Infine, la funzione caratteristica della y vale, essendo

$$y = \sum_{k=1}^{N} \xi_k$$

e ricordando l'equazione (6.11), applicabile perché anche le ξ_k sono indipendenti tra loro, otteniamo

$$\begin{aligned}
\phi_y(t) &= \prod_{k=1}^{N} \phi_{\xi_k}(t) \\
&= \prod_{k=1}^{N} e^{\left(ita_k\mu_k - \frac{1}{2}t^2 a_k{}^2\sigma_k{}^2\right)} \\
&= e^{\left[it\left(\sum_k a_k\mu_k\right) - \frac{1}{2}t^2\left(\sum_k a_k{}^2\sigma_k{}^2\right)\right]} \\
&= e^{\left(it\mu - \frac{t^2\sigma^2}{2}\right)}
\end{aligned}$$

ove si è posto

$$\mu = \sum_{k=1}^{N} a_k\,\mu_k \qquad \text{e} \qquad \sigma^2 = \sum_{k=1}^{N} a_k{}^2\sigma_k{}^2 \ .$$

Questa è appunto la funzione caratteristica di una nuova distribuzione normale; e, in virtù di uno dei teoremi enunciati nel paragrafo 6.4, quanto dimostrato prova la tesi.

8.3 La distribuzione di Cauchy

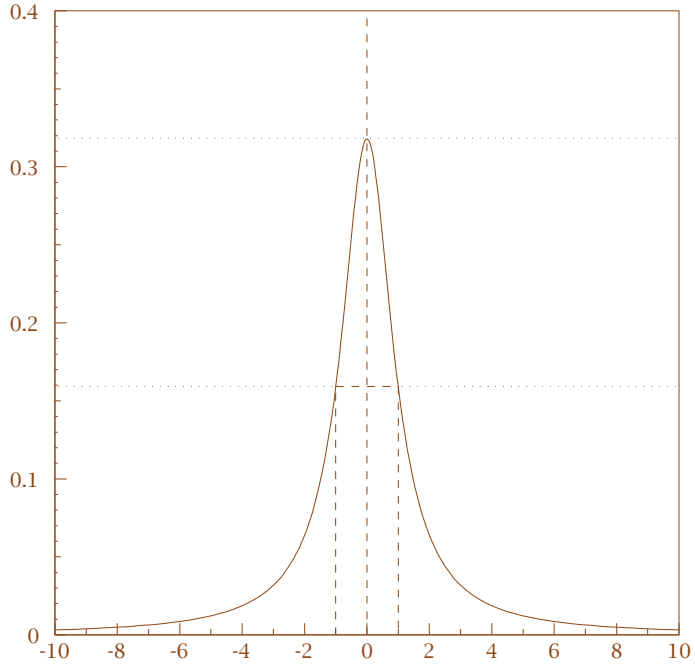
La distribuzione di Cauchy (o *distribuzione di Breit-Wigner*, nome con il quale è più nota nel mondo della fisica) è definita da una densità di probabilità che corrisponde alla funzione, dipendente da due parametri θ e d (con la condizione $d > 0$),

$$f(x;\theta,d) \ = \ \frac{1}{\pi d}\frac{1}{1+\left(\frac{x-\theta}{d}\right)^2} \ . \tag{8.6}$$

Anche se la (8.6) è integrabile, e la sua funzione integrale, ovverosia la funzione di distribuzione della x, vale

$$F(x;\theta,d) \ = \ \int_{-\infty}^{x} f(t)\,\mathrm{d}t \ = \ \frac{1}{2} + \frac{1}{\pi}\arctan\left(\frac{x-\theta}{d}\right)$$

FIGURA 8d - L'andamento della distribuzione di Cauchy, per $\theta = 0$ e $d = 1$.



nessuno dei momenti esiste, nemmeno la media.

θ è la mediana della distribuzione e d ne misura la larghezza a metà altezza, come è rilevabile ad esempio dalla figura 8d. La funzione caratteristica della distribuzione di Cauchy è la

$$\phi_x(t;\theta,d) = e^{i\theta t - |t|\,d} \quad ;$$

per la cosiddetta *variabile standardizzata*,

$$u = \frac{x-\theta}{d}$$

funzione di frequenza, funzione di distribuzione e funzione caratteristica valgono rispettivamente

$$\begin{cases} f(u) &= \dfrac{1}{\pi\,(1+u^2)} \\ F(u) &= \dfrac{1}{2} + \dfrac{1}{\pi}\arctan u \\ \phi_u(t) &= e^{-|t|} \end{cases}$$

Secondo la funzione (8.6) sono, ad esempio, distribuite le intensità nelle righe spettrali di emissione e di assorbimento degli atomi (che hanno una ampiezza non nulla); e la massa invariante delle risonanze nella fisica delle particelle elementari. È evidente però come nella fisica la distribuzione di Cauchy possa descrivere questi fenomeni solo in prima approssimazione: infatti essa si annulla solo per $x \to \pm\infty$, ed è chiaramente priva di significato fisico una probabilità non nulla di emissione spettrale per frequenze negative, o di masse invarianti anch'esse negative nel caso delle risonanze.

Per la distribuzione di Cauchy *troncata*, ossia quella descritta dalla funzione di frequenza (per la variabile standardizzata)

$$f(u|-K \le u \le K) = \begin{cases} 0 & |u| > K \\ \dfrac{1}{2\,\arctan K}\,\dfrac{1}{(1+u^2)} & |u| \le K \end{cases}$$

(*discontinua* in $u = \pm K$), esistono invece i momenti: i primi due valgono

$$E(u|-K \le u \le K) = 0$$

e

$$\mathrm{Var}(u|-K \le u \le K) = \frac{K}{\arctan K} - 1$$

Se le x_k sono N variabili casuali indipendenti che seguono la distribuzione di Cauchy con parametri θ_k e d_k, una generica loro combinazione lineare

$$y = \sum_{k=1}^{N} a_k\,x_k$$

segue la stessa distribuzione: infatti la funzione generatrice per le x_k è

$$\phi_{x_k}(t) = e^{i\theta_k t - |t| d_k}$$

e, definendo $\xi_k = a_k x_k$ e ricordando la (6.17),

$$\phi_{\xi_k}(t) \;=\; \phi_{x_k}(a_k t) \;=\; e^{i a_k \theta_k t - |t|\cdot|a_k|\,d_k} \quad ;$$

infine, applicando la (6.11),

$$\phi_y(t) \;=\; \prod_{k=1}^{N} \phi_{\xi_k}(t) \;=\; e^{i\theta_y t - |t| d_y}$$

ove si è posto

$$\theta_y = \sum_{k=1}^{N} a_k\theta_k \qquad \text{e} \qquad d_y = \sum_{k=1}^{N} |a_k| d_k \quad .$$

Una conseguenza importante è che il valore medio di un campione di misure proveniente da una popolazione che segua la distribuzione di Cauchy con certi parametri θ e d (in questo caso tutte le a_k sono uguali e valgono $1/N$) è distribuito anch'esso secondo Cauchy *e con gli stessi parametri*; in altre parole *non si guadagna nessuna informazione* accumulando più di una misura (e calcolando la media aritmetica del campione)[6].

8.3.1 Il rapporto di due variabili normali

Siano due variabili casuali x ed y che seguano la distribuzione normale standardizzata $N(0,1)$; e sia inoltre la y definita su tutto l'asse reale ad eccezione dell'origine ($y \neq 0$). La densità di probabilità congiunta di x e y è la

$$f(x,y) \;=\; N(x;0,1)\cdot N(y;0,1) \;=\; \frac{1}{2\pi}\,e^{-\frac{1}{2}x^2}e^{-\frac{1}{2}y^2} \quad ;$$

[6]Esistono altre tecniche, basate però sull'uso della mediana, che permettono di migliorare la conoscenza del valore di θ disponendo di più di una misura.

definendo

$$u = \frac{x}{y} \qquad \text{e} \qquad v = y$$

e ricordando la (7.4), la densità di probabilità $\varphi(u)$ della u è la

$$\begin{aligned}
\varphi(u) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}u^2v^2} e^{-\frac{1}{2}v^2} \, |v| \, \mathrm{d}v \\
&= \frac{1}{\pi} \int_{0}^{+\infty} e^{-\frac{1}{2}v^2(1+u^2)} \, v \, \mathrm{d}v \\
&= \frac{1}{\pi\,(1+u^2)} \int_{0}^{+\infty} e^{-t} \, \mathrm{d}t \\
&= \frac{1}{\pi\,(1+u^2)} \left[-e^{-t} \right]_0^{+\infty} \\
&= \frac{1}{\pi\,(1+u^2)}
\end{aligned}$$

Per eseguire l'integrazione si è effettuata la sostituzione

$$t = \frac{1}{2} v^2 \left(1 + u^2\right) \qquad \Longrightarrow \qquad \mathrm{d}t = \left(1 + u^2\right) v \, \mathrm{d}v$$

e si riconosce immediatamente nella $\varphi(u)$ la densità di probabilità di una variabile (standardizzata) di Cauchy: *il rapporto tra due variabili normali segue la distribuzione di Cauchy.*

8.4 La distribuzione di Bernoulli

Consideriamo un evento casuale ripetibile E, avente probabilità costante p di verificarsi; indichiamo con $q = 1-p$ la probabilità del non verificarsi di E (cioè la probabilità dell'evento complementare $\overline{E}$). Vogliamo ora determinare la probabilità $P(x;N)$ che in N prove ripