- (1) “-i”: il comando agisce in modo interattivo e chiede conferma prima di cancellare ogni *file*; caldamente consigliata quando si lavora a notte fonda!
- (2) “-f”: il comando agisce in modo forzato senza chiedere conferma.
- (3) “-r”: il comando agisce in modo ricorsivo; con questa può accettare il nome di una *directory* come argomento ed in tal caso cancella ricorsivamente tutto il contenuto della *directory* stessa.


 Può essere buona abitudine utilizzare d'abitudine il comando *rm* in modo interattivo (con l'opzione “-i”) e, cosa forse più utile, evitare di lavorare fino a notte fonda dopo una giornata faticosa ed una cena pesante!

8.6.4 cp

È l'analogo del comando *copy* in Microsoft DOS e serve per copiare un *file* di partenza (che viene passato come primo argomento) in un *file* o in una *directory* di destinazione (che costituisce viceversa il secondo argomento). Ad esempio il comando:

```
>cp *.doc Temp
>
```

copia tutti i *file* il cui nome termina per “.doc” nella *directory* “Temp”.

 Sotto Linux il simbolo “*” è un carattere speciale che sostanzialmente sta per “qualsiasi sequenza di lettere e numeri”. È la shell stessa che si occupa di operare la sostituzione, sulla base del contenuto del filesystem nel momento in cui il comando viene impartito.

L'esempio successivo:

```
>cp Ciao.doc Salve.doc
>
```

crea una copia del *file* “Ciao.doc” nella *directory* corrente chiamandola “Salve.doc”.

8.6.5 mv

Serve per rinominare un *file*:

```
>mv Ciao.doc Salve.doc
>
```

oppure per spostarlo in una *directory* di destinazione, se quest'ultima viene specificata:

```
>mv Ciao.doc Temp
>
```

Per inciso notiamo che, dopo il primo comando, il *file* “Ciao.doc” non esiste più.

8.6.6 chmod

Serve essenzialmente per cambiare i permessi su un *file*. L’utente deve passare al comando il livello del permesso che vuole cambiare (“u” per il proprietario, “g” per il gruppo e “o” per l’utente generico), l’operazione da eseguire (“+” per aggiungere il permesso e “-” per toglierlo), il tipo di permesso (“r” per la lettura, “w” per la scrittura e “x” per l’esecuzione) ed infine il nome del *file* su cui impostare i nuovi permessi.

È più semplice di quello che sembra; il seguente comando:

```
>chmod u-w Cluster.eps
>
```

toglie al proprietario i privilegi necessari per poter modificare il *file* “Cluster.eps”. Digitando “ls -l” è possibile verificarlo facilmente.

8.7 VISUALIZZARE I *file* DI TESTO

Al di là dei *file* di testo propriamente detti, molte applicazioni Linux (ed il sistema stesso) utilizzano *file* di configurazione di tipo testuale, per cui è essenziale avere strumenti per visualizzare (ed eventualmente modificare) gli oggetti di questo tipo.

Linux offre un ampio insieme di comandi dedicati alla visualizzazione dei *file* di tipo testuale. Di seguito elenchiamo i principali; al solito le pagine man sono uno strumento utile per esplorare le svariate opzioni che ognuno di questi comandi prevede.

8.7.1 more

Accetta come argomento un *file* e ne visualizza il contenuto sullo schermo; per esempio

```
>more Cluster.eps
>
```

visualizza il contenuto del *file* “Cluster.eps”. La barra spaziatrice fa avanzare di una pagina, mentre il tasto “b” riporta indietro di una pagina; si torna alla *shell* premendo il tasto “q”. Le pagine man dovrebbero illustrare l’elenco completo dei comandi che si possono usare all’interno di more.

8.7.2 less

La sintassi ed il funzionamento sono essenzialmente identici a quelli di more:

```
>less Cluster.eps
>
```

La differenza è che dovrebbe essere più facile muoversi all'interno del documento (in particolare si possono usare le frecce direzionali).

8.7.3 head

Visualizza le prime 10 righe (ma il numero può essere modificato utilizzando l'opportuna opzione) del *file* che gli viene passato come argomento. La sintassi è la solita:

```
>head Cluster.eps
>
```

È utile quando si vuole dare una sola occhiata ad un *file*.

8.7.4 tail

Al contrario del precedente visualizza le ultime righe di un *file*.

8.8 MODIFICARE I *file* DI TESTO

La visualizzazione non è certo l'unica cosa di cui uno ha bisogno, lavorando con i *file* di tipo testuale; emacs è probabilmente l'editor di testo più diffuso tra gli utenti Linux, ormai. Descrivere, sia pur in modo sommario, le sterminate funzionalità di emacs ci porterebbe lontano dal filo della discussione; d'altra parte il programma ha un ampio sistema di documentazione integrata che dovrebbe essere più che sufficiente per cavarsela in ogni situazione.

Si può aprire un *file* di testo con emacs semplicemente digitando dalla *shell*:

```
>emacs sample.txt
```

(ovviamente sostituite "sample.txt" con il nome del file di cui avete bisogno). Fatto questo, emacs funziona più o meno come un qualsiasi *editor* di testo; da notare che offre un grandissimo numero di *shortcut* da tastiera (ad esempio <CTRL + x> <CTRL + s> per salvare) che rendono veramente veloce ed efficiente (per lo meno per gli esperti) l'*editing*. Al solito: lavorare con emacs è il miglior modo per imparare ad usarlo.

8.9 I PROCESSI

Ogni comando Linux crea un *processo* che il sistema operativo esegue fino a quando esso non termina o viene interrotto. Il punto fondamentale è che, essendo multitasking, Linux permette di mandare in esecuzione più di un processo contemporaneamente; permette inoltre di mandare in esecuzione processi in *background* (o, come si dice, dietro le quinte) semplicemente facendo seguire un carattere "&" al comando.

Un esempio può essere utile per chiarire le idee. Supponiamo di voler aprire con emacs il *file* "Cluster.txt"; possiamo farlo normalmente:


```
>emacs Cluster.txt
```

oppure possiamo farlo in *background*:

```
>emacs Cluster.txt &
[2] 12563
>
```

La differenza, come si vede, è che nel secondo caso il sistema torna alla linea di comando e la *shell* è di nuovo utilizzabile, mentre nel primo la *shell* è inutilizzabile fino a che non chiudiamo emacs (anche se, a dirla tutta, possiamo sempre aprire una nuova *shell*...). Per completezza, i numeri “[2] 12563” identificano il processo all’interno del sistema operativo.

Linux offre un *set* di comandi per monitorare i processi ed eventualmente intervenire su di essi. Al solito vedremo i fondamentali.

8.9.1 ps

Visualizza l’elenco dei processi in memoria dei quali l’utente è proprietario. Non servono argomenti e la sintassi è banale:

```
>ps
>
```

Il risultato è una lista dei processi con il proprio numero di identificazione, lo stato (*running*, *stopped*, etc.) ed altre informazioni per le quali si rimanda alle pagine *man*.

8.9.2 top

Visualizza un elenco, aggiornato continuamente, di tutti i processi in memoria, fornendo inoltre alcune informazioni importanti come il numero di identificazione, il proprietario, il tempo di esecuzione, la priorità, etc. Anche qui non servono argomenti

```
>top
>
```

e le pagine *man* riportano tutte le opzioni valide. Si torna alla *shell* digitando <CTRL + c>.

8.9.3 kill

Capita a volte, per qualche ragione, di voler interrompere forzatamente un processo; in tal caso si utilizza il comando *kill*, che accetta come parametro il numero di identificazione del processo da eliminare. Ad esempio per uccidere il processo avviato quando prima avevamo lanciato emacs in background dovremmo fare:

```
>kill -9 12563
[2]+  Killed      emacs
>
```

Si rimanda alle pagine *man* per il significato dell’opzione “-9”.

8.10 LA STAMPA

L'installazione, la configurazione e la gestione delle stampanti sono punti non proprio banali di cui generalmente si occupa l'amministratore di sistema. All'utente è sufficiente conoscere pochi comandi, che fortunatamente sono gli stessi in tutti i casi (stampante locale o stampante condivisa in rete).

8.10.1 lpr

Permette di stampare un *file* e richiede come argomento il nome del *file* da stampare. Se non si passano opzioni, il documento viene inviato alla stampante di *default*; è tuttavia buona regola specificare esplicitamente il nome della stampante di destinazione mediante l'opzione "-P". Supponiamo ad esempio di avere una stampante postscript che si chiama "hp" e di voler stampare il *file* "Cluster.eps". La sintassi è:

```
>lpr -P hp Cluster.eps
>
```

8.10.2 lpq

Accetta come parametro il nome di una stampante e mostra i processi nella coda di stampa per la stampante in questione. Esempio:

```
>lpq -P hp
```

8.10.3 lprm

Permette di rimuovere un processo dalla coda di stampa, nel caso ve ne sia bisogno; supponiamo di voler eliminare il processo di stampa "123" (il numero di identificazione è fornito dal comando lpq):

```
>lprm -P hp 123
>
```

8.11 ACCESSO ALLE PERIFERICHE

Abbiamo già detto che Linux non fa alcuna distinzione, dal punto di vista dell'utente che accede al *filesystem*, tra dispositivi (ad esempio *hard disk*) che sono fisicamente diversi tra loro. Questo è vero anche per le periferiche come *cd rom* o *floppy disk*. L'unica cosa che l'utente deve fare è dire al sistema come accedere a queste periferiche, dopo di che i dati in esse contenuti divengono disponibili nel *filesystem* esattamente come i *file* contenuti sull'*hard disk*.



Tutto ciò che diremo nel resto del capitolo, ed in particolare questa sezione sull'accesso alle periferiche, dipende fortemente dalla particolare distribuzione di Linux utilizzata e dalla

configurazione impostata dall'amministratore di sistema. Il lettore non deve rimanere sorpreso se alcuni dei comandi non funzionano immediatamente per la sua particolare installazione; nella maggior parte dei casi saranno necessarie solo piccole modifiche.

8.11.1 mount

Questo comando dice al sistema come accedere ad una periferica e come essa deve essere visualizzata all'interno del *filesystem*. Supponiamo per esempio di voler leggere il contenuto di un *cd rom* (che è probabilmente una delle operazioni più comuni che l'utente medio esegue). I comandi:

```
>mount /media/cdrom
>cd /media/cdrom
>
```

dovrebbero essere sufficienti per *montare* il dispositivo (renderlo cioè disponibile nel *filesystem*) e spostarsi nella *directory* corrispondente. Per alcune distribuzioni potrebbe rendersi necessaria la variante:

```
>mount /mnt/cdrom
>cd /mnt/cdrom
>
```

A questo punto il contenuto del *cd rom* dovrebbe essere disponibile nella *directory* stessa di modo che si possono eseguire senza particolari problemi i comandi che già conosciamo.

Il procedimento è sostanzialmente lo stesso nel caso di una *memory stick* USB, anche se qui la situazione si complica leggermente in quanto distribuzioni diverse (o anche distribuzioni identiche configurate in modo differente) possono comportarsi in modo molto diverso. In alcuni casi un utente generico potrebbe anche non avere i permessi necessari per montare il dispositivo. Nel caso più semplice i comandi:

```
>mount /media/pen
>cd /media/pen
>
```

potrebbero funzionare (ovviamente a patto di aver inserito correttamente la *memory stick* in una delle porte USB). Se così non fosse, con ogni probabilità il sistema ha riconosciuto il dispositivo con un nome diverso da "pen" (ad esempio "usb" o "USB_DISK"). Se il comando:

```
ls /media
```

non mostra niente oltre al *cd rom* ed al *floppy*, è probabilmente una buona idea chiedere aiuto all'amministratore di sistema.

8.11.2 umount

Non è il caso di preoccuparsi se non si riesce più ad aprire il drive del *cd rom*! Probabilmente dipende dal fatto che non si è fatto correttamente l'*umount* (cioè lo *smontaggio*) del dispositivo. È sufficiente digitare:

```
>umount /media/cdrom  
>
```

oppure

```
>umount /mnt/cdrom  
>
```

per poter riprendere il controllo della situazione e poter estrarre fisicamente il disco.

SCRIVERE DOCUMENTI SCIENTIFICI: L^AT_EX

INDICE

9.1	Introduzione	179
9.2	Dalla stesura alla stampa	180
9.3	Il primo documento L ^A T _E X	182
9.4	Un documento realistico	184
9.5	Elenchi	187
9.6	Tabelle	188
9.7	L ^A T _E X e la matematica	190
9.8	Inserire le figure	192

L^AT_EX è un programma estremamente potente per la formattazione di documenti di testo, fortemente orientato alla matematica, che gode di una notevole popolarità nella comunità scientifica e tra i fisici in particolare. In questo capitolo il nostro approccio sarà, in un certo senso, minimale: impareremo come si scrive e compila un documento L^AT_EX e come si introducono, all'interno del testo, liste tabelle, equazioni e figure. È tutto ciò che verosimilmente può risultare utile nella redazione di una breve relazione didattica. Per il resto rimandiamo alla documentazione specifica, che è sconfinata, ed in particolare ai riferimenti bibliografici.

9.1 INTRODUZIONE

Prima di entrare nel vivo della discussione, cerchiamo di chiarire alcuni punti fondamentali che spesso confondono il neofita e che distinguono L^AT_EX dai *word processor* più comunemente usati, come OpenOffice o Microsoft Word.

L^AT_EX *non* fa parte dei programmi di formattazione di tipo WYSIWYG¹. L'idea fondamentale che, consapevolmente o meno, sta alla base di questo tipo di programmi (quelli a la Microsoft Word, tanto per intenderci) è la seguente: quello che appare sullo schermo *mentre* scriviamo è sostanzialmente identico al documento finale, come apparirà *dopo* la stampa. Questa idea, che pure sembra così naturale, tende purtroppo a nascondere il fatto che la stesura (cioè la scrittura vera e propria del documento, l'organizzazione concettuale del materiale e la scelta delle parole) e la formattazione (cioè la scelta dei

¹ Questo orrendo acronimo, che pure si trova piuttosto diffusamente nei manuali e nella documentazione, sta per l'espressione inglese *What You See Is What You Get* che, tradotta alla lettera, significa: ciò che vedi è ciò che ottieni.

caratteri, l'organizzazione della veste grafica e l'impaginazione) costituiscono due fasi distinte e concettualmente diverse nella produzione di un documento scritto. È pure evidente che la stesura e la formattazione di un testo richiedono capacità e competenze diverse e che un ottimo scrittore può essere un pessimo tipografo e viceversa.

Facciamo un esempio concreto, tanto per chiarire le idee. Se volessimo dare inizio ad un nuovo capitolo di un nostro ipotetico documento, quel che la maggior parte di noi farebbe con un *word processor* di tipo WYSIWYG sarebbe con ogni probabilità scrivere il titolo, poi selezionarlo, metterlo in grassetto, aumentare un po' la dimensione del carattere e magari, a seconda dei gusti, centrarlo nella pagina. Si tratta di una tipica situazione in cui si rischia, per così dire, di *mischiare* il lavoro dello scrittore e quello del tipografo. Ma vi è un'altra cosa importante: nonostante tutti nostri sforzi e le operazioni di formattazione che abbiamo compiuto, non vi è ancora niente nel nostro titolo che lo qualifichi come tale, nel senso della struttura del documento, e che lo distingua dal resto del testo.

In L^AT_EX l'approccio è radicalmente diverso. Un *file* L^AT_EX è un misto di testo e comandi che sostanzialmente *non* contiene informazioni riguardanti la formattazione (e dunque, per definizione, non ha nemmeno la vaga pretesa di apparire simile all'esito finale della stampa). La formattazione vera e propria avviene invece in un secondo momento, quando il *file* stesso viene processato con un apposito programma. Tanto per tornare all'esempio di prima, il titolo del nostro capitolo in un documento L^AT_EX verrà racchiuso, come vedremo, entro un apposito comando che lo qualifica effettivamente come titolo di un capitolo e che il programma di formattazione è in grado di riconoscere ed interpretare correttamente a seconda delle impostazioni definite dall'utente.

Intendiamoci: in alcuni casi (una lettera di poche righe, ad esempio) vi è talmente poco lavoro di tipografia che probabilmente si può sopravvivere senza L^AT_EX (anche se L^AT_EX ha un bellissimo *template* per le lettere). D'altra parte anche i programmi di formattazione WYSIWYG, se usati *propriamente*, consentono di soddisfare la maggior parte delle esigenze. Ma quando si deve affrontare la redazione di un documento complesso e voluminoso, L^AT_EX ha, tra gli altri, l'indubbio vantaggio di rendere difficile il realizzarlo in modo tipograficamente mal strutturato. Un motivo sufficiente, quanto meno, per provarlo².

9.2 DALLA STESURA ALLA STAMPA

Come abbiamo già detto, un documento L^AT_EX è composto sostanzialmente da un misto di testo e di comandi L^AT_EX validi; ne vedremo numerosi esempi nelle prossime sezioni. In ogni caso la realizzazione del prodotto finale, pronto per la stampa, passa attraverso le operazioni fondamentali (cfr. figura 28) elencate di seguito:

(1) Stesura del documento L^AT_EX. Qualsiasi editor di testo³ va bene allo scopo, ov-

² L^AT_EX è un programma *freeware* (cioè può essere scaricato da web, installato e addirittura redistribuito in modo assolutamente libero) ed *open source* (cioè il codice sorgente è aperto, consultabile e modificabile da tutti).

³ Possiamo citare emacs sotto Linux e notepad (blocco note) sotto Microsoft Windows. È importante notare che i word processor più avanzati non sono editor di testo e non servono quindi allo scopo (provare per

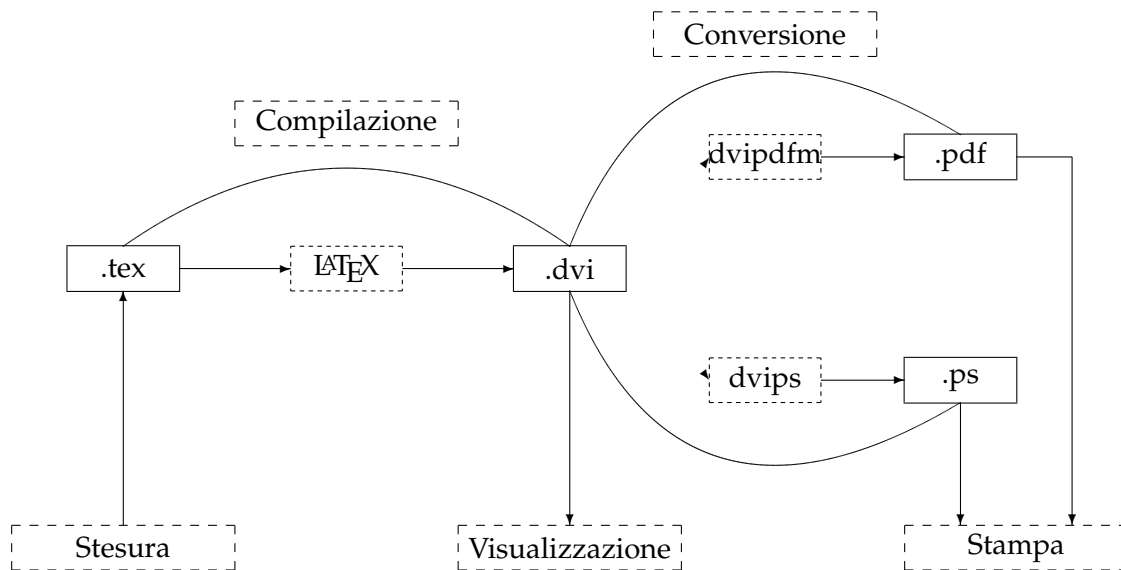


FIGURA 28.. Schema concettuale della preparazione di un documento con \LaTeX . La prima è costituita dalla stesura del documento stesso, che avviene utilizzando un qualsiasi editor di testo. Il programma di compilazione permette di creare un'anteprima di come il documento finale apparirà al momento della stampa. In questa prima fase il documento \LaTeX viene tipicamente compilato un numero cospicuo di volte, fino a che l'autore non è soddisfatto del prodotto finale. A questo punto l'anteprima può essere convertita nei formati pdf o ps per l'archiviazione, la condivisione e la stampa.

viamente, anche se vale la pena di notare che alcuni (sia sotto Linux che sotto Microsoft Windows) possiedono funzionalità specifiche molto utili⁴. Per fissare le idee, se vogliamo creare un nuovo *file* \LaTeX chiamato "Relazione.tex" usando emacs sotto Linux, digiteremo nella *shell*:

```
>emacs Relazione.tex &
```

(2) Compilazione. Si esegue dal terminale⁵ attraverso il comando `latex`. Ad esempio per compilare il *file* "Relazione.tex" digiteremo:

```
>latex Relazione.tex
```

In questo modo, a meno che non vi siano errori di sintassi all'interno del documento \LaTeX (nel qual caso il compilatore si lamenterà e potrete tornare al terminale premendo <CTRL + C>), verrà prodotto un certo numero di *file* con diverse estensioni (dvi, log, credere!).

⁴ Tra questi ricordiamo Kile sotto Linux e TeXniCenter sotto Microsoft Windows.

⁵ Per terminale intendiamo qui sia la *shell* di Linux che il *prompt* di Microsoft DOS nel caso in cui si lavori sotto Microsoft Windows. Notiamo anche esplicitamente che, nel caso in cui si utilizzino editor orientati a \LaTeX come quelli citati prima, vi sono in generale appositi pulsanti nelle rispettive interfacce grafiche per l'utente che permettono di eseguire questa operazione automaticamente.

aux). Il più importante è senza dubbio il *file* con estensione dvi⁶ che, come vedremo tra un attimo, rappresenta una sorta di *anteprima di stampa*.

(3) Visualizzazione dell’anteprima. È necessario utilizzare un programma in grado di leggere correttamente il *file* dvi. Sotto Linux⁷ si utilizza tipicamente il comando `xdvi`:

```
>xdvi Relazione.dvi &
```

(4) Conversione in un formato adatto per la stampa. Quando siamo soddisfatti del nostro documento, come appare nel formato dvi, possiamo convertirlo in un formato adatto per essere archiviato, condiviso o stampato. Tipicamente questo formato può essere il postscript:

```
>dvips Relazione.dvi -o Relazione.ps
```

oppure il pdf:

```
>dvipdfm Relazione.dvi Relazione.pdf
```

Nel primo caso creiamo il *file* postscript “Relazione.ps”, nel secondo il *file* pdf “Relazione.pdf”, che possono essere visualizzati con GhostView ed Acrobat Reader, rispettivamente, oltre che essere stampati con il comando usuale.

9.3 IL PRIMO DOCUMENTO L^AT_EX

Tanto abbiamo detto sulla compilazione e la stampa che probabilmente è giunto il momento di cercare di capire come effettivamente questo documento L^AT_EX vada scritto.

Cominciamo dunque con il più semplice documento che si possa immaginare e scriviamo un *file* di testo contenente le 4 linee seguenti:

```
\documentclass[11pt, a4paper]{article}
\begin{document}
Il mio primo documento\ldots
\end{document}
```

Il risultato della compilazione, essenzialmente pronto per essere stampato, è mostrato in figura 29. Non è niente di esaltante, bisogna ammetterlo... Ma prima di andare avanti e familiarizzare con funzionalità più potenti ed avanzate, cerchiamo di capire il significato di ciascuna delle quattro righe del nostro documento. Una per una.

```
\documentclass[11pt, a4paper]{article}
```

⁶ L’estensione .dvi è l’acronimo dell’espressione inglese *DeVice Independent* che, tradotta letteralmente, significa: indipendente dal dispositivo.

⁷ Sotto Microsoft Windows si usa tipicamente un programma che si chiama `yap` e che viene installato automaticamente al momento dell’installazione di L^AT_EX. Se si usano editor di testo con funzionalità specifiche, essi includono spesso dei bottoni per compiere l’operazione automaticamente.

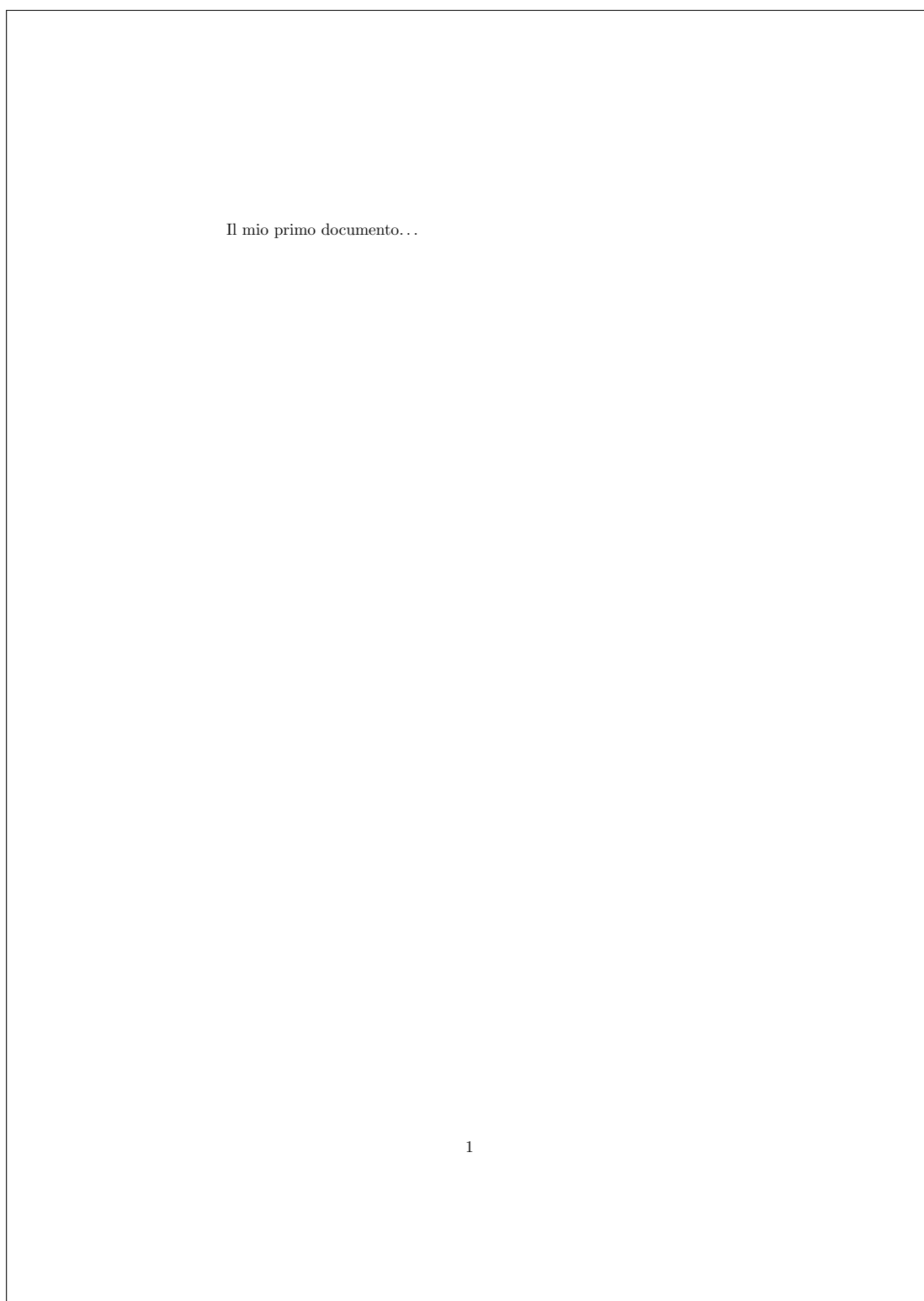


FIGURA 29.. Risultato della compilazione del documento \LaTeX riportato nel paragrafo 9.3.

Qui dichiariamo sostanzialmente il tipo di documento. Diciamo che si tratta di un documento del tipo *article*⁸, che la dimensione dei caratteri del testo ordinario è 11 punti⁹, e che la dimensione della pagina è quella di un foglio A4¹⁰ (che è la scelta tipica).

```
\begin{document}
```

Segna l'inizio del documento vero e proprio. In generale questo comando separa quello che è chiamato *preambolo* (nel nostro caso è sostanzialmente la linea precedente) dal *corpo* vero e proprio del documento, che è *sempre* contenuto tra un comando di inizio ed un comando di fine documento.

```
Il mio primo documento\ldots
```

Si tratta del corpo del nostro documento, ed in effetti è la parte che compare nella versione pronta per la stampa, come si può vedere in figura 29.

```
\end{document}
```

Questa linea chiude il documento ed è sempre richiesta alla fine di un *file* L^AT_EX valido.

Nel prossimo paragrafo vedremo un esempio di documento più realistico e cominceremo ad apprezzare alcune funzionalità più utili ed avanzate.

9.4 UN DOCUMENTO REALISTICO

Consideriamo attentamente il documento L^AT_EX (valido) che segue...

```
\documentclass[11pt, a4paper]{article}
\title{Il mio primo documento}
\author{Luca Baldini}
\date{20 Giugno 2006}
```

```
\begin{document}
\maketitle
```

```
\section{Introduzione}
```

Questo vuole essere un esempio di documento realistico, con lo scopo di mostrare alcune tra le funzionali\`a di base di \LaTeX.

```
\section{Un nuovo capitolo}
```

Una prima cosa da notare \`e che, una volta dichiarate le sezioni del

⁸ Esistono moltissimi tipi diversi di documenti predefiniti in L^AT_EX tra cui *book*, *report*, *letter* e *slides*; altri possono essere aggiunti, anche se non si tratta esattamente della cosa più semplice del mondo. Il tipo *article* si presta piuttosto bene alla stesura di una relazione di media lunghezza su una esperienza didattica ed è questo il motivo per cui lo abbiamo scelto per i nostri esempi.

⁹ Esistono anche le varianti 10 e 12 punti.

¹⁰ Altri formati validi, solo per citarne un paio, sono *a5paper* e *letterpaper*.

documento, `\LaTeX` si occupa automaticamente della numerazione. Questo può risultare estremamente utile nel caso in cui si voglia aggiungere una nuova sezione in mezzo ad un documento in stato avanzato di stesura (non è necessario rinumerare ciò che viene dopo, `\LaTeX` fa tutto da solo).

```
\subsection{Una sotto-sezione}
```

Come vedete esiste anche un livello più basso di sezionamento del documento.

```
\subsection{Ancora una sottosezione}
```

Anche a questo livello la numerazione è automatica. Questo consente anche, in documenti più complessi, la generazione automatica dell'indice.

```
\section{Conclusioni}
```

Per il momento ci fermiamo qui, abbiamo abbastanza di cui discutere...

```
\end{document}
```

...e vediamo come appare l'uscita del compilatore, pronta per la stampa, che è riportata in figura 30. Cominciamo con alcuni commenti di carattere generale per poi passare all'analisi di alcuni tra i nuovi comandi che abbiamo appena incontrato.

Tutti i comandi `LaTeX` cominciano con il carattere `"\"` (*backslash*). Alcuni di essi accettano (o richiedono, a seconda del caso) *opzioni* (che sono sempre contenute tra parentesi quadre) e *argomenti* (racchiusi tra parentesi graffe). I comandi ed il testo sono, come è evidente, mischiati insieme all'interno del documento. È importante notare come gli spazi multipli siano trattati come spazi singoli e come i ritorni a capo siano semplicemente ignorati, se non che una linea vuota segna l'inizio di un nuovo paragrafo. Esiste ovviamente un comando per andare a capo ed è `"\"` (doppio *backslash*).

Tra i comandi rivestono particolare importanza quelli di sezionamento, che sono utilizzati molto frequentemente. I documenti di tipo *article* ammettono i due comandi di sezionamento:

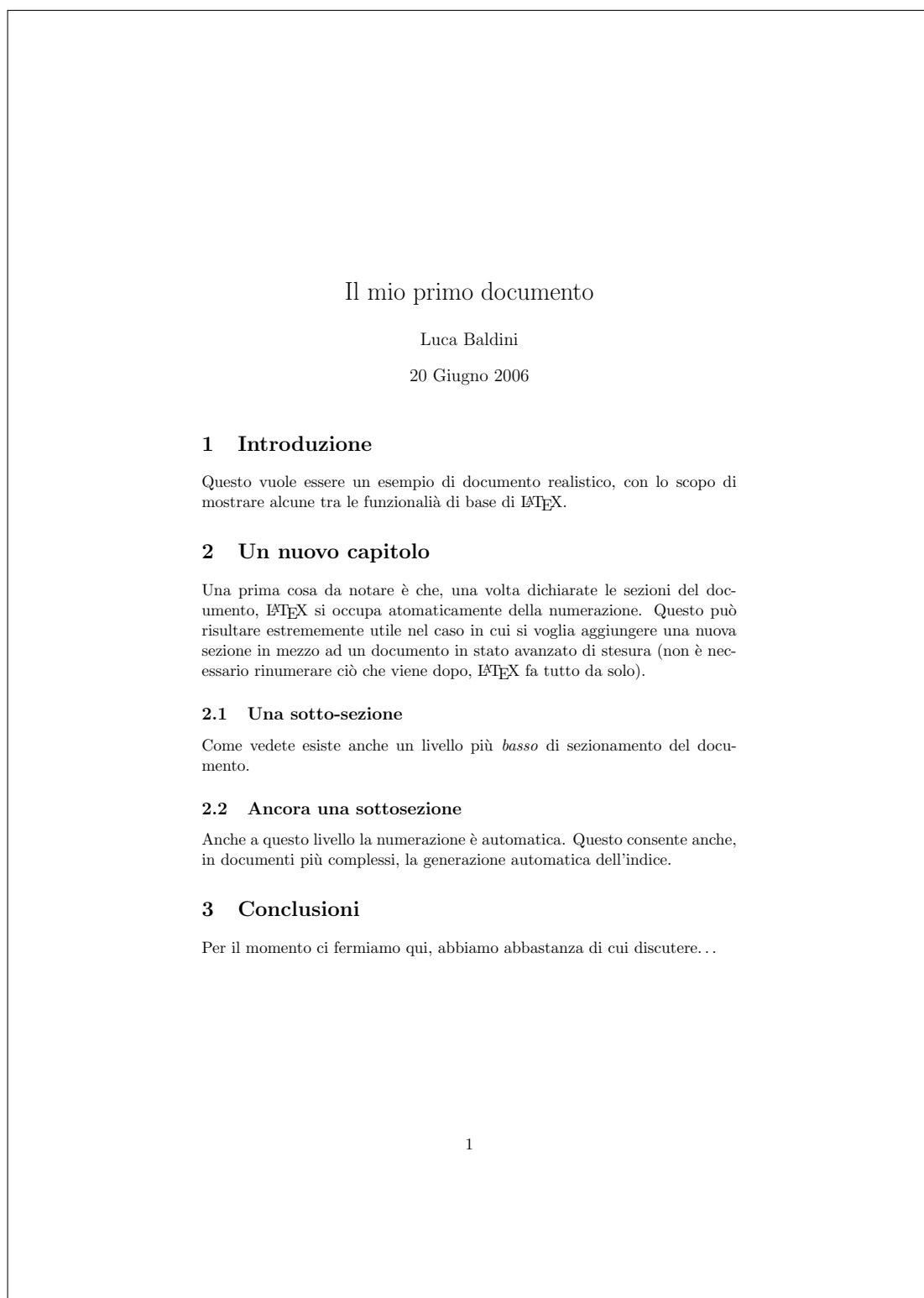
```
\section{}
```

```
\subsection{}
```

che accettano come argomento il titolo della sezione o della sotto-sezione, rispettivamente.

Una volta definito il tipo di documento, `LaTeX` si occupa automaticamente sia di numerare le sezioni e le sotto-sezioni che di fissare le dimensioni e le altre caratteristiche dei caratteri per il testo ordinario e per i titoli. Si tratta di un punto qualificante della filosofia di `LaTeX`: il documento non contiene, al suo interno, alcuna informazione riguardante la formattazione. Ovviamente è possibile fare ogni genere di cambiamento, ma questo avviene ordinariamente nel preambolo attraverso comandi specifici. Rimandiamo ai riferimenti bibliografici per questi argomenti, che sono, in un certo senso, più avanzati.

Veniamo adesso alla descrizione (sintetica, per necessità) di alcuni dei nuovi comandi che abbiamo introdotto.

FIGURA 30.. Risultato della compilazione del documento L^AT_EX riportato nel paragrafo 9.4.

```
\title{Il mio primo documento}
\author{Luca Baldini}
\date{20 Giugno 2006}
```

Questo blocco definisce, come è naturale aspettarsi, il titolo, l'autore e la data di composizione del documento, rispettivamente. Se il comando `date` non viene dato esplicitamente, \LaTeX scrive automaticamente la data del giorno, mutuata dal tempo di sistema del calcolatore. Di *default* questa data viene scritta in lingua inglese, ma si può ovviare a questo inconveniente mettendo nel preambolo il comando `"\usepackage[italian]{babel}"`. Se non si vuole la data è sufficiente scrivere `"\date{}"`.

Vale la pena notare che, di per sé, questi comandi non hanno alcun effetto (provare per credere!), a meno che non siano seguiti, nel corpo del documento, dal comando:

```
\maketitle
```

I più attenti avranno notato come sono state *prodotte* le lettere accentate. Ad esempio i due comandi `"\`e"` e `"\'e"` generano, dopo la compilazione, i due caratteri `"è"` ed `"é"` (cogliamo l'occasione per ricordare che la lingua Italiana possiede due accenti distinti: quello grave e quello acuto). Si tratta di un altro tratto distintivo della filosofia di \LaTeX , per cui si tende a costruire esplicitamente gli oggetti complicati (ad esempio una lettera accentata) a partire da oggetti più semplici (la lettera e l'accento), anziché avere un numero divergente di oggetti (e di tasti, in questo caso...) fin dall'inizio¹¹. Non serve aggiungere che il gioco funziona non solo con la *a*, ma con tutte le vocali (ed anche con le consonanti, sebbene il risultato non sia in genere particolarmente utile).

9.5 ELENCHI

Molte delle esigenze più avanzate del semplice testo vengono gestite in \LaTeX attraverso specifici *ambienti* predefiniti. Tecnicamente un ambiente è una porzione di documento inclusa tra un comando `"\begin{nome_ambiente}"` ed un comando `"\end{nome_ambiente}"`.

Esistono *ambienti*, ad esempio, per la creazione di elenchi puntati e numerati. Le seguenti linee:

```
\begin{itemize}
\item{Punto primo.}
\item{Punto secondo.}
\end{itemize}
```

producono, dopo la compilazione:

- Punto primo.
- Punto secondo.

¹¹ Ad essere onesti \LaTeX offre supporto completo, previa dichiarazione di un apposito pacchetto nel preambolo, per la tastiera Italiana e, quindi, per le ordinarie lettere accentate. Il lettore interessato può trovare le informazioni relative nella documentazione specifica.

D'altra parte questo frammento:

```
\begin{enumerate}
\item{Punto primo.}
\item{Punto secondo.}
\end{enumerate}
```

genera:

1. Punto primo.
2. Punto secondo.

9.6 TABELLE

Le tabelle si costruiscono, secondo la filosofia di L^AT_EX, a partire da elementi *semplici*: essenzialmente linee e caratteri come mostrato nel frammento di codice che segue:

```
\begin{tabular}{cc}
\hline
Colonna 1 & Colonna 2 \\
\hline
\hline
a & b \\
c & d \\
\hline
\end{tabular}
```

e che, compilato, genera:

Colonna 1	Colonna 2
a	b
c	d

Analizziamo in dettaglio queste poche righe una alla volta. La prima:

```
\begin{tabular}{cc}
```

è piuttosto densa di significato. Sostanzialmente segna l'inizio dell'ambiente "tabular" e dichiara una tabella di due colonne. Ognuna delle due "c" sta qui per *center* e sta ad indicare che il contenuto della colonna corrispondente sarà centrato. Il lettore non sarà stupito di sentire che è possibile allineare a destra o a sinistra il contenuto delle colonne usando "r" (*right*) o "l" (*left*), rispettivamente¹².

Le linee orizzontali debbono essere esplicitamente dichiarate tramite il comando "\hline" ed il contenuto vero e proprio della tabella è strutturato come nella riga seguente:

¹² Per completezza ricordiamo che in L^AT_EX è possibile utilizzare il delimitatore "|" per separare le colonne di una tabella con linee verticali—in questo caso particolare potremmo scrivere, ad esempio, "|c|c|" anziché "cc". Tuttavia l'uso di separatori verticali nelle tabelle non è considerata una buona pratica tipografica.

a & b \\\

Il carattere “&” serve da separatore tra celle contigue appartenenti alla stessa linea, mentre il doppio *backslash* (“\\”) funge da terminatore di linea.

Vale la pena, trovandoci in argomento, notare che con poche righe in più:

```
\begin{table}[!htb]
\begin{center}
\begin{tabular}{cc}
\hline
Colonna 1 & Colonna 2 \\\
\hline
\hline
a & b \\\
c & d \\\
\hline
\end{tabular}
\caption{Questa \‘e una tabella di esempio.}
\end{center}
\end{table}
```

è possibile ottenere un risultato estremamente più appagante da un punto di vista estetico (cfr. tabella 1). Esaminiamo le differenze una per una. La più evidente è che

Colonna 1	Colonna 2
a	b
c	d

TABELLA 1.. Questa è una tabella di esempio.

adesso la tabella è centrata. Questo si ottiene banalmente includendola all’interno dei comandi:

```
\begin{center}
\end{center}
```

I più attenti avranno anche notato che abbiamo introdotto un ulteriore nuovo ambiente attraverso le due linee:

```
\begin{table}[!htb]
\end{table}
```

Questo ha due (benefici) effetti distinti. Il primo è che diventa possibile aggiungere una didascalia (automaticamente numerata) alla tabella attraverso il comando:

```
\caption{}
```

Il secondo (e più importante) è che la tabella diventa un oggetto *flottante*, cioè la sua posizione all'interno del documento pronto per la stampa (dopo la compilazione) non è più fissata. L'ambiente "table" è in effetti il primo esempio che incontriamo di ambiente flottante; ne vedremo almeno un altro ("figure") nel seguito.

Molti, all'inizio, si trovano in imbarazzo di fronte agli oggetti flottanti ed hanno l'impressione di non riuscire a prendere il controllo sul posizionamento di figure e tabelle. Si tratta effettivamente di un argomento piuttosto complicato che non abbiamo il tempo di trattare approfonditamente qui—rimandiamo il lettore alle indicazioni bibliografiche. Ci limitiamo a notare che in generale l'inserimento di figure e tabelle in posizioni *fisse* rispetto al corpo del testo è considerata una pessima pratica tipografica. In altre parole non è saggio pretendere di posizionare una figura *esattamente* di seguito ad una determinata linea di testo; dopo l'impaginazione finale quella linea potrebbe essere tanto vicina al margine inferiore della pagina da non lasciare fisicamente abbastanza spazio. E di sicuro non è una cosa a cui si deve pensare *durante* la stesura di un documento—è un lavoro da tipografo, non da scrittore. Il consiglio, qui, è di rilassarsi e di lasciare a L^AT_EX il lavoro sporco; nella maggior parte dei casi sarà perfettamente in grado di inserire gli oggetti flottanti nel modo più *naturale* possibile all'interno del testo. L'autore può aumentare la leggibilità del documento, come vedremo, aggiungendo riferimenti incrociati.

Adesso esaminiamo di nuovo la linea:

```
\begin{table}[!htb]
```

e cerchiamo di capire il significato dell'opzione "!htb" che abbiamo passato al comando. Le tre lettere "h" (*here*), "t" (*top*) e "b" (*bottom*), in ordine, dicono a L^AT_EX dov'è che vorremmo il nostro oggetto flottante; in questo caso L^AT_EX proverà ad inserirlo esattamente dove esso è definito nel corpo del testo (*here*) e, se questo per qualche ragione non fosse possibile, lo inserirà all'inizio (*top*) o alla fine (*bottom*) della prima pagina successiva disponibile. Il punto esclamativo ("!"), in un certo senso, rafforza l'opzione ed *obbliga* L^AT_EX a mettere per un attimo da parte il proprio spiccato senso estetico e ad assecondare la richiesta dell'autore, a meno che non manchi fisicamente lo spazio per farlo. L'opzione "!htb" dovrebbe essere sufficiente per la maggior parte delle applicazioni non troppo avanzate—per quello che può contare tutte le figure di queste dispense sono state inserite in questo modo.

9.7 L^AT_EX E LA MATEMATICA

Veniamo adesso ad uno dei motivi per cui L^AT_EX ha riscosso tanto successo all'interno della comunità scientifica: l'estensivo supporto che fornisce alla scrittura di simboli ed equazioni, che lo rende particolarmente indicato per la scrittura di articoli scientifici¹³.

L^AT_EX offre numerosi ambienti per la matematica; ne citeremo qui solamente due: "math" ed "equation". Il primo permette di inserire simboli matematici all'interno del testo e si apre e chiude essenzialmente con un simbolo "\$". Ad esempio:

¹³ In effetti è probabilmente sensato, di passaggio, ricordare che l'autore originario di T_EX il *motore tipografico* che sta alla base di L^AT_EX è proprio un matematico, Donald E. Knuth


```
\sigma_{x} = \sqrt{\sigma_{x}^2}
```

diviene, dopo la compilazione, $\sigma_x = \sqrt{\sigma_x^2}$. Notiamo, per inciso, il comando “`\sqrt{}`” che permette di inserire radici quadrate di espressioni. Dovrebbe essere altresì chiaro come inserire lettere greche e caratteri in apice e pedice.

L’ambiente `equation` permette di scrivere vere e proprie equazioni (numerate), separate dal corpo del testo¹⁴. Ad esempio:

```
\begin{equation}\label{osc_armonico}
m \frac{d^2x}{dt^2} = m \ddot{x} = -k x
\end{equation}
```


diviene:

$$m \frac{d^2x}{dt^2} = m\ddot{x} = -kx \quad (161)$$

Il lettore potrà essere incuriosito dal comando “`\label{}`”; sostanzialmente esso permette di mettere un’etichetta in corrispondenza di un oggetto¹⁵, o in altre parole di assegnargli un nome comprensibile di modo che in seguito, nel corpo del testo, ci si possa riferire ad esso in modo univoco con l’apposito comando “`\ref{}`”. Se, ad esempio, adesso scriviamo:

```
la (\ref{osc_armonico}) \‘e l’equazione di moto di un oscillatore armonico.
```

dopo la compilazione ciò che otteniamo è: la (161) è l’equazione di moto di un oscillatore armonico. I due comandi “`\label{}`” e “`\ref{}`” sono incredibilmente utili. Pensate a quanti riferimenti numerati ad ogni tipo di oggetto vi sono in queste dispense... senza un meccanismo del genere tenere oggetti e riferimenti sincronizzati sarebbe stato letteralmente un inferno!

 Gli spazi all’interno degli ambienti matematici vengono ignorati. L^AT_EX ha un suo algoritmo piuttosto sofisticato per determinare automaticamente le spaziature necessarie tra i vari oggetti in modalità matematica.

Chiudiamo il paragrafo con un’ultima equazione, che mostra un certo numero di comandi utili:

```
\begin{equation}
\left( \int_{-\infty}^{\infty} e^{-x^2} dx \right)^2 =
\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} =
\prod_{n=1}^{\infty} \left( \frac{n+1}{n} \right)^{(-1)^{n-1}} =
\pi
\end{equation}
```

¹⁴ Potrà apparire curioso, ma non è permesso (provare per credere) lasciare linee vuote all’interno dell’ambiente `equation`.

¹⁵ In questo caso l’oggetto in questione è un’equazione, ma il tutto funziona altrettanto bene con tabelle, figure, titoli di sezioni e paragrafi etc.

che appare come:

$$\left(\int_{-\infty}^{\infty} e^{-x^2} dx\right)^2 = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = \prod_{n=1}^{\infty} \left(\frac{n+1}{n}\right)^{(-1)^{n-1}} = \pi \quad (162)$$

Rimandiamo il lettore alla documentazione specifica per una trattazione esaustiva delle funzionalità offerte da L^AT_EX in modalità matematica.

9.8 INSERIRE LE FIGURE

Eccoci all'ultimo argomento di questa breve rassegna sulle funzionalità di base di L^AT_EX: l'inserimento delle figure. La prima cosa che dobbiamo fare è includere un pacchetto

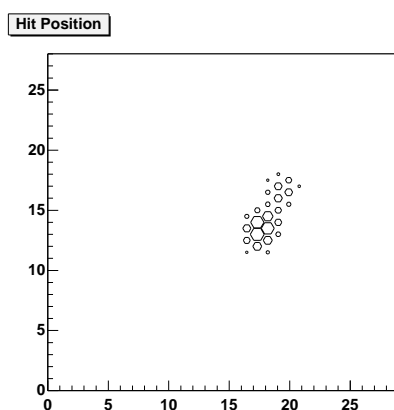


FIGURA 31.. Questa è una figura di esempio.

specifico, e questo si fa inserendo all'interno del preambolo il comando:

```
\usepackage[dvips]{graphicx}
```

Fatto questo, grafici ed immagini in formato eps¹⁶ possono essere inseriti con poche linee (il risultato della compilazione appare in figura 31):

```
\begin{figure}[!htb]
\begin{center}
\includegraphics[width=5.8cm]{./ps_LaTeX/figure/Test.eps}
\end{center}
\caption{Questa \e una figura di esempio.}
\label{figura_di_esempio}
\end{figure}
```

¹⁶ Il formato eps (*Encapsulated PostScript*) è piuttosto usato, specialmente nella comunità scientifica, per la creazioni di grafici, schemi sintetici ed oggetti di grafica vettoriale in generale (sostanzialmente tutto ciò che è costituito da linee, punti e lettere). Vedremo nel prossimo capitolo come creare grafici eps utilizzando il programma di analisi dati gnuplot.

Sostanzialmente non vi è niente di nuovo se non l'ambiente (flottante) "figure", per cui vale quanto detto a proposito di "table", ed il comando:

```
\includegraphics[width=8cm]{./ps_LaTeX/figure/Test.eps}
```

cui è passato il nome del *file* eps come argomento e la larghezza con cui l'immagine in esso contenuta deve apparire nel documento. Notiamo, per inciso, che "\includegraphics" è il comando che richiede il pacchetto "graphicx".

Si possono ovviamente anche inserire figure in formato diverso dal postscript (ad esempio jpg, png etc.). L'unica differenza è che in questo caso bisogna specificare esplicitamente le dimensioni dell'immagine¹⁷ tra le opzioni di "\includegraphics". Se, per esempio, supponiamo di voler inserire una figura in formato jpg di 800×600 pixel, dovremo scrivere qualcosa del tipo:

```
\includegraphics[width=8cm, bb=0 0 800 600]{./Figura.jpg}
```

in cui il frammento "bb=0 0 800 600" fornisce appunto la dimensione della *cornice* che racchiude l'immagine (cioè della *bounding box*), nella forma delle coordinate (espresse in pixel) degli angoli in alto a sinistra ("0 0") ed in basso a destra ("800 600") della cornice stessa. .

¹⁷ Tecnicamente questo è conseguenza del fatto che il formato eps, a differenza degli altri, contiene al suo interno la definizione della cosiddetta *bounding box*.

VISUALIZZARE ED ANALIZZARE DATI: GNUPLOT

INDICE

10.1	Lanciare gnuplot	196
10.2	Due cose semplici (ma utili)...	197
10.3	...e due cose meno ovvie	199
10.4	Visualizzare una serie di dati	201
10.5	Realizzare un istogramma	204
10.6	Salvare i grafici sotto forma di immagini	205
10.7	Il concetto di macro	206
10.8	Eseguire un fit ad una serie di dati	207
10.9	Operazioni con le colonne	212

Inutile dire che di programmi per la visualizzazione e l'analisi di dati scientifici ce sono per tutti i gusti. Nell'imbarazzo della scelta abbiamo optato per gnuplot, guidati da una serie di considerazioni che vale la pena di esporre brevemente.

Per prima cosa gnuplot è *freeware*, *open source* ed è sviluppato ed utilizzato da una comunità decisamente ampia, il che è garanzia di affidabilità e supporto estensivo per l'utente, a partire dalla documentazione.

Inoltre gnuplot è, se paragonato alla maggior parte dei *prodotti* concorrenti, sorprendentemente *semplice*, nel senso che tipicamente permette di eseguire con poca difficoltà tutte quelle operazioni a cui il fisico si trova spesso di fronte (visualizzazione di dati e funzioni, fit, etc.), per lo meno nelle analisi non troppo sofisticate.

Terzo ed ultimo (anche se non in importanza) punto: gnuplot si gestisce completamente da linea di comando attraverso un *macrolinguaggio* nativo e *non* attraverso un'interfaccia grafica. Sebbene questo possa sembrare a prima vista uno svantaggio (retaggio di un tempo ormai andato in cui i calcolatori erano dei deprimenti schermi neri con su qualche scritta incomprensibile) avremo occasione di vedere nel seguito che si tratta invece di una notevole facilitazione quando si voglia eseguire molte volte la stessa sequenza (anche complessa) di operazioni.

10.1 LANCIARE GNUPLOT

Se stiamo lavorando sotto Linux¹, digitare il comando “gnuplot” dovrebbe essere sufficiente per avere indietro una schermata *simile* a questa:

```
>gnuplot
```

```

G N U P L O T
Version 4.0 patchlevel 0
last modified Thu Apr 15 14:44:22 CEST 2004
System: Linux 32 bit

Copyright (C) 1986 - 1993, 1998, 2004
Thomas Williams, Colin Kelley and many others

This is gnuplot version 4.0. Please refer to the documentation
for command syntax changes. The old syntax will be accepted
throughout the 4.0 series, but all savefiles use the new syntax.

Type 'help' to access the on-line reference manual.
The gnuplot FAQ is available from
    http://www.gnuplot.info/faq/

Send comments and requests for help to
    <gnuplot-info@lists.sourceforge.net>
Send bugs, suggestions and mods to
    <gnuplot-bugs@lists.sourceforge.net>

```

```
Terminal type set to 'x11'
gnuplot>
```

Se ciò non accade vi è una probabilità prossima all’unità che il programma non sia correttamente installato (o non sia installato affatto); in questo caso porre rimedio all’inconveniente è facilissimo, essendo gnuplot incluso praticamente in tutte le distribuzioni Linux: sarà sufficiente andare a ripescare i cd da cui si è installato il sistema operativo.

Supponiamo dunque di aver eseguito con successo il nostro comando. L’ultima linea sullo schermo:

```
gnuplot>
```

¹ Sotto Microsoft Windows gnuplot è disponibile sotto forma di eseguibile e tipicamente si lancia con un doppio click di mouse sull’eseguibile stesso. Se le variabili di ambiente sono impostate correttamente (ed in particolare il percorso completo all’eseguibile di gnuplot è contenuto nella variabile d’ambiente PATH), allora il programma si può anche lanciare dal prompt di Microsoft DOS e l’uso diventa essenzialmente identico a quello sotto Linux. Nel seguito cercheremo via via di mettere in evidenza le differenze di impiego tra i due sistemi operativi.

ci dice che ci troviamo all'interno della *shell* di gnuplot e che il programma è pronto in attesa di comandi. Se, come ci viene suggerito nella schermata iniziale, digitiamo il comando²:

```
gnuplot> help
```

questo dovrebbe essere sufficiente per visualizzare l'aiuto interattivo, che è sempre un buon punto di partenza.

10.2 DUE COSE SEMPLICI (MA UTILI)...

Supponiamo di voler disegnare grafico della funzione:

$$f(x) = \frac{\sin x}{x} \quad (163)$$

Ebbene: in gnuplot è sufficiente il singolo comando:

```
gnuplot> plot sin(x)/x
```

per ottenere qualcosa di non troppo lontano da ciò che è mostrato in figura 32:

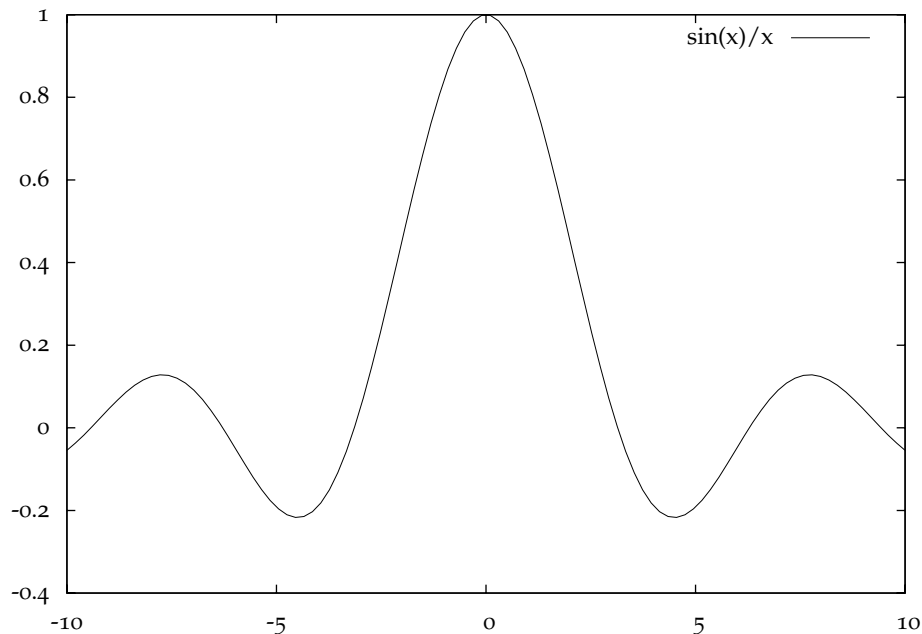


FIGURA 32.. Grafico della funzione (163), prodotto da gnuplot con il semplice comando "plot sin(x)/x".

² Per dovere di chiarezza notiamo esplicitamente, anche se non dovrebbe essercene stretto bisogno, che, qui ed in tutto il seguito del capitolo, il comando propriamente detto è tutto e solo ciò che segue "gnuplot>" (che serve semplicemente ad indicare che, in un determinato istante, ci troviamo entro la *shell* di gnuplot). In questo caso particolare, dunque, l'utente digiterà fisicamente "help" e non "gnuplot> help".

Teniamo a mente il comando “plot”, che abbiamo appena utilizzato, perché si tratta di uno dei comandi fondamentali e lo incontreremo ancora sovente, nel seguito.

Trovandoci a parlare di funzioni, notiamo subito come gnuplot supporti in modo estensivo essenzialmente tutte le funzioni con cui si ha a che fare di solito nell’analisi di dati scientifici (queste includono, ma non sono limitate a, funzioni trigonometriche, esponenziali e logaritmi). Se siete incerti su come si possa scrivere il nome di una funzione in modo che gnuplot *capisca* che cosa volete dire, provate semplicemente a digitare:

```
gnuplot> help functions
```

ed otterrete tutte le informazioni di cui avete bisogno. Notiamo anche come le funzioni possano essere poi *combinare* ulteriormente tra di loro con gli usuali operatori aritmetici che si trovano elencati in tabella 2.

Operatore	Simbolo
Addizione	+
Sottrazione	-
Moltiplicazione	*
Divisione	/
Elevamento a potenza	**

TABELLA 2.. Associazione tra gli operatori aritmetici di base ed il simbolo corrispondente all’interno della *shell* di gnuplot.

Spulciando tra le funzioni disponibili all’interno di gnuplot troviamo anche, come era logico aspettarsi, la *error function* che abbiamo introdotto a proposito della distribuzione di Gauss e che si trova descritta in dettaglio in appendice A.1 e tabulata nelle appendici A.2 e A.3. In effetti possiamo leggere direttamente dalle tavole (cfr. A.2):

$$\text{erf}(1) = 0.8413$$

Proviamo allora ad ottenere lo stesso risultato utilizzando gnuplot; per far questo ci servirà, come vedremo, il comando “print”, che sostanzialmente consente di stampare sullo schermo³ il valore di un’espressione, passata come argomento. Digitiamo:

```
gnuplot> print erf(1)
0.842700790029219
```

Vi è chiaramente qualcosa che non va, perché il numero che otteniamo non è corretto. Ma se leggiamo attentamente il contenuto dell’appendice A.1, vi troviamo un esplicito riferimento al fatto che ci sono due definizioni differenti della *error function*, $\text{erf}(x)$ ed $\text{Erf}(x)$, legate dalla relazione:

$$\text{erf}(x) = \frac{1}{2} \text{Erf} \left(\frac{x}{\sqrt{2}} \right)$$

³ Ad essere rigorosi questo non è necessariamente vero. Volendo, l’uscita del comando print può essere rediretta su un *file* attraverso l’apposito comando “set print”.

Nelle tavole numeriche noi utilizziamo la prima, ma potrebbe darsi che gnuplot usi l'altra. Proviamo allora a digitare:

```
gnuplot> print 0.5*erf(1/sqrt(2))
0.341344745680573
```

Questa volta abbiamo il risultato corretto, il che conferma la nostra ipotesi. Notiamo per inciso che, sebbene non costituisca il modo più efficiente, il comando "print" permette, in linea di principio, di rigenerare le tavole numeriche riportate nelle sopraccitate appendici. Più semplicemente, esso permette di evitare l'interpolazione tra valori tabulati, ogni qual volta si abbia un calcolatore a disposizione.

10.3 ...E DUE COSE MENO OVVIE

Vi sono situazioni in cui il calcolatore sembra fornire all'utente dei risultati inaspettati, apparentemente errati, o addirittura assurdi. Lo scopo dichiarato di questo paragrafo è di mostrare come, nella stragrande maggioranza di questi casi, non sia il computer a sbagliare; più semplicemente l'utilizzatore ha posto al computer una domanda che è effettivamente *diversa* da quella che egli stesso, in buona fede, credeva aver posto. Si può dire, a buon diritto, che in questa difficoltà di comunicazione tra utente e calcolatore risieda l'origine di un gran parte degli errori di programmazione. Che è come dire che, una volta pensata una *domanda* che si vuol porre al calcolatore, il punto fondamentale è formularla in modo corretto nel linguaggio del calcolatore stesso.

Supponiamo dunque di voler disegnare il grafico della funzione (163) in un intervallo specifico della variabile indipendente x . Per fissare le idee, supponiamo che questo intervallo sia $[-170, 170]$. Vi è un comando specifico, "set xrange", per fissare i limiti di variabilità della x , per cui proveremo a scrivere:

```
gnuplot> set xrange [-170:170]
gnuplot> plot sin(x)/x
```

Il risultato, che è mostrato in figura 33, non può non destare perplessità. Sicuramente il gran numero di cuspidi confligge palesemente con il fatto ben noto che la funzione è in effetti derivabile su tutta la retta reale. Eppure ciò che abbiamo fatto è così semplice (si tratta in fondo di due soli comandi) e naturale che non è ovvio, a prima vista, il punto esatto in cui abbiamo sbagliato.

La questione ci riporta per un attimo indietro ad analizzare più in dettaglio il comando "plot". La domanda che ci facciamo è, in altre parole: che cosa fa effettivamente gnuplot quando noi impartiamo un comando "plot"? Beh, per dirla in breve, essenzialmente divide l'intervallo di variabilità della variabile indipendente x in un certo numero di intervalli, calcola il valore della funzione in corrispondenza degli estremi di questi intervalli (in termini tecnici si dice che *campiona* la funzione) e quindi unisce con delle spezzate i valori calcolati. Questo ci suggerisce che l'aspetto del grafico in figura 33 possa essere in effetti dovuto ad un semplice problema di visualizzazione, connesso con il fatto che abbiamo campionato la funzione in modo troppo grossolano.

4 Se non dichiariamo esplicitamente questo intervallo, gnuplot lo sceglie da solo, preliminarmente.

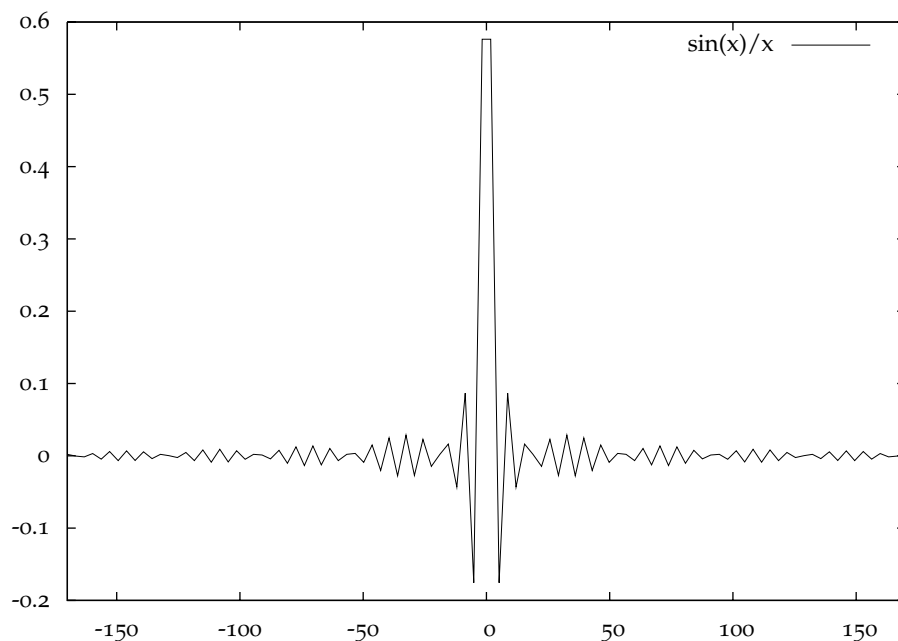


FIGURA 33.. Grafico della funzione (163), prodotto da gnuplot, nell'intervallo $[-170, 170]$.

Leggendo la documentazione impariamo che, di *default*, il numero di punti in cui una funzione è campionata per il grafico è esattamente 100. Questo numero può, come ogni altra cosa, essere modificata dall'utente; maggiore è il numero di punti richiesti, più lungo il tempo necessario per la visualizzazione, come è logico aspettarsi. Ad ogni modo il comando specifico per questa operazione è "set samples", che accetta come argomento il numero di campionamenti. Possiamo allora provare:

```
gnuplot> set samples 10000
gnuplot> set xrange [-170:170]
gnuplot> plot sin(x)/x
```

ed il risultato, finalmente ragionevole, è mostrato in figura 34.

Torniamo allora per un istante a quanto detto all'inizio del paragrafo e cerchiamo di interpretare i due ultimi grafici nel nostro nuovo schema concettuale. In buona fede eravamo convinti di chiedere al computer di disegnare il grafico di una funzione; in realtà: nel primo caso gli abbiamo chiesto di campionare la funzione in 100 punti e di unire i valori ottenuti con linee rette, mentre nel secondo gli abbiamo chiesto di fare lo stesso campionando in 10000 punti. In entrambi i casi il calcolatore ha eseguito gli ordini alla lettera; il fatto che il primo grafico non ci piaccia, mentre il secondo sia esteticamente appagante dipende solo da noi.

Avevamo promesso esattamente *due* cose meno ovvie nel titolo del paragrafo ed ecco qui la seconda. Proviamo a digitare nella *shell* di gnuplot il comando "print 1/2". Riuscite ad immaginare che cosa succederà? Ebbene, il risultato è, ancora una volta, sorprendente:

```
gnuplot> print 1/2
```

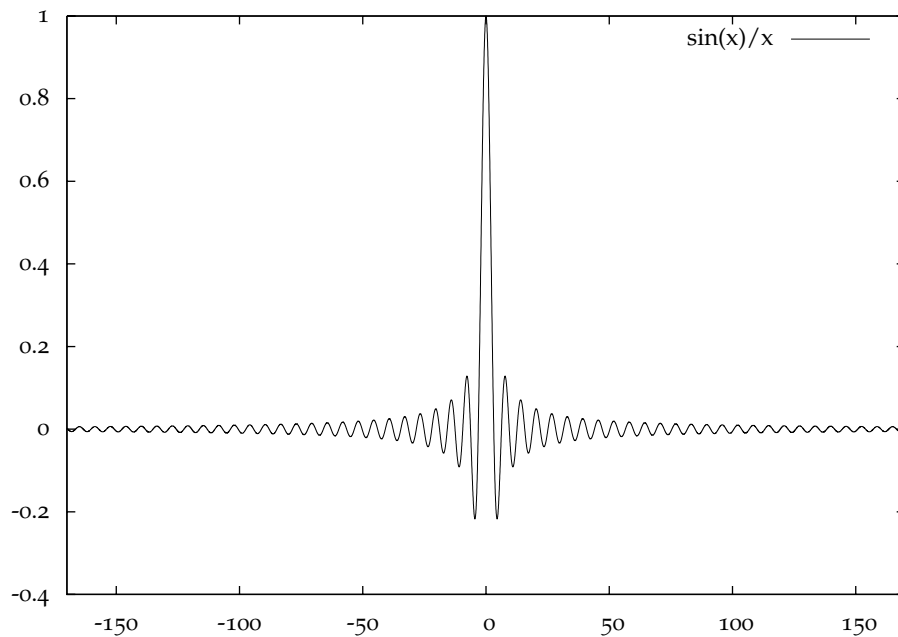


FIGURA 34.. Grafico della funzione (163), prodotto da gnuplot, nell'intervallo $[-170, 170]$, con un numero di punti di campionamento pari a 10000.

0

Senza perdere troppo tempo, il punto fondamentale è che gnuplot interpreta numeratore e denominatore come numeri interi e restituisce pure il risultato della divisione come intero (in effetti 0 è la parte intera di $\frac{1}{2}$). Si tratta di una *trappola* tipica in cui tutti sono caduti almeno una volta nella vita ed è questo il motivo per cui l'abbiamo mostrata qui. Se adesso rendiamo manifestamente reali il numeratore ed il denominatore, tutto torna come ci aspettiamo:

```
gnuplot> print 1.0/2.0
0.5
```

10.4 VISUALIZZARE UNA SERIE DI DATI

gnuplot è capace di visualizzare in modo molto semplice i dati contenuti in un *file*, purché organizzati per colonne e separati con spazi o <TAB>. Tutte le linee che cominciano con un cancelletto (“#”) sono automaticamente ignorate da gnuplot (e tecnicamente prendono il nome di *commenti*). Questo permette, tra le altre cose, di inserire in un generico *file* di dati una breve descrizione del contenuto delle colonne, di modo che esso sia più facilmente comprensibile.

Supponiamo ad esempio di avere un *file* di testo (denominato “LeggeOraria.txt”) del tipo:

```
# Tempo (sec)    Posizione (cm)    Errore (cm)
```

5	10.1	2.0
10	17.8	1.5
15	26.6	1.5
20	32.8	2.0
25	40.2	2.5
30	47.4	2.0

Il comando:

```
gnuplot> plot 'LeggeOraria.txt' using 1:2
```

dovrebbe generare qualcosa di non troppo diverso dalla figura 35. In questo contesto

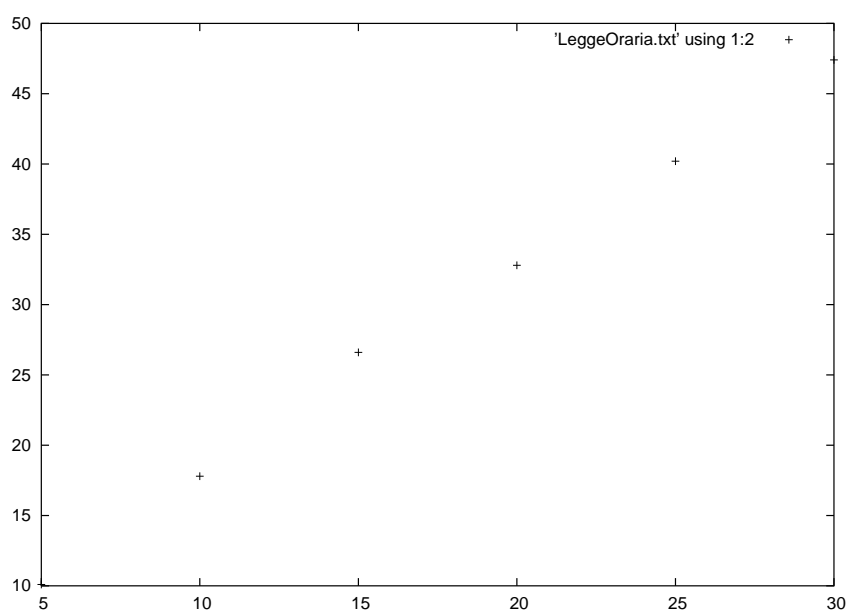


FIGURA 35.. Visualizzazione, mediante gnuplot, dei dati contenuti nel file “LeggeOraria.txt” (riportato poche righe sopra).

il comando “plot” accetta come argomento il nome del *file* di testo contenente i dati (racchiuso tra virgolette o apici) e l’opzione “using” specifica le colonne che saranno graficate sull’asse delle x e su quello delle y , rispettivamente. Si tratta di una sintassi tipica che useremo spesso; sottolineiamo anche esplicitamente il fatto che il comando “plot” supporta una notevole quantità di opzioni, come si può facilmente verificare digitando “help plot”.

⚠ È molto frequente, specialmente all’inizio, far confusione tra percorsi relativi ed assoluti ai file. Quando scriviamo “plot ‘LeggeOraria.txt’” assumiamo implicitamente che un file denominato “LeggeOraria.txt” sia presente nella directory corrente.

⚠ Se non abbiamo mai usato il comando “cd” all’interno della shell di gnuplot, la directory corrente sarà la directory da cui abbiamo lanciato il programma. In ogni caso essa può

essere visualizzata tramite il comando “pwd”, anche questo da digitare nella shell di gnuplot.

Quando cerchiamo di accedere ad un file che fisicamente non esiste (o che non si trova nella posizione esatta cui stiamo puntando all'interno del *filesystem*) otteniamo in generale un messaggio d'errore. Tanto per fissare le idee, supponiamo di aver lanciato gnuplot da una certa *directory* e di avere il file di dati “LeggeOraria.txt” all'interno di una *directory* denominata “dati” e situata a sua volta all'interno della *directory* di partenza. Digitando il comando precedente provocheremo ovviamente un errore:

```
gnuplot> plot 'LeggeOraria.txt' using 1:2
~
can't read data file "LeggeOraria.txt"
```

perché gnuplot non troverà fisicamente il *file*. Abbiamo allora due possibilità distinte: scrivere il percorso completo al *file*:

```
gnuplot> plot './data/LeggeOraria.txt' using 1:2
```

oppure cambiare *directory* e successivamente impartire il comando “plot”:

```
gnuplot> cd './data'
gnuplot> plot 'LeggeOraria.txt' using 1:2
```



Il lettore non deve far confusione tra il comando “cd” di gnuplot ed il comando standard cd di Linux. In particolare il primo richiede che il nome della directory di destinazione sia racchiuso tra virgolette o apici.

Ogni volta che si ha un messaggio di errore del tipo “can't read data file ...” significa che, in un modo o nell'altro, stiamo puntando ad un *file* che fisicamente non si trova dove noi pensiamo (o che magari non ci siamo nemmeno posti il problema). In tal caso le cose da fare sono due: capire dove siamo con il comando “pwd” (da digitare nella *shell* di gnuplot) e quindi indirizzare il *file* correttamente in uno dei due modi appena mostrati.

Torniamo adesso alla nostra serie di dati e vediamo brevemente alcuni comandi specifici di formattazione dei grafici. Supponiamo di digitare, nella *shell* di gnuplot:

```
gnuplot> set title 'Legge oraria per una particella libera'
gnuplot> set xlabel 'Tempo (sec)'
gnuplot> set ylabel 'Posizione (cm)'
gnuplot> set xrange [0: 40]
gnuplot> set yrange [0: 50]
gnuplot> unset key
gnuplot> plot 'LeggeOraria.txt' using 1:2:3 with yerrorbars
```

Il risultato è mostrato in figura 36. La maggior parte dei nuovi comandi introdotti dovrebbe essere auto-esplicativa. In particolare dovrebbe essere chiaro come inserire il titolo del grafico, impostare gli estremi dei due assi e mettere delle *etichette* di testo (in inglese si chiamano *label*) sugli assi stessi. Pur tuttavia commenteremo esplicitamente gli ultimi due. Il primo:

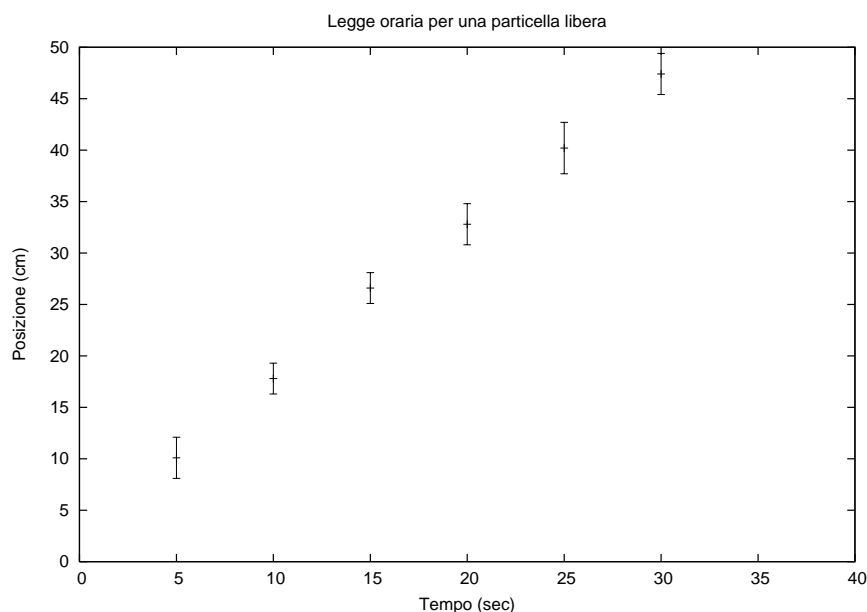



FIGURA 36.. Visualizzazione, mediante gnuplot, dei dati contenuti nel file “LeggeOraria.txt”, dopo alcuni semplici comandi di formattazione.

```
gnuplot> unset key
```

sostanzialmente elimina la legenda, che altrimenti viene inserita nel grafico in alto a destra (cfr. figura 35). L'altro è l'usuale comando “plot”:

```
gnuplot> plot 'LeggeOraria.txt' using 1:2:3 with yerrorbars
```

che viene però usato, questa volta, con l'opzione “with yerrorbars”, la quale permette di rappresentare sul grafico le barre d'errore sull'asse delle y ; la colonna dei valori per le barre d'errore viene passata come terzo argomento dell'opzione “using”. Esiste anche un'opzione “with xerrorbars” che permette di rappresentare le barre d'errore su entrambi gli assi; in questo caso le quattro colonne che sono passate all'opzione “using” sono interpretate come x , y , Δx e Δy , rispettivamente.

 Tutti i comandi di formattazione hanno effetto solamente nel momento in cui si impartisce un nuovo comando “plot” e non prima; di converso il loro effetto cessa solamente nel momento in cui vengono esplicitamente sovrascritti oppure quando si esce da gnuplot. Vale anche la pena di menzionare il comando “replot” il quale non fa altro che ripetere esattamente l'ultimo comando “plot”.

10.5 REALIZZARE UN ISTOGRAMMA

Diciamo subito che gnuplot *non* supporta la realizzazione di istogrammi, nel senso che non ha le funzionalità per creare i canali e popolarli a partire da una serie di dati. Pur tuttavia esiste un'opzione di *visualizzazione* del comando “plot” dedicata

specificamente alla creazione istogrammi⁵. Senza entrare troppo nei dettagli diremo che se forniamo a gnuplot un *file* di dati (supponiamo, per fissare le idee, che si chiami “Istogramma.txt”) contenente una colonna con le coordinate centrali dei canali ed una con il contenuto dei canali stessi, possiamo disegnare l’istogramma corrispondente con un comando del tipo:

```
gnuplot> plot 'Istogramma.txt' using 1:2 with histeps
```

10.6 SALVARE I GRAFICI SOTTO FORMA DI IMMAGINI

Il lettore potrà chiedersi come sia stato possibile salvare i grafici generati da gnuplot in una forma che permettesse di inserirli nel paragrafo precedente (e precisamente sotto forma di *file* postscript). Prima di mostrare esplicitamente i comandi che serviranno è tuttavia necessario introdurre due concetti nuovi di fondamentale importanza:

(1) “output”: è sostanzialmente il luogo in cui vengono mostrate (o, più precisamente *redirette*) le schermate di gnuplot (in Italiano diremmo *uscita*). Nel nostro caso esso può essere, a seconda dei casi, il monitor del calcolatore oppure un *file*⁶. Vedremo tra breve come passare da una modalità all’altra.

(2) “terminal”: è il *linguaggio* con cui le schermate di gnuplot vengono redirette sull’uscita (qualunque essa sia). Nel caso l’uscita sia un *file*, questo linguaggio può essere, ad esempio postscript, gif, png, etc. Nel caso in cui invece vogliamo, come accade tipicamente, le schermate sul monitor, questo linguaggio sarà, in un qualche senso, il linguaggio del sistema operativo; come vedremo nel seguito, il nome da usare sarà “x11” sotto Linux e “windows” sotto Microsoft Windows.

A questo punto non dovrebbe essere troppo difficile capire come si possa salvare un grafico sotto forma di immagine: dovremo indirizzare l’uscita su di un *file* specifico ed impostare opportunamente il terminale. Vale a dire qualcosa del genere:

```
gnuplot> set terminal postscript
gnuplot> set output 'Immagine.eps'
gnuplot> replot
gnuplot> set output
gnuplot> set terminal x11
```



I comandi “set output” e “set terminal” non sono indipendenti l’uno dall’altro: non tutti i terminali sono appropriati per tutti i tipi di output e viceversa. Il lettore potrà divertirsi, per curiosità, a provare alcune delle combinazioni proibite (come terminale postscript su output standard oppure terminale “x11” su di un file). Gli effetti saranno interessanti.

Analizziamo i comandi appena scritti uno per uno, alla luce di quanto detto sino ad ora. I primi due

⁵ Sottolineiamo esplicitamente che si tratta solo di un’opzione di visualizzazione e che la definizione ed il popolamento dei canali debbono essere eseguiti dall’utente.

⁶ Vi sono in effetti alcune altre possibilità un poco più esotiche che potete trovare nella documentazione specifica.

```
gnuplot> set terminal postscript
gnuplot> set output 'Immagine.eps'
```

redirigono l'uscita su di una *file* che si chiamerà “Immagine.eps”⁷ e che verrà scritto in linguaggio postscript, come specificato nella impostazione del terminale. Il comando successivo:

```
gnuplot> replot
```

esegue di nuovo l'ultimo comando plot che è stato digitato. In questo caso, essendo l'*output* rediretto su *file*, il grafico non apparirà sullo schermo; al contrario verrà fisicamente *scritto* sul *file* stesso. Quando torniamo sull'*output* standard, vale a dire sullo schermo del calcolatore:

```
gnuplot> set output
```

il *file* di uscita viene *chiuso*. L'ultimo comando permette di impostare di nuovo il terminale nella modalità di default:

```
gnuplot> set terminal x11
```

e siamo di nuovo al punto di partenza, in cui i comandi “plot” producono bellissimi grafici sullo schermo. Vale la pena notare che, se lavorassimo sotto Microsoft Windows, quest'ultimo comando (e solo questo) sarebbe diverso:

```
gnuplot> set terminal windows
```

10.7 IL CONCETTO DI MACRO

Adesso che il nostro vocabolario, in termini di comandi di gnuplot, si è fatto abbastanza ampio da consentirci di produrre una certa quantità di cose interessanti, diventa davvero indispensabile introdurre il concetto di *macro*. In termini semplici una macro è una sequenza di comandi (validi) inseriti, uno dopo l'altro, in un *file* di testo. Il vantaggio fondamentale del lavorare con le macro è che, mentre tutti i comandi che digitiamo nel prompt vanno perduti non appena usciamo dal programma, una macro, una volta salvata, può essere riutilizzata (eventualmente dopo modifiche) a distanza di tempo. Come vedremo, è sufficiente un solo comando per eseguirla.

Una tipica macro che utilizzi i comandi imparati sino ad ora potrebbe apparire come segue:

```
# Inizializzazione.
set output
set terminal x11

# Definizione e formattazione del grafico.
```

⁷ Se non specifichiamo il percorso completo il *file* verrà creato nella *directory* corrente. È importante anche notare che, nel caso esista già un *file* con lo stesso nome, esso verrà sovrascritto senza che ne venga richiesta conferma.


```

set title 'Legge oraria per una particella libera'
set xlabel 'Tempo (sec)'
set ylabel 'Posizione (cm)'
set xrange [0: 40]
set yrange [0: 50]
unset key
plot 'LeggeOraria.txt' using 1:2:3 with yerrorbars

# Creazione del file di uscita.
set terminal postscript
set output 'Immagine.eps'
replot

# Ritorno allo schermo.
set output
set terminal x11

```

Dato che, come già sappiamo, le linee che cominciano con un cancelletto vengono ignorate da gnuplot, nella macro qui sopra ne abbiamo usate alcune per commentare la macro stessa. Si tratta di una buona abitudine che, unita all'uso di linee vuote per separare blocchi logici distinti, aumenta la leggibilità. Notiamo esplicitamente come i primi due comandi:

```

gnuplot> set output
gnuplot> set terminal x11

```

pur se non formalmente necessari, costituiscano pure una buona abitudine, che può salvare da apparenti inconsistenze nel caso in cui le impostazioni di partenza non siano, per qualche ragione, quelle che ci aspettiamo.

Supponendo dunque che il *file* di testo in cui abbiamo salvato la macro si chiami "macro.txt", essa si esegue digitando nella *shell* di gnuplot:

```

gnuplot> load 'macro.txt'

```

Questo dovrebbe produrre sostanzialmente la figura 36.

10.8 ESEGUIRE UN FIT AD UNA SERIE DI DATI

gnuplot permette di eseguire un fit del minimo χ^2 ad una serie di dati⁸ con una generica funzione definita dall'utente. I comandi coinvolti in questa operazione sono, più o meno, i seguenti:

```

gnuplot> f(x) = a + b*x

```

⁸ Per essere precisi, gnuplot utilizza una propria implementazione dell'algoritmo di Marquardt-Levenberg per il fit non lineare dei minimi quadrati (NLLS). I punti sono pesati con l'inverso del loro errore, per cui questo metodo è in effetti concettualmente affine a quello che noi conosciamo come metodo del minimo χ^2 .

```
gnuplot> a = 1
gnuplot> b = 1
gnuplot> fit f(x) 'LeggeOraria.txt' using 1:2:3 via a, b
```


Al solito li analizzeremo brevemente uno ad uno. Il primo:

```
gnuplot> f(x) = a + b*x
```

definisce la funzione con cui vogliamo eseguire il fit. In questo caso è semplicemente una retta, ma si possono utilizzare tutte le funzioni predefinite in gnuplot, eventualmente combinate attraverso gli operatori in tabella 2. Il che è sufficiente per la stragrande maggioranza delle applicazioni. I due comandi successivi:

```
gnuplot> a = 1
gnuplot> b = 1
```


inizializzano i parametri della funzione con cui si fa il fit.

 Il modo in cui il calcolatore esegue un fit è, per molti aspetti, diverso da come faremmo noi, seguendo i metodi descritti nel capitolo 5. In particolare, anche nei casi in cui le equazioni a cui porta il metodo del minimo χ^2 si possono risolvere analiticamente, il calcolatore utilizza dei metodi numerici per cui i valori iniziali dei parametri sono essenziali. Vi sono casi, nemmeno troppi complicati, in cui il fit può convergere o non convergere, a seconda di come inizializziamo i parametri. Vi sono anche casi, per certi versi ancor più pericolosi, in cui valori iniziali diversi fanno convergere il fit a valori finali diversi. In ogni caso, è una buona abitudine fornire una stima sensata dei parametri (magari solo guardando il grafico) prima di cominciare il fit.

L'ultimo comando:

```
gnuplot> fit f(x) 'LeggeOraria.txt' using 1:2:3 via a, b
```

è quello che esegue il fit propriamente detto. Esso accetta come parametro la funzione con cui fittare i dati ed il nome del *file* in cui i dati stessi sono contenuti. Le tre colonne che vengono passate all'opzione "using" sono interpretate come valori della x , della y e degli errori sulla y , rispettivamente. La terza colonna (quella degli errori sulla y) non è strettamente richiesta.

 Se gli errori sulla y non vengono forniti, gnuplot assume che essi siano tutti pari ad 1; questo è rilevante (è importante ricordarlo) per il valore del χ^2 che, come vedremo, il fit fornisce come sottoprodotto; gnuplot non supporta fit di tipo generale (vale a dire con errori su entrambe le variabili).

L'opzione "via" definisce quali tra i parametri della funzione siano liberi di variare durante il fit (quelli che eventualmente non sono menzionati sono considerati fissi al loro valore iniziale).

Se le cose non fossero ancora perfettamente chiare, la macro seguente:

```

# Inizializzazione.
set output
set terminal x11

# Definizione e formattazione del grafico.
set title 'Legge oraria per una particella libera'
set xlabel 'Tempo (sec)'
set ylabel 'Posizione (cm)'
set xrange [0: 40]
set yrange [0: 50]
unset key

# Esecuzione del fit
f(x) = a + b*x
a = 1
b = 1
fit f(x) 'LeggeOraria.txt' using 1:2:3 via a, b

# Creazione del grafico.
plot 'LeggeOraria.txt' using 1:2:3 with yerrorbars, f(x)

# Creazione del file di uscita.
set terminal postscript
set output 'Immagine.eps'
replot

# Ritorno allo schermo.
set output
set terminal x11

```

dovrebbe produrre, ammesso che tutto funzioni correttamente, il grafico in figura 37. Notiamo come il comando “plot” permetta di graficare più oggetti (in questo caso una serie di dati ed una funzione), a patto che essi siano separati da una virgola.

Contestualmente al grafico (mostrato sullo schermo e salvato su di un *file*) il comando “fit” produrrà una serie di linee sul terminale di gnuplot, contenenti informazioni sull’esito del fit stesso⁹:

```

*****
Wed Feb 12 09:45:22 2003

FIT:    data read from 'LeggeOraria.txt' using 1:2:3
        #datapoints = 6

```

⁹ Queste informazioni sono anche salvate, per futura referenza, su un *file* di testo chiamato “fit.log” e salvato nella *directory* corrente.

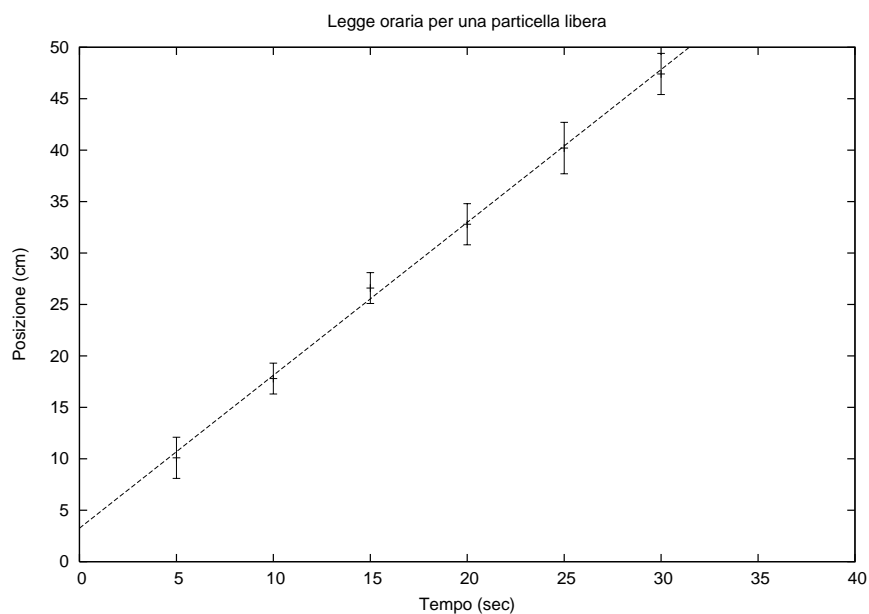


FIGURA 37.. Grafico prodotto dalla macro riportata in questa sezione.

```
function used for fitting: f(x)
fitted parameters initialized with current variable values
```

```
Iteration 0
WSSR      : 209.004          delta(WSSR)/WSSR   : 0
delta(WSSR) : 0             limit for stopping : 1e-05
lambda    : 6.93718
```

```
initial set of free parameter values
```

```
a          = 1
b          = 1
```

```
After 5 iterations the fit converged.
final sum of squares of residuals : 0.686015
rel. change during last iteration : -2.87956e-12
```

```
degrees of freedom (ndf) : 4
rms of residuals      (stdfit) = sqrt(WSSR/ndf)      : 0.41413
variance of residuals (reduced chisquare) = WSSR/ndf : 0.171504
```

```
Final set of parameters          Asymptotic Standard Error
=====                          =====
```

```

a          = 3.25533          +/- 0.698          (21.44%)
b          = 1.48589          +/- 0.03902         (2.626%)

```

correlation matrix of the fit parameters:

```

          a      b
a          1.000
b        -0.897  1.000

```

Le informazioni più importanti sono i valori finali dei parametri che si trovano dopo la riga:

Final set of parameters	Asymptotic Standard Error
=====	=====
a = 3.25533 +/- 0.698 (21.44%)	
b = 1.48589 +/- 0.03902 (2.626%)	

la matrice di correlazione dei parametri, che abbiamo definito nel paragrafo 7.1 e che leggiamo sotto la riga:

correlation matrix of the fit parameters:

```

          a      b
a          1.000
b        -0.897  1.000

```

ed il valore del χ^2 diviso il numero di gradi di libertà:

variance of residuals (reduced chisquare) = WSSR/ndf : 0.171504

Vale la pena di ricordare che, nel caso in cui non si specifichi esplicitamente una colonna di dati contenente gli errori sulle y , il valore del χ^2 (che gnuplot fornisce comunque) perde completamente di significato (in tal caso, come abbiamo già detto, il programma assume semplicemente che tutti gli errori siano pari all'unità).

Come commento finale vorremmo sottolineare che il fatto che il calcolatore scriva esplicitamente:

```

a          = 3.25533          +/- 0.698          (21.44%)

```

non ci autorizza a dimenticare le regole consuete sulle cifre significative. Noi scriveremo sempre:

$$a = 3.25 \pm 0.70$$

10.9 OPERAZIONI CON LE COLONNE

Vi è un'ultima cosa che, essendo utile in molte situazioni, vale la pena di discutere brevemente e si tratta delle operazioni che gnuplot consente di eseguire con le colonne di un generico *file* di dati.

Supponiamo, per fissare le idee, di avere un *file* di dati (come al solito organizzato per colonne) e di voler fare un grafico con la colonna numero 1 sull'asse delle x ed il *quadrato* della colonna numero 2 su quello delle y . Ebbene, in questo caso scriveremmo:

```
gnuplot> plot 'LeggeOraria.txt' using 1:($2**2)
```

Non è il caso, di nuovo, di dilungarsi in spiegazioni. Essenzialmente la scrittura “(\$2)” identifica la seconda colonna come oggetto nel suo insieme; dopo di che si può utilizzare la consueta sintassi (cioè le funzioni e gli operatori che già conosciamo) per fare operazioni sulle colonne stesse, elemento per elemento. A volte, specialmente se le operazioni da fare sulle colonne sono complesse, se si vuole vedere il risultato di tali operazioni (quanto meno come controllo che tutto proceda correttamente) conviene generare un *file* di dati trasformati (utilizzando un programma che lo consenta, come *scilab*) e quindi lavorare su di essi.

TAVOLE NUMERICHE

A.1 DEFINIZIONE DELLA FUNZIONE $\text{erf}(x)$

La funzione erf (nota anche come *error function*) è solitamente definita¹ come:

$$\text{erf}(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\xi^2/2} d\xi \quad (166)$$

ed è essenzialmente l'integrale, eseguito tra 0 ed un estremo variabile x , di una distribuzione Gaussiana di media 0 e deviazione standard 1:

$$\text{erf}(x) = \int_0^x G(z) dz \quad (167)$$

► ESEMPIO A.1. $\text{erf}(1) = 0.3413$. Questo significa che la probabilità che il valore di una variabile Gaussiana standard sia compreso tra 0 ed 1 vale 0.3413.

Il risultato può essere facilmente generalizzato ad una qualsiasi distribuzione Gaussiana: la probabilità che la variabile sia compresa tra μ (0 per la variabile standard) e $\mu + \sigma$ ($0 + 1 = 1$ per la variabile standard) vale 0.3413. \square

► ESEMPIO A.2. Se moltiplichiamo per due il valore appena trovato ($0.3413 \cdot 2 = 0.6826$) riotteniamo il ben noto risultato che la probabilità che una variabile gaussiana disti meno di una deviazione standard dal suo valor medio è circa il 68%. \square

Nel capitolo 3 abbiamo già anticipato che questo integrale (che rappresenta la probabilità di trovare una variabile normale in forma standard entro l'intervallo $[0, x]$) non ha espressione analitica e deve essere valutato numericamente. La funzione erf si trova

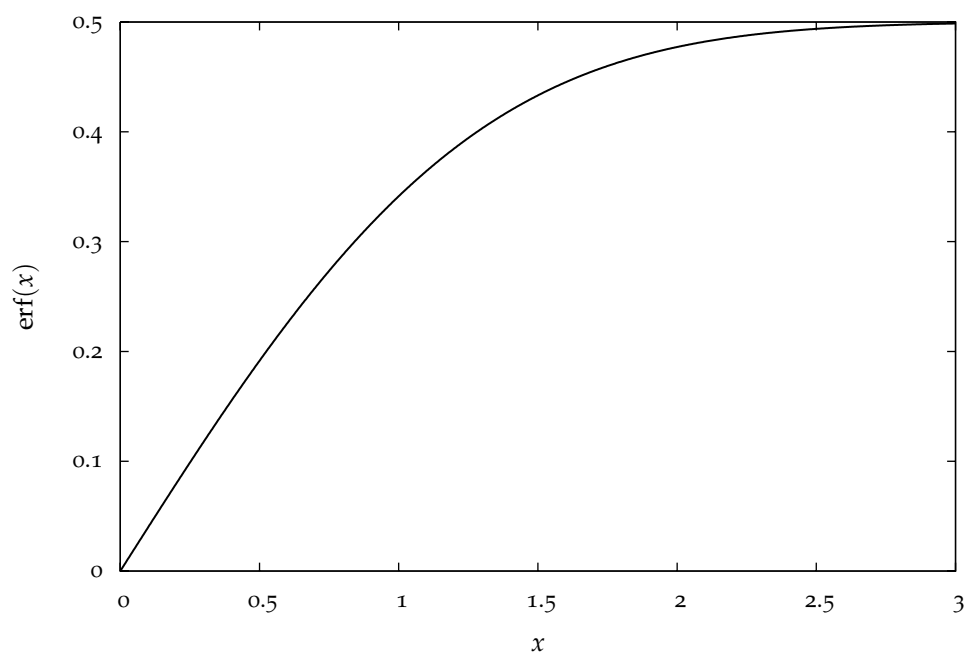
¹ Purtroppo alcuni autori preferiscono definire la funzione erf come:

$$\text{Erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-\xi^2} d\xi \quad (164)$$

che è legata alla 166 da un semplice cambio di variabile:

$$\text{erf}(x) = \frac{1}{2} \text{Erf}\left(\frac{x}{\sqrt{2}}\right) \quad (165)$$

Prima di utilizzare delle tavole numeriche è dunque essenziale verificarne il significato esatto.

FIGURA 38.. Grafico della funzione $\text{erf}(x)$ definita dall'equazione 166.

tabulata qui di seguito per un insieme significativo di valori di x . Il procedimento base per estrarre dalle tavole i valori di probabilità è stato già descritto nel capitolo 3 e è sostanzialmente fondato sulla relazione, valida per una variabile Gaussiana standard²:

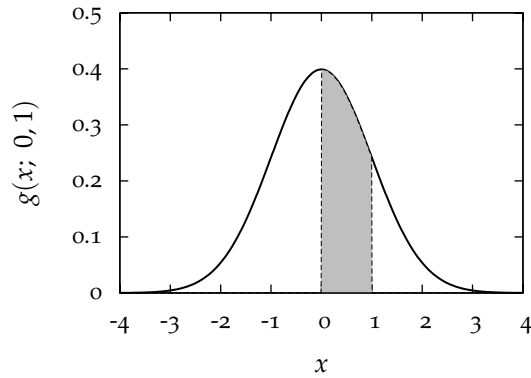
$$P(a \leq x \leq b) = \text{erf}(b) - \text{erf}(a) \quad (168)$$

Per una distribuzione Gaussiana con media e deviazione standard arbitrarie la (168) può ancora essere utilizzata passando alla variabile in forma standard (cfr. capitolo 3).

² Alcuni dei programmi di analisi dati comunemente usati (ad esempio scilab e gnuplot) utilizzano la definizione (164). In questo caso si ha, ovviamente:

$$P(a \leq x \leq b) = \frac{1}{2} \left[\text{Erf} \left(\frac{b}{\sqrt{2}} \right) - \text{Erf} \left(\frac{a}{\sqrt{2}} \right) \right]$$

A.2 INTEGRALE NORMALE DEGLI ERRORI - I



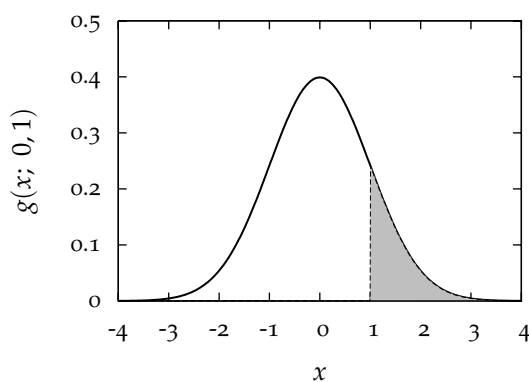
Tavole della funzione $f(x) = \text{erf}(x)$.

Nota: le righe identificano le prime due cifre della x , mentre le colonne identificano la seconda cifra decimale della x stessa.

ESEMPIO: la probabilità che una variabile gaussiana standard sia compresa tra 0.0 e 1.52 è 0.43574 (cfr. sedicesima riga, terza colonna).

x	0	1	2	3	4	5	6	7	8	9
0.0	0.00000	0.00399	0.00798	0.01197	0.01595	0.01994	0.02392	0.02790	0.03188	0.03586
0.1	0.03983	0.04380	0.04776	0.05172	0.05567	0.05962	0.06356	0.06749	0.07142	0.07535
0.2	0.07926	0.08317	0.08706	0.09095	0.09483	0.09871	0.10257	0.10642	0.11026	0.11409
0.3	0.11791	0.12172	0.12552	0.12930	0.13307	0.13683	0.14058	0.14431	0.14803	0.15173
0.4	0.15542	0.15910	0.16276	0.16640	0.17003	0.17364	0.17724	0.18082	0.18439	0.18793
0.5	0.19146	0.19497	0.19847	0.20194	0.20540	0.20884	0.21226	0.21566	0.21904	0.22240
0.6	0.22575	0.22907	0.23237	0.23565	0.23891	0.24215	0.24537	0.24857	0.25175	0.25490
0.7	0.25804	0.26115	0.26424	0.26730	0.27035	0.27337	0.27637	0.27935	0.28230	0.28524
0.8	0.28814	0.29103	0.29389	0.29673	0.29955	0.30234	0.30511	0.30785	0.31057	0.31327
0.9	0.31594	0.31859	0.32121	0.32381	0.32639	0.32894	0.33147	0.33398	0.33646	0.33891
1.0	0.34134	0.34375	0.34614	0.34849	0.35083	0.35314	0.35543	0.35769	0.35993	0.36214
1.1	0.36433	0.36650	0.36864	0.37076	0.37286	0.37493	0.37698	0.37900	0.38100	0.38298
1.2	0.38493	0.38686	0.38877	0.39065	0.39251	0.39435	0.39617	0.39796	0.39973	0.40147
1.3	0.40320	0.40490	0.40658	0.40824	0.40988	0.41149	0.41309	0.41466	0.41621	0.41774
1.4	0.41924	0.42073	0.42220	0.42364	0.42507	0.42647	0.42785	0.42922	0.43056	0.43189
1.5	0.43319	0.43448	0.43574	0.43699	0.43822	0.43943	0.44062	0.44179	0.44295	0.44408
1.6	0.44520	0.44630	0.44738	0.44845	0.44950	0.45053	0.45154	0.45254	0.45352	0.45449
1.7	0.45543	0.45637	0.45728	0.45818	0.45907	0.45994	0.46080	0.46164	0.46246	0.46327
1.8	0.46407	0.46485	0.46562	0.46638	0.46712	0.46784	0.46856	0.46926	0.46995	0.47062
1.9	0.47128	0.47193	0.47257	0.47320	0.47381	0.47441	0.47500	0.47558	0.47615	0.47670
2.0	0.47725	0.47778	0.47831	0.47882	0.47932	0.47982	0.48030	0.48077	0.48124	0.48169
2.1	0.48214	0.48257	0.48300	0.48341	0.48382	0.48422	0.48461	0.48500	0.48537	0.48574
2.2	0.48610	0.48645	0.48679	0.48713	0.48745	0.48778	0.48809	0.48840	0.48870	0.48899
2.3	0.48928	0.48956	0.48983	0.49010	0.49036	0.49061	0.49086	0.49111	0.49134	0.49158
2.4	0.49180	0.49202	0.49224	0.49245	0.49266	0.49286	0.49305	0.49324	0.49343	0.49361
2.5	0.49379	0.49396	0.49413	0.49430	0.49446	0.49461	0.49477	0.49492	0.49506	0.49520
2.6	0.49534	0.49547	0.49560	0.49573	0.49585	0.49598	0.49609	0.49621	0.49632	0.49643
2.7	0.49653	0.49664	0.49674	0.49683	0.49693	0.49702	0.49711	0.49720	0.49728	0.49736
2.8	0.49744	0.49752	0.49760	0.49767	0.49774	0.49781	0.49788	0.49795	0.49801	0.49807
2.9	0.49813	0.49819	0.49825	0.49831	0.49836	0.49841	0.49846	0.49851	0.49856	0.49861

A.3 INTEGRALE NORMALE DEGLI ERRORI - II

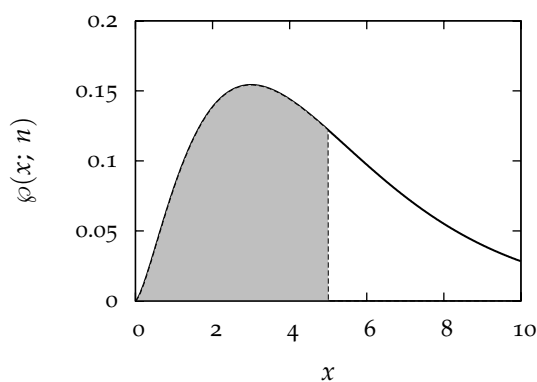


Tavole della funzione $f(x) = 0.5 - \text{erf}(x)$.

Nota: le righe identificano le prime due cifre della x , mentre le colonne identificano la seconda cifra decimale della x stessa.

ESEMPIO: la probabilità che una variabile gaussiana in forma standard sia maggiore di 3.32 è $4.50 \cdot 10^{-4}$ (cfr. quarta riga, terza colonna).

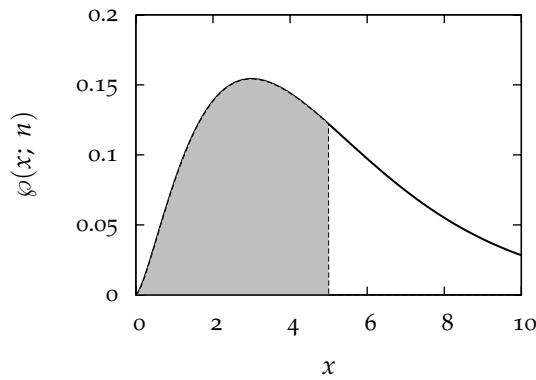
x	0	1	2	3	4	5	6	7	8	9
3.0	1.35 e-3	1.31 e-3	1.26 e-3	1.22 e-3	1.18 e-3	1.14 e-3	1.11 e-3	1.07 e-3	1.04 e-3	1.00 e-3
3.1	9.68 e-4	9.35 e-4	9.04 e-4	8.74 e-4	8.45 e-4	8.16 e-4	7.89 e-4	7.62 e-4	7.36 e-4	7.11 e-4
3.2	6.87 e-4	6.64 e-4	6.41 e-4	6.19 e-4	5.98 e-4	5.77 e-4	5.57 e-4	5.38 e-4	5.19 e-4	5.01 e-4
3.3	4.83 e-4	4.66 e-4	4.50 e-4	4.34 e-4	4.19 e-4	4.04 e-4	3.90 e-4	3.76 e-4	3.62 e-4	3.49 e-4
3.4	3.37 e-4	3.25 e-4	3.13 e-4	3.02 e-4	2.91 e-4	2.80 e-4	2.70 e-4	2.60 e-4	2.51 e-4	2.42 e-4
3.5	2.33 e-4	2.24 e-4	2.16 e-4	2.08 e-4	2.00 e-4	1.93 e-4	1.85 e-4	1.78 e-4	1.72 e-4	1.65 e-4
3.6	1.59 e-4	1.53 e-4	1.47 e-4	1.42 e-4	1.36 e-4	1.31 e-4	1.26 e-4	1.21 e-4	1.17 e-4	1.12 e-4
3.7	1.08 e-4	1.04 e-4	9.96 e-5	9.57 e-5	9.20 e-5	8.84 e-5	8.50 e-5	8.16 e-5	7.84 e-5	7.53 e-5
3.8	7.23 e-5	6.95 e-5	6.67 e-5	6.41 e-5	6.15 e-5	5.91 e-5	5.67 e-5	5.44 e-5	5.22 e-5	5.01 e-5
3.9	4.81 e-5	4.61 e-5	4.43 e-5	4.25 e-5	4.07 e-5	3.91 e-5	3.75 e-5	3.59 e-5	3.45 e-5	3.30 e-5
4.0	3.17 e-5	3.04 e-5	2.91 e-5	2.79 e-5	2.67 e-5	2.56 e-5	2.45 e-5	2.35 e-5	2.25 e-5	2.16 e-5
4.1	2.07 e-5	1.98 e-5	1.89 e-5	1.81 e-5	1.74 e-5	1.66 e-5	1.59 e-5	1.52 e-5	1.46 e-5	1.39 e-5
4.2	1.33 e-5	1.28 e-5	1.22 e-5	1.17 e-5	1.12 e-5	1.07 e-5	1.02 e-5	9.77 e-6	9.34 e-6	8.93 e-6
4.3	8.54 e-6	8.16 e-6	7.80 e-6	7.46 e-6	7.12 e-6	6.81 e-6	6.50 e-6	6.21 e-6	5.93 e-6	5.67 e-6
4.4	5.41 e-6	5.17 e-6	4.94 e-6	4.71 e-6	4.50 e-6	4.29 e-6	4.10 e-6	3.91 e-6	3.73 e-6	3.56 e-6
4.5	3.40 e-6	3.24 e-6	3.09 e-6	2.95 e-6	2.81 e-6	2.68 e-6	2.56 e-6	2.44 e-6	2.32 e-6	2.22 e-6
4.6	2.11 e-6	2.01 e-6	1.92 e-6	1.83 e-6	1.74 e-6	1.66 e-6	1.58 e-6	1.51 e-6	1.43 e-6	1.37 e-6
4.7	1.30 e-6	1.24 e-6	1.18 e-6	1.12 e-6	1.07 e-6	1.02 e-6	9.68 e-7	9.21 e-7	8.76 e-7	8.34 e-7
4.8	7.93 e-7	7.55 e-7	7.18 e-7	6.83 e-7	6.49 e-7	6.17 e-7	5.87 e-7	5.58 e-7	5.30 e-7	5.04 e-7
4.9	4.79 e-7	4.55 e-7	4.33 e-7	4.11 e-7	3.91 e-7	3.71 e-7	3.52 e-7	3.35 e-7	3.18 e-7	3.02 e-7
5.0	2.87 e-7	2.72 e-7	2.58 e-7	2.45 e-7	2.33 e-7	2.21 e-7	2.10 e-7	1.99 e-7	1.89 e-7	1.79 e-7
5.1	1.70 e-7	1.61 e-7	1.53 e-7	1.45 e-7	1.37 e-7	1.30 e-7	1.23 e-7	1.17 e-7	1.11 e-7	1.05 e-7
5.2	9.96 e-8	9.44 e-8	8.95 e-8	8.48 e-8	8.03 e-8	7.60 e-8	7.20 e-8	6.82 e-8	6.46 e-8	6.12 e-8
5.3	5.79 e-8	5.48 e-8	5.19 e-8	4.91 e-8	4.65 e-8	4.40 e-8	4.16 e-8	3.94 e-8	3.72 e-8	3.52 e-8
5.4	3.33 e-8	3.15 e-8	2.98 e-8	2.82 e-8	2.66 e-8	2.52 e-8	2.38 e-8	2.25 e-8	2.13 e-8	2.01 e-8
5.5	1.90 e-8	1.79 e-8	1.69 e-8	1.60 e-8	1.51 e-8	1.43 e-8	1.35 e-8	1.27 e-8	1.20 e-8	1.14 e-8
5.6	1.07 e-8	1.01 e-8	9.55 e-9	9.01 e-9	8.50 e-9	8.02 e-9	7.57 e-9	7.14 e-9	6.73 e-9	6.35 e-9
5.7	5.99 e-9	5.65 e-9	5.33 e-9	5.02 e-9	4.73 e-9	4.46 e-9	4.21 e-9	3.96 e-9	3.74 e-9	3.52 e-9
5.8	3.32 e-9	3.12 e-9	2.94 e-9	2.77 e-9	2.61 e-9	2.46 e-9	2.31 e-9	2.18 e-9	2.05 e-9	1.93 e-9
5.9	1.82 e-9	1.71 e-9	1.61 e-9	1.51 e-9	1.43 e-9	1.34 e-9	1.26 e-9	1.19 e-9	1.12 e-9	1.05 e-9

A.4 DISTRIBUZIONE DEL χ^2 - I

Valori del χ_p^2 a ν gradi di libertà per cui l'area ombreggiata è p .

ESEMPIO: nel caso di $\nu = 10$ gradi di libertà, la probabilità che il χ^2 sia minore di 9.34 è il 50% (cfr. decima riga, settima colonna).

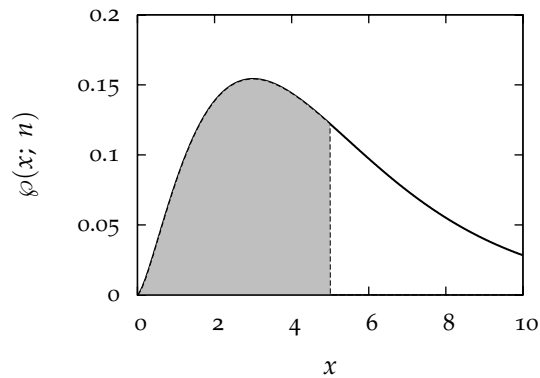
ν	$\chi_{0.005}^2$	$\chi_{0.01}^2$	$\chi_{0.025}^2$	$\chi_{0.05}^2$	$\chi_{0.10}^2$	$\chi_{0.25}^2$	$\chi_{0.50}^2$	$\chi_{0.75}^2$	$\chi_{0.90}^2$	$\chi_{0.95}^2$	$\chi_{0.975}^2$	$\chi_{0.99}^2$	$\chi_{0.999}^2$
1	3.9 e-5	1.6 e-4	9.8 e-4	3.9 e-3	1.6 e-2	0.10	0.45	1.32	2.71	3.84	5.02	6.63	10.8
2	1.0 e-2	2.0 e-2	5.1 e-2	0.10	0.21	0.58	1.39	2.77	4.61	5.99	7.38	9.21	13.8
3	7.2 e-2	0.11	0.22	0.35	0.58	1.21	2.37	4.11	6.25	7.81	9.35	11.3	16.3
4	0.21	0.30	0.48	0.71	1.06	1.92	3.36	5.39	7.78	9.49	11.1	13.3	18.5
5	0.41	0.55	0.83	1.15	1.61	2.67	4.35	6.63	9.24	11.1	12.8	15.1	20.5
6	0.68	0.87	1.24	1.64	2.20	3.45	5.35	7.84	10.6	12.6	14.4	16.8	22.5
7	0.99	1.24	1.69	2.17	2.83	4.25	6.35	9.04	12.0	14.1	16.0	18.5	24.3
8	1.34	1.65	2.18	2.73	3.49	5.07	7.34	10.2	13.4	15.5	17.5	20.1	26.1
9	1.73	2.09	2.70	3.33	4.17	5.90	8.34	11.4	14.7	16.9	19.0	21.7	27.9
10	2.16	2.56	3.25	3.94	4.87	6.74	9.34	12.5	16.0	18.3	20.5	23.2	29.6
11	2.60	3.05	3.82	4.57	5.58	7.58	10.3	13.7	17.3	19.7	21.9	24.7	31.3
12	3.07	3.57	4.40	5.23	6.30	8.44	11.3	14.8	18.5	21.0	23.3	26.2	32.9
13	3.57	4.11	5.01	5.89	7.04	9.30	12.3	16.0	19.8	22.4	24.7	27.7	34.5
14	4.07	4.66	5.63	6.57	7.79	10.2	13.3	17.1	21.1	23.7	26.1	29.1	36.1
15	4.60	5.23	6.26	7.26	8.55	11.0	14.3	18.2	22.3	25.0	27.5	30.6	37.7
16	5.14	5.81	6.91	7.96	9.31	11.9	15.3	19.4	23.5	26.3	28.8	32.0	39.3
17	5.70	6.41	7.56	8.67	10.1	12.8	16.3	20.5	24.8	27.6	30.2	33.4	40.8
18	6.26	7.01	8.23	9.39	10.9	13.7	17.3	21.6	26.0	28.9	31.5	34.8	42.3
19	6.84	7.63	8.91	10.1	11.7	14.6	18.3	22.7	27.2	30.1	32.9	36.2	43.8
20	7.43	8.26	9.59	10.9	12.4	15.5	19.3	23.8	28.4	31.4	34.2	37.6	45.3
21	8.03	8.90	10.3	11.6	13.2	16.3	20.3	24.9	29.6	32.7	35.5	38.9	46.8
22	8.64	9.54	11.0	12.3	14.0	17.2	21.3	26.0	30.8	33.9	36.8	40.3	48.3
23	9.26	10.2	11.7	13.1	14.8	18.1	22.3	27.1	32.0	35.2	38.1	41.6	49.7
24	9.89	10.9	12.4	13.8	15.7	19.0	23.3	28.2	33.2	36.4	39.4	43.0	51.2
25	10.5	11.5	13.1	14.6	16.5	19.9	24.3	29.3	34.4	37.7	40.6	44.3	52.6
26	11.2	12.2	13.8	15.4	17.3	20.8	25.3	30.4	35.6	38.9	41.9	45.6	54.1
27	11.8	12.9	14.6	16.2	18.1	21.7	26.3	31.5	36.7	40.1	43.2	47.0	55.5
28	12.5	13.6	15.3	16.9	18.9	22.7	27.3	32.6	37.9	41.3	44.5	48.3	56.9
29	13.1	14.3	16.0	17.7	19.8	23.6	28.3	33.7	39.1	42.6	45.7	49.6	58.3
30	13.8	15.0	16.8	18.5	20.6	24.5	29.3	34.8	40.3	43.8	47.0	50.9	59.7

A.5 DISTRIBUZIONE DEL χ^2 - II

Valori del χ_p^2 a ν gradi di libertà per cui l'area ombreggiata è p .

ESEMPIO: nel caso $\nu = 31$ gradi di libertà, la probabilità che il χ^2 sia minore di 21.4 è il 10% (cfr. prima riga, quinta colonna).

ν	$\chi_{0.005}^2$	$\chi_{0.01}^2$	$\chi_{0.025}^2$	$\chi_{0.05}^2$	$\chi_{0.10}^2$	$\chi_{0.25}^2$	$\chi_{0.50}^2$	$\chi_{0.75}^2$	$\chi_{0.90}^2$	$\chi_{0.95}^2$	$\chi_{0.975}^2$	$\chi_{0.99}^2$	$\chi_{0.999}^2$
31	14.5	15.7	17.5	19.3	21.4	25.4	30.3	35.9	41.4	45.0	48.2	52.2	61.1
32	15.1	16.4	18.3	20.1	22.3	26.3	31.3	37.0	42.6	46.2	49.5	53.5	62.5
33	15.8	17.1	19.0	20.9	23.1	27.2	32.3	38.1	43.7	47.4	50.7	54.8	63.9
34	16.5	17.8	19.8	21.7	24.0	28.1	33.3	39.1	44.9	48.6	52.0	56.1	65.2
35	17.2	18.5	20.6	22.5	24.8	29.1	34.3	40.2	46.1	49.8	53.2	57.3	66.6
36	17.9	19.2	21.3	23.3	25.6	30.0	35.3	41.3	47.2	51.0	54.4	58.6	68.0
37	18.6	20.0	22.1	24.1	26.5	30.9	36.3	42.4	48.4	52.2	55.7	59.9	69.3
38	19.3	20.7	22.9	24.9	27.3	31.8	37.3	43.5	49.5	53.4	56.9	61.2	70.7
39	20.0	21.4	23.7	25.7	28.2	32.7	38.3	44.5	50.7	54.6	58.1	62.4	72.1
40	20.7	22.2	24.4	26.5	29.1	33.7	39.3	45.6	51.8	55.8	59.3	63.7	73.4
41	21.4	22.9	25.2	27.3	29.9	34.6	40.3	46.7	52.9	56.9	60.6	65.0	74.7
42	22.1	23.7	26.0	28.1	30.8	35.5	41.3	47.8	54.1	58.1	61.8	66.2	76.1
43	22.9	24.4	26.8	29.0	31.6	36.4	42.3	48.8	55.2	59.3	63.0	67.5	77.4
44	23.6	25.1	27.6	29.8	32.5	37.4	43.3	49.9	56.4	60.5	64.2	68.7	78.7
45	24.3	25.9	28.4	30.6	33.4	38.3	44.3	51.0	57.5	61.7	65.4	70.0	80.1
46	25.0	26.7	29.2	31.4	34.2	39.2	45.3	52.1	58.6	62.8	66.6	71.2	81.4
47	25.8	27.4	30.0	32.3	35.1	40.1	46.3	53.1	59.8	64.0	67.8	72.4	82.7
48	26.5	28.2	30.8	33.1	35.9	41.1	47.3	54.2	60.9	65.2	69.0	73.7	84.0
49	27.2	28.9	31.6	33.9	36.8	42.0	48.3	55.3	62.0	66.3	70.2	74.9	85.4
50	28.0	29.7	32.4	34.8	37.7	42.9	49.3	56.3	63.2	67.5	71.4	76.2	86.7
51	28.7	30.5	33.2	35.6	38.6	43.9	50.3	57.4	64.3	68.7	72.6	77.4	88.0
52	29.5	31.2	34.0	36.4	39.4	44.8	51.3	58.5	65.4	69.8	73.8	78.6	89.3
53	30.2	32.0	34.8	37.3	40.3	45.7	52.3	59.5	66.5	71.0	75.0	79.8	90.6
54	31.0	32.8	35.6	38.1	41.2	46.7	53.3	60.6	67.7	72.2	76.2	81.1	91.9
55	31.7	33.6	36.4	39.0	42.1	47.6	54.3	61.7	68.8	73.3	77.4	82.3	93.2
56	32.5	34.3	37.2	39.8	42.9	48.5	55.3	62.7	69.9	74.5	78.6	83.5	94.5
57	33.2	35.1	38.0	40.6	43.8	49.5	56.3	63.8	71.0	75.6	79.8	84.7	95.8
58	34.0	35.9	38.8	41.5	44.7	50.4	57.3	64.9	72.2	76.8	80.9	86.0	97.0
59	34.8	36.7	39.7	42.3	45.6	51.4	58.3	65.9	73.3	77.9	82.1	87.2	98.3
60	35.5	37.5	40.5	43.2	46.5	52.3	59.3	67.0	74.4	79.1	83.3	88.4	99.6

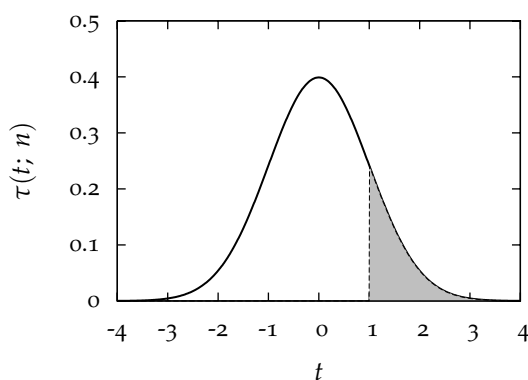
A.6 DISTRIBUZIONE DEL χ^2 - III

Valori del χ_p^2 a ν gradi di libertà per cui l'area ombreggiata è p .

ESEMPIO: nel caso di $\nu = 61$ gradi di libertà, la probabilità che il χ^2 sia minore di 60.3 è il 50% (cfr. prima riga, settima colonna).

ν	$\chi_{0.005}^2$	$\chi_{0.01}^2$	$\chi_{0.025}^2$	$\chi_{0.05}^2$	$\chi_{0.10}^2$	$\chi_{0.25}^2$	$\chi_{0.50}^2$	$\chi_{0.75}^2$	$\chi_{0.90}^2$	$\chi_{0.95}^2$	$\chi_{0.975}^2$	$\chi_{0.99}^2$	$\chi_{0.999}^2$
61	36.3	38.3	41.3	44.0	47.3	53.2	60.3	68.0	75.5	80.2	84.5	89.6	101
62	37.1	39.1	42.1	44.9	48.2	54.2	61.3	69.1	76.6	81.4	85.7	90.8	102
63	37.8	39.9	43.0	45.7	49.1	55.1	62.3	70.2	77.7	82.5	86.8	92.0	103
64	38.6	40.6	43.8	46.6	50.0	56.1	63.3	71.2	78.9	83.7	88.0	93.2	105
65	39.4	41.4	44.6	47.4	50.9	57.0	64.3	72.3	80.0	84.8	89.2	94.4	106
66	40.2	42.2	45.4	48.3	51.8	57.9	65.3	73.3	81.1	86.0	90.3	95.6	107
67	40.9	43.0	46.3	49.2	52.7	58.9	66.3	74.4	82.2	87.1	91.5	96.8	109
68	41.7	43.8	47.1	50.0	53.5	59.8	67.3	75.5	83.3	88.3	92.7	98.0	110
69	42.5	44.6	47.9	50.9	54.4	60.8	68.3	76.5	84.4	89.4	93.9	99.2	111
70	43.3	45.4	48.8	51.7	55.3	61.7	69.3	77.6	85.5	90.5	95.0	100	112
71	44.1	46.2	49.6	52.6	56.2	62.6	70.3	78.6	86.6	91.7	96.2	102	114
72	44.8	47.1	50.4	53.5	57.1	63.6	71.3	79.7	87.7	92.8	97.4	103	115
73	45.6	47.9	51.3	54.3	58.0	64.5	72.3	80.7	88.8	93.9	98.5	104	116
74	46.4	48.7	52.1	55.2	58.9	65.5	73.3	81.8	90.0	95.1	99.7	105	117
75	47.2	49.5	52.9	56.1	59.8	66.4	74.3	82.9	91.1	96.2	101	106	119
76	48.0	50.3	53.8	56.9	60.7	67.4	75.3	83.9	92.2	97.4	102	108	120
77	48.8	51.1	54.6	57.8	61.6	68.3	76.3	85.0	93.3	98.5	103	109	121
78	49.6	51.9	55.5	58.7	62.5	69.3	77.3	86.0	94.4	99.6	104	110	122
79	50.4	52.7	56.3	59.5	63.4	70.2	78.3	87.1	95.5	101	105	111	124
80	51.2	53.5	57.2	60.4	64.3	71.1	79.3	88.1	96.6	102	107	112	125
81	52.0	54.4	58.0	61.3	65.2	72.1	80.3	89.2	97.7	103	108	114	126
82	52.8	55.2	58.8	62.1	66.1	73.0	81.3	90.2	98.8	104	109	115	127
83	53.6	56.0	59.7	63.0	67.0	74.0	82.3	91.3	99.9	105	110	116	129
84	54.4	56.8	60.5	63.9	67.9	74.9	83.3	92.3	101	106	111	117	130
85	55.2	57.6	61.4	64.7	68.8	75.9	84.3	93.4	102	108	112	118	131
86	56.0	58.5	62.2	65.6	69.7	76.8	85.3	94.4	103	109	114	119	132
87	56.8	59.3	63.1	66.5	70.6	77.8	86.3	95.5	104	110	115	121	134
88	57.6	60.1	63.9	67.4	71.5	78.7	87.3	96.5	105	111	116	122	135
89	58.4	60.9	64.8	68.2	72.4	79.7	88.3	97.6	106	112	117	123	136
90	59.2	61.8	65.6	69.1	73.3	80.6	89.3	98.6	108	113	118	124	137

A.7 DISTRIBUZIONE T DI STUDENT - I



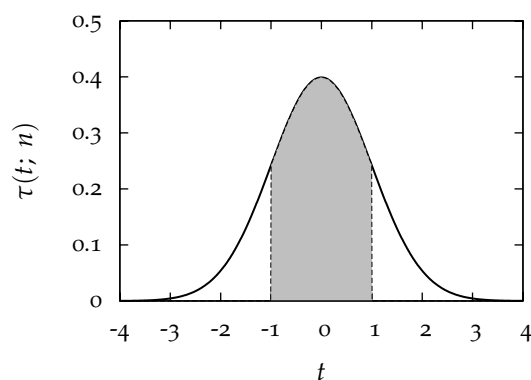
Valori di t_p a ν gradi di libertà per cui l'area ombreggiata è p .

Nota: il caso $\nu = \infty$ corrisponde ad una distribuzione di Gauss.

ESEMPIO: nel caso $\nu = 10$ gradi di libertà, la probabilità che la variabile t sia maggiore di 2.764 è l'1% (cfr. nona riga, terza colonna).

ν	$t_{0.001}$	$t_{0.005}$	$t_{0.010}$	$t_{0.025}$	$t_{0.050}$	$t_{0.0159}$	$t_{0.200}$	$t_{0.300}$	$t_{0.400}$	$t_{0.425}$
2	22.33	9.925	6.965	4.303	2.920	5.473	1.061	0.617	0.289	0.215
3	10.21	5.841	4.541	3.182	2.353	3.809	0.978	0.584	0.277	0.206
4	7.173	4.604	3.747	2.776	2.132	3.236	0.941	0.569	0.271	0.202
5	5.893	4.032	3.365	2.571	2.015	2.952	0.920	0.559	0.267	0.199
6	5.208	3.707	3.143	2.447	1.943	2.785	0.906	0.553	0.265	0.197
7	4.785	3.499	2.998	2.365	1.895	2.674	0.896	0.549	0.263	0.196
8	4.501	3.355	2.896	2.306	1.860	2.596	0.889	0.546	0.262	0.195
9	4.297	3.250	2.821	2.262	1.833	2.538	0.883	0.543	0.261	0.195
10	4.144	3.169	2.764	2.228	1.812	2.493	0.879	0.542	0.260	0.194
11	4.025	3.106	2.718	2.201	1.796	2.458	0.876	0.540	0.260	0.194
12	3.930	3.055	2.681	2.179	1.782	2.429	0.873	0.539	0.259	0.193
13	3.852	3.012	2.650	2.160	1.771	2.405	0.870	0.538	0.259	0.193
14	3.787	2.977	2.624	2.145	1.761	2.384	0.868	0.537	0.258	0.193
15	3.733	2.947	2.602	2.131	1.753	2.367	0.866	0.536	0.258	0.192
16	3.686	2.921	2.583	2.120	1.746	2.352	0.865	0.535	0.258	0.192
17	3.646	2.898	2.567	2.110	1.740	2.339	0.863	0.534	0.257	0.192
18	3.610	2.878	2.552	2.101	1.734	2.328	0.862	0.534	0.257	0.192
19	3.579	2.861	2.539	2.093	1.729	2.317	0.861	0.533	0.257	0.192
20	3.552	2.845	2.528	2.086	1.725	2.308	0.860	0.533	0.257	0.192
25	3.450	2.787	2.485	2.060	1.708	2.274	0.856	0.531	0.256	0.191
30	3.385	2.750	2.457	2.042	1.697	2.252	0.854	0.530	0.256	0.191
40	3.307	2.704	2.423	2.021	1.684	2.225	0.851	0.529	0.255	0.190
50	3.261	2.678	2.403	2.009	1.676	2.209	0.849	0.528	0.255	0.190
100	3.174	2.626	2.364	1.984	1.660	2.177	0.845	0.526	0.254	0.190
200	3.131	2.601	2.345	1.972	1.653	2.162	0.843	0.525	0.254	0.189
300	3.118	2.592	2.339	1.968	1.650	2.157	0.843	0.525	0.254	0.189
400	3.111	2.588	2.336	1.966	1.649	2.154	0.843	0.525	0.254	0.189
500	3.107	2.586	2.334	1.965	1.648	2.153	0.842	0.525	0.253	0.189
∞	3.090	2.576	2.326	1.960	1.645	2.147	0.842	0.524	0.253	0.189

A.8 DISTRIBUZIONE T DI STUDENT - II



Valori di t_p a ν gradi di libertà per cui l'area ombreggiata è p .

Nota: il caso $\nu = \infty$ corrisponde ad una distribuzione di Gauss.

ESEMPIO: nel caso $\nu = 10$ gradi di libertà la probabilità che t sia compreso tra -3.169 e 3.169 è il 99% (cfr. nona riga, ottava colonna).

ν	$t_{0.05}$	$t_{0.10}$	$t_{0.25}$	$t_{0.50}$	$t_{0.75}$	$t_{0.6827}$	$t_{0.95}$	$t_{0.99}$	$t_{0.999}$	$t_{0.9999}$
2	0.071	0.142	0.365	0.816	1.604	1.321	4.303	9.925	31.60	99.99
3	0.068	0.137	0.349	0.765	1.423	1.197	3.182	5.841	12.92	28.00
4	0.067	0.134	0.341	0.741	1.344	1.142	2.776	4.604	8.610	15.54
5	0.066	0.132	0.337	0.727	1.301	1.111	2.571	4.032	6.869	11.18
6	0.065	0.131	0.334	0.718	1.273	1.091	2.447	3.707	5.959	9.082
7	0.065	0.130	0.331	0.711	1.254	1.077	2.365	3.499	5.408	7.885
8	0.065	0.130	0.330	0.706	1.240	1.067	2.306	3.355	5.041	7.120
9	0.064	0.129	0.329	0.703	1.230	1.059	2.262	3.250	4.781	6.594
10	0.064	0.129	0.328	0.700	1.221	1.053	2.228	3.169	4.587	6.211
11	0.064	0.129	0.327	0.697	1.214	1.048	2.201	3.106	4.437	5.921
12	0.064	0.128	0.326	0.695	1.209	1.043	2.179	3.055	4.318	5.694
13	0.064	0.128	0.325	0.694	1.204	1.040	2.160	3.012	4.221	5.513
14	0.064	0.128	0.325	0.692	1.200	1.037	2.145	2.977	4.140	5.363
15	0.064	0.128	0.325	0.691	1.197	1.034	2.131	2.947	4.073	5.239
16	0.064	0.128	0.324	0.690	1.194	1.032	2.120	2.921	4.015	5.134
17	0.064	0.128	0.324	0.689	1.191	1.030	2.110	2.898	3.965	5.044
18	0.064	0.127	0.324	0.688	1.189	1.029	2.101	2.878	3.922	4.966
19	0.064	0.127	0.323	0.688	1.187	1.027	2.093	2.861	3.883	4.897
20	0.063	0.127	0.323	0.687	1.185	1.026	2.086	2.845	3.850	4.837
25	0.063	0.127	0.322	0.684	1.178	1.020	2.060	2.787	3.725	4.619
30	0.063	0.127	0.322	0.683	1.173	1.017	2.042	2.750	3.646	4.482
40	0.063	0.126	0.321	0.681	1.167	1.013	2.021	2.704	3.551	4.321
50	0.063	0.126	0.320	0.679	1.164	1.010	2.009	2.678	3.496	4.228
100	0.063	0.126	0.320	0.677	1.157	1.005	1.984	2.626	3.390	4.053
200	0.063	0.126	0.319	0.676	1.154	1.003	1.972	2.601	3.340	3.970
300	0.063	0.126	0.319	0.675	1.153	1.002	1.968	2.592	3.323	3.944
400	0.063	0.126	0.319	0.675	1.152	1.001	1.966	2.588	3.315	3.930
500	0.063	0.126	0.319	0.675	1.152	1.001	1.965	2.586	3.310	3.922
∞	0.063	0.126	0.319	0.674	1.150	1.000	1.960	2.576	3.291	3.891

UN SEMPLICE DOCUMENTO L^AT_EX

Riportiamo di seguito il codice sorgente di un breve documento L^AT_EX, che può costituire un esempio utile per redigere la relazione di un'esperienza didattica. Il codice è volutamente non commentato, in modo da non risultare troppo appesantito; le informazioni contenute nel capitolo 9 dovrebbero essere più che sufficienti per ricostruire il significato dei singoli comandi. Il risultato della compilazione è riportato di seguito, in figura 39.

```
\documentclass[a4paper, 10pt]{article}
\usepackage[dvips]{graphicx}
\usepackage[italian]{babel}

\begin{document}
\title{\LaTeX: un \emph{template} per le relazioni}
\author{L. Baldini - INFN Pisa}
\maketitle

\section{La nostra prima sezione}
In questa sezione parliamo di equazioni. Ma prima notiamo che in
\LaTeX\ si pu\`o andare a capo con un \verb|\\|\\
oppure lasciando una linea vuota, nel qual caso la linea
successiva \`e indentata.

\subsection{Equazioni}
Ma veniamo dunque alle equazioni\ldots
\begin{equation}
\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}
\end{equation}

\section{Tabelle, elenchi e figure}

La tabella \ref{UnaTabella} \`e il nostro primo esempio di tabella,
mentre la figura \ref{UnaFigura} \`e, appunto, una figura.

\begin{table}[htb!]
\begin{center}
\begin{tabular}{cc}
```

```

\hline
Colonna 1 & Colonna 2\\
\hline
\hline
a & b \\
c & d \\
\hline
\end{tabular}
\caption{Un esempio di tabella\ldots}
\label{UnaTabella}
\end{center}
\end{table}

\noindent In \LaTeX\ 'e facile fare elenchi:
\begin{enumerate}
\item{Punto primo.}
\item{Punto secondo.}
\end{enumerate}

\begin{figure}[htb!]
\begin{center}
\includegraphics[height=3cm, angle=270]{figura.eps}
\caption{Questa 'e una figura. Piccola, lo ammetto, ma pur sempre
una figura.}
\label{UnaFigura}
\end{center}
\end{figure}

\end{document}

```

L^AT_EX: un *template* per le relazioni

L. Baldini - INFN Pisa

27 agosto 2008

1 La nostra prima sezione

In questa sezione parliamo di equazioni. Ma prima notiamo che in L^AT_EX si può andare a capo con un `\\` oppure lasciando una linea vuota, nel qual caso la linea successiva è indentata.

1.1 Equazioni

Ma veniamo dunque alle equazioni...

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi} \quad (1)$$

2 Tabelle, elenchi e figure

La tabella 1 è il nostro primo esempio di tabella, mentre la figura 1 è, appunto, una figura.

Colonna 1	Colonna 2
a	b
c	d

Tabella 1: Un esempio di tabella...

In L^AT_EX è facile fare elenchi:

1. Punto primo.
2. Punto secondo.

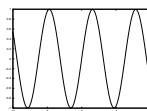


Figura 1: Questa è una figura. Piccola, lo ammetto, ma pur sempre una figura.

FIGURA 39.. Risultato della compilazione del documento L^AT_EX di esempio.

SIMBOLI MATEMATICI IN \LaTeX

Le tabelle che seguono illustrano i comandi per ottenere in \LaTeX alcuni tra simboli matematici più comunemente utilizzati. Ricordiamo che tutti questi comandi sono disponibili *solamente* in modalità matematica—se dovessero essere necessari all’interno del testo vanno racchiusi tra due “\$”.

Si rimanda il lettore alla (ponderosa) guida [18] per una lista esaustiva dei simboli disponibili in \LaTeX .

LETTERE GRECHE									
α	<code>\alpha</code>	β	<code>\beta</code>	γ	<code>\gamma</code>	δ	<code>\delta</code>	ϵ	<code>\epsilon</code>
ε	<code>\varepsilon</code>	ζ	<code>\zeta</code>	η	<code>\eta</code>	θ	<code>\theta</code>	ϑ	<code>\vartheta</code>
γ	<code>\gamma</code>	κ	<code>\kappa</code>	λ	<code>\lambda</code>	μ	<code>\mu</code>	ν	<code>\nu</code>
ξ	<code>\xi</code>	π	<code>\pi</code>	ϖ	<code>\varpi</code>	ρ	<code>\rho</code>	ϱ	<code>\varrho</code>
σ	<code>\sigma</code>	ς	<code>\varsigma</code>	τ	<code>\tau</code>	υ	<code>\upsilon</code>	ϕ	<code>\phi</code>
φ	<code>\varphi</code>	χ	<code>\chi</code>	ψ	<code>\psi</code>	ω	<code>\omega</code>		
Γ	<code>\Gamma</code>	Δ	<code>\Delta</code>	Θ	<code>\Theta</code>	Λ	<code>\Lambda</code>	Ξ	<code>\Xi</code>
Π	<code>\Pi</code>	Σ	<code>\Sigma</code>	Υ	<code>\Upsilon</code>	Φ	<code>\Phi</code>	Ψ	<code>\Psi</code>
Ω	<code>\Omega</code>								

TABELLA 3.. Tabella dei comandi \LaTeX per le lettere Greche in modalità matematica.

OPERATORI BINARI E DI RELAZIONE									
\pm	<code>\pm</code>	\mp	<code>\mp</code>	\times	<code>\times</code>	\div	<code>\div</code>	\circ	<code>\circ</code>
\cdot	<code>\cdot</code>	\cap	<code>\cap</code>	\cup	<code>\cup</code>	\oplus	<code>\oplus</code>	\ominus	<code>\ominus</code>
\otimes	<code>\otimes</code>	\oslash	<code>\oslash</code>	\odot	<code>\odot</code>				
\leq	<code>\leq</code>	\geq	<code>\geq</code>	\ll	<code>\ll</code>	\gg	<code>\gg</code>	\subset	<code>\subset</code>
\subseteq	<code>\subseteq</code>	\supset	<code>\supset</code>	\supseteq	<code>\supseteq</code>	\in	<code>\in</code>	\ni	<code>\ni</code>
\equiv	<code>\equiv</code>	\sim	<code>\sim</code>	\simeq	<code>\simeq</code>	\approx	<code>\approx</code>	\propto	<code>\propto</code>
\cong	<code>\cong</code>	\doteq	<code>\doteq</code>	\parallel	<code>\parallel</code>				

TABELLA 4.. Tabella degli operatori binari e degli operatori di relazione più comuni.

ALTRI SIMBOLI				
\cdots <code>\ldots</code>	\cdots <code>\cdots</code>	\vdots <code>\vdots</code>	\ddots <code>\ddots</code>	\forall <code>\forall</code>
∞ <code>\infty</code>	\hbar <code>\hbar</code>	\emptyset <code>\emptyset</code>	\exists <code>\exists</code>	∇ <code>\nabla</code>
∂ <code>\partial</code>	\sum <code>\sum</code>	\prod <code>\prod</code>	\int <code>\int</code>	\oint <code>\oint</code>

TABELLA 5.. Altri simboli disponibili in \LaTeX .

BREVE GLOSSARIO DI GNUPLOT

Questa appendice include alcuni dei comandi di gnuplot più comunemente usati, in combinazione con alcune tra le opzioni più utili. Può servire da riferimento veloce durante l'elaborazione di dati. Si rimanda comunque al capitolo 10 per informazioni più dettagliate sul significato dei singoli comandi.

```
gnuplot> set output
gnuplot> set terminal x11
```

Questi due comandi preparano gnuplot ad operare nella modalità comunemente impiegata, in cui i grafici vengono rediretti sullo schermo. È una buona abitudine includerli all'inizio ed alla fine di ogni macro. Ricordiamo anche la variante per Microsoft Windows "set terminal windows".

```
gnuplot> set terminal postscript
gnuplot> set output 'prova.eps'
```

Questi comandi reindirizzano l'uscita su di un file postscript denominato "prova.eps". È essenziale ricordarsi di tornare allo schermo con i due comandi precedenti, una volta scritto il file.

```
gnuplot> help
```

Visualizza le pagine di documentazione interattiva. Può accettare come argomento comandi o parole chiave di gnuplot su cui si desidera avere informazioni.

```
gnuplot> set title 'Titolo del grafico'
```

Imposta il titolo del grafico, che è accettato come parametro tra virgolette o tra apici.

```
gnuplot> set xlabel 'x (cm)'
```

Imposta l'etichetta di testo sull'asse delle x , passata come parametro tra virgolette o apici.

```
gnuplot> set ylabel 'y (cm)'
```

Imposta l'etichetta di testo sull'asse delle y , passata come parametro tra virgolette o apici.

```
gnuplot> set xrange [0.5:1.5]
```

Imposta gli estremi del grafico sull'asse delle x . Essi vengono passati tra parentesi quadre, separati da due punti.

```
gnuplot> set yrange [3:4]
```

Imposta gli estremi del grafico sull'asse delle y . Funziona esattamente come il comando precedente.

```
gnuplot> set autoscale x
```

Abilita l'impostazione automatica degli estremi dell'asse x (che, in questo modo, si adattano ai dati).

```
gnuplot> set autoscale y
```

Abilita l'impostazione automatica degli estremi dell'asse y (che, in questo modo, si adattano ai dati).

```
gnuplot> unset key
```

Elimina la legenda (se presente).

```
gnuplot> set key
```

Ripristina la legenda (se era stata eliminata). È il contrario del comando precedente.

```
gnuplot> set logscale x
```

Seleziona la scala logaritmica sull'asse delle x .

```
gnuplot> unset logscale x
```

Ripristina la scala lineare sull'asse delle x (se era selezionata quella logaritmica). È il contrario del comando precedente.

```
gnuplot> set logscale y
```

Seleziona la scala logaritmica sull'asse delle y .

```
gnuplot> unset logscale y
```

Ripristina la scala lineare sull'asse delle y (se era selezionata quella logaritmica). È il contrario del comando precedente.


```
gnuplot> lambda = 3.0
```

Crea una variabile chiamata “lambda” e le assegna il valore di 3.0. Le variabili possono essere identificate con una qualsiasi combinazione di lettere e numeri (purché cominci con una lettera). Questa sintassi è spesso usata quando si eseguono dei fit.

```
gnuplot> print lambda
```

Stampa sullo schermo il valore della variabile “lambda” (che deve essere stata precedentemente definita).

```
gnuplot> plot 'dati.txt' using 1:2
```

Crea un grafico a partire dai dati contenuti nel file “dati.txt”. In questo caso particolare la prima colonna è messa sull’asse delle x , la seconda su quello delle y .

```
gnuplot> plot 'dati.txt' using 1:2:3 with yerrorbars
```

Come il comando precedente, ma la terza colonna è questa volta interpretata come colonna degli errori sull’asse delle y , che sono correttamente riportati sul grafico.

```
gnuplot> plot 'dati.txt' using 1:2:3:4 with xyerrorbars
```

Come i due comandi precedenti, con le barre d’errore su entrambi gli assi.

```
gnuplot> plot 'dati.txt' using 1:2 with histeps
```

I dati contenuti nel file “dati.txt” vengono rappresentati sotto forma di istogramma.

```
gnuplot> plot 'dati.txt' using 1:($2)**2
```

Con questa sintassi sull’asse delle y vengono rappresentati i valori contenuti nella seconda colonna, elevati al quadrato.

```
gnuplot> replot
```

Esegue di nuovo l’ultimo comando “plot” che è stato impartito nella *shell* di gnuplot.

```
gnuplot> f(x) = a*sin(x)
```

Definisce una funzione della variabile indipendente x (che non necessita di essere definita a sua volta). La funzione può poi essere usata in un fit.

```
gnuplot> fit f(x) 'dati.txt' using 1:2 via a
```

Esegue un fit ai dati contenuti nel file "dati.txt" con la funzione (precedentemente definita) $f(x)$. Il parametro a è lasciato libero di variare (ricordiamo che la sua inizializzazione è essenziale per una corretta convergenza del fit).

```
gnuplot> fit f(x) 'dati.txt' using 1:2:3 via a
```

Come il comando precedente, ma in questo caso i punti (x_i, y_i) sono pesati con gli errori Δy_i , contenuti nella colonna numero 3 del file di dati.

UN ESEMPIO DI FIT CON SCILAB

gnuplot utilizza una variante del metodo di minimo χ^2 per eseguire i fit. Il procedimento in questione è di tipo iterativo: si parte da una stima iniziale dei parametri (fornita dall'utente) e si cambiano questi ultimi fino a che il χ^2 non converge ad un minimo. Questo è il modo in cui, in genere, funzionano i pacchetti di analisi dati per effettuare dei fit. In realtà, per lo meno nei casi più semplici, non è affatto complicato scrivere dei piccoli programmi che eseguano dei fit.

E.1 MINIMO χ^2 NEL CASO LINEARE

Supponiamo di avere una serie di dati sperimentali e di volerli fittare con la funzione più semplice in assoluto:

$$f(x) = ax \quad (169)$$

(abbiamo scelto una funzione puramente lineare, è bene notarlo, solamente per semplicità; in generale azzerare l'intercetta non è un'operazione lecita). Al solito, nell'ipotesi che gli errori sulla x siano trascurabili possiamo eseguire il fit minimizzando la quantità:

$$S = \sum_{i=0}^n \frac{(y_i - ax_i)^2}{(\Delta y_i)^2} \quad (170)$$

Nel nostro caso l'equazione da risolvere è:

$$\frac{\partial S}{\partial a} = -2 \sum_{i=0}^n \frac{(y_i - ax_i) \cdot x_i}{(\Delta y_i)^2} = 0 \quad (171)$$

che conduce all'equazione per a :

$$a = \frac{\sum_{i=0}^n \frac{y_i x_i}{(\Delta y_i)^2}}{\sum_{i=0}^n \frac{(x_i)^2}{(\Delta y_i)^2}} \quad (172)$$

Notiamo, per inciso, che introducendo i vettori n -dimensionali:

$$\begin{aligned} \mathbf{v} &= \left(\frac{x_1}{\Delta y_1}, \frac{x_2}{\Delta y_2}, \dots, \frac{x_n}{\Delta y_n} \right) \\ \mathbf{w} &= \left(\frac{y_1}{\Delta y_1}, \frac{y_2}{\Delta y_2}, \dots, \frac{y_n}{\Delta y_n} \right) \end{aligned} \quad (173)$$

l'equazione per a si può scrivere in forma compatta come:

$$a = \frac{\mathbf{v} \cdot \mathbf{w}}{\mathbf{v} \cdot \mathbf{v}} \quad (174)$$

in cui il puntino sta ad indicare l'usuale prodotto scalare tra vettori.

E.2 IMPLEMENTARE IL FIT CON SCILAB

A questo punto eseguire il fit è una cosa semplice: basta prendere i dati ed eseguire, nella giusta sequenza, i prodotti, le divisioni e le somme necessarie a calcolare a secondo l'equazione scritta sopra. Una calcolatrice va benissimo, ma il tutto si può fare in modo più efficiente (specialmente nel caso si abbia un gran numero di dati) scrivendo un piccolo programma. Qualsiasi linguaggio di programmazione può servire egregiamente allo scopo; noi proveremo a farlo con il pacchetto *software* scilab, che offre tutta una serie di funzioni utili per maneggiare matrici e vettori.


scilab si lancia dalla *shell* con il comando:

```
>scilab
```

La prima cosa che dobbiamo fare è importare entro scilab la nostra matrice di dati, che chiameremo "M". Supponendo che il nostro *file* si chiami "LeggeOraria.txt", useremo il comando di scilab "read" come segue:

```
scilab>[M] = read('LeggeOraria.txt', 6, 3)
[M] =
| 5   10.1   2.0 |
| 10  17.8   1.5 |
| 15  26.6   1.5 |
| 20  32.8   2.0 |
| 25  40.2   2.5 |
| 30  47.4   2.0 |
scilab>
```

Essenzialmente questo significa: leggiamo il *file* "LeggeOraria.txt" ed estraiamo una matrice di sei righe e tre colonne che chiamiamo "M" e che da questo momento è disponibile in memoria per effettuare le operazioni desiderate; ad ogni passo scilab dovrebbe stampare sullo schermo ciò che abbiamo appena fatto, che è utile per verificarne la correttezza.

 *scilab non ama i commenti (le righe precedute da "#", che invece non creavano problemi a gnuplot) per cui dobbiamo avere la precauzione di cancellarle dal file contenente i dati, prima di eseguire il comando.*

A questo punto dobbiamo estrarre dalla matrice i vettori che ci interessano, cioè quello delle x , quello delle y e quello degli errori sulla y . Il modo per farlo (con ovvio significato dei termini) è:

```

scilab>[x] = M(:, 1)
[x] =
| 5 |
| 10 |
| 15 |
| 20 |
| 25 |
| 30 |
scilab>[y] = M(:, 2)
| 10.1 |
| 17.8 |
| 26.6 |
| 32.8 |
| 40.2 |
| 47.4 |
scilab>[dy] = M(:,3)
| 2.0 |
| 1.5 |
| 1.5 |
| 2.0 |
| 2.5 |
| 2.0 |
scilab>

```

Cioè: definiamo il vettore x come la prima colonna della matrice “ M ”, il vettore y come la seconda colonna e dy come la terza colonna. Costruiamo quindi i vettori v e w definiti nel paragrafo precedente:

```

scilab>v = x./dy
scilab>w = y./dy
scilab>

```

(notare che l’operatore “./” è quello che consente, dentro scilab, di eseguire la divisione membro a membro di due vettori). Non ci resta che eseguire i prodotti scalari per calcolare a :

```

scilab>(v'*w)/(v'*v)

```

(l’apostrofo “'” indica il vettore trasposto, dopo di che il prodotto scalare si esegue con l’usuale simbolo di moltiplicazione); scilab dovrebbe visualizzare sullo schermo il valore del parametro cercato.

BIBLIOGRAFIA

- [1] Philip R. Bevington, D. Keith Robinson, *Data Reduction and Error Analysis for the Physical Sciences*, McGraw Hill, Third Edition, 2003.
- [2] Tarcisio Del Prete, *Methods of Statistical Data Analysis in High Energy Physics*, disponibile alla pagina web:
<http://www.pi.infn.it/atlas/documenti/note/statistica.ps.gz>.
- [3] John R. Taylor, *Introduzione all'analisi degli errori*, Zanichelli Editore, Seconda Edizione, 2000.
- [4] Maurizio Loreti, *Teoria degli errori e fondamenti di statistica (introduzione alla fisica sperimentale)*, disponibile alla pagina web:
<http://wwwcdf.pd.infn.it/labo/INDEX.html>.
- [5] Kenneth A. Brownlee, *Statistical theory and methodology in science and engineering*, Wiley publications in applied statistics, Second Edition, 1984.
- [6] Harald Cramer, *Mathematical Methods of Statistics*, Princeton University Press, Paperback Edition, 1999.
- [7] Niels Arley, K. Rander Buch, *Introduction to the theory of probability and statistics*, Wiley & Sons, 1950.
- [8] Granvil C. Kyker Jr., *Am. J. Phys.* 51 (9), 852 (1983).
- [9] Wayne E. McGovern, *Am. J. Phys.* 60 (10), 943 (1992).
- [10] M. S. Lafleur, P. F. Hinrichsen, P. C. Landry, R. B. Moore, *The Poisson Distribution*, *Physics Teacher* 10, 314 (1972).
- [11] John. M. Mc Cormick, Mario G. Salvadori, *Numerical Methods in Fortran*, Prentice-Hall, 1964.
- [12] Lars Wirzenius, *Guida dell'amministratore di sistema Linux*, disponibile alla pagina web:
<ftp://ftp.pluto.it/pub/pluto/ildp/guide/sag/sag.pdf>.
- [13] Larry Greenfield, *Linux guida dell'utente*, disponibile alla pagina web:
<ftp://ftp.pluto.it/pub/pluto/ildp/guide/Guida-Utente/Guida-Utente.pdf>.
- [14] Home page di L^AT_EX:
<http://www.latex-project.org>
- [15] Home page del Gruppo utilizzatori Italiani di T_EX:
<http://www.guit.sssup.it>

- [16] Marc Baudoin, *Impara L^AT_EX! (... e mettilo da parte)*, disponibile alla pagina web:
http://www.mat.uniroma1.it/centro-calcolo/manuali/impara_latex.pdf.
- [17] Tobias Oetiker, *The Not So Short Introduction to L^AT_EX₂ ϵ* , disponibile alla pagina web:
<http://www.ctan.org/tex-archive/info/lshort/english/lshort.pdf>.
- [18] Scott Pakin, *The Comprehensive L^AT_EX Symbol List*, disponibile alla pagina web:
www.ctan.org/tex-archive/info/symbols/comprehensive/symbols-a4.pdf.
- [19] Home page di gnuplot:
<http://www.gnuplot.info>
- [20] Home page di scilab:
<http://www.scilab.org>
- [21] Home page di sourceforge:
<http://sourceforge.net>
- [22] Home page di GNU (GNU's Not Unix):
<http://www.gnu.org>
- [23] Home page di python:
<http://www.python.org>



Tutti i link a pagine web contenuti tra le indicazioni bibliografiche erano correttamente funzionanti la mattina del 14 Agosto 2007.

COLOPHON

Questo lavoro è stato realizzato con $\text{\LaTeX 2}_{\epsilon}$ sulla distribuzione di Linux Fedora 11. Tutti i grafici sono stati realizzati con gnuplot versione 4.2 *patchlevel* 2. Le tavole numeriche sono state generate utilizzando pyROOT versione 5.18.00.

INDICE ANALITICO

- χ^2
 - distribuzione del, 88
 - test del, 130
- accuratezza, 5
- aspettazione
 - valore di, 46
- binomiale
 - distribuzione, 60, 84
- Cauchy
 - distribuzione di, 90
- cifre significative, 17
- confidenza
 - intervalli di, 114
- correlazione, 141
 - coefficiente di, 143
- covarianza, 103
- curtosi, 52
- deviazione standard
 - definizione, 47
- distribuzione
 - funzione di, 32
- errore
 - casuale, 6
 - di misura, 6
 - relativo, 16
 - sistematico, 6
- esponenziale
 - distribuzione, 75
- fit
 - problemi di, 117
- Gauss
 - distribuzione di, 80
- HWHM
 - definizione, 49
 - della distribuzione di Cauchy, 93
 - della distribuzione di Gauss, 81
- incertezza
 - di misura, 6
 - massima, 7
- istogramma, 37
- limite centrale
 - teorema del, 109
- media
 - definizione, 42
 - della distribuzione binomiale, 63
 - della distribuzione del χ^2 , 89
 - della distribuzione di Cauchy, 92
 - della distribuzione di Gauss, 82
 - della distribuzione di Poisson, 70
 - della distribuzione esponenziale, 77
 - della distribuzione uniforme, 73
 - di una funzione di variabile casuale, 99
- media campione, 107
 - varianza della, 108
- media pesata, 123
- mediana
 - definizione, 43
 - della distribuzione di Cauchy, 92
- minimi quadrati
 - metodo di fit dei, 118
- minimo χ^2
 - metodo di fit del, 123
- moda, 45
- momenti
 - di una distribuzione, 51
- normalizzazione
 - di una funzione di distribuzione, 42
- Poisson

- distribuzione di, 65, 86
- precisione, 7
- probabilità, 26, 27
 - addizione delle, 28
 - densità di, 34
 - leggi della, 28
 - moltiplicazione delle, 29
- ripetitività, 5
- riproducibilità, 6
- skewness, 51
- t
 - distribuzione, 146
 - test, 147
 - variabile di Student, 145
- Tschebyscheff
 - teorema di, 52
- uniforme
 - distribuzione, 72
- variabile casuale, 32
- varianza
 - calcolo della, 48
 - definizione, 47
 - della distribuzione binomiale, 63
 - della distribuzione del χ^2 , 89
 - della distribuzione di Cauchy, 93
 - della distribuzione di Gauss, 83
 - della distribuzione di Poisson, 70
 - della distribuzione esponenziale, 77
 - della distribuzione uniforme, 74
 - di una funzione di variabile casuale, 99
- varianza campione, 107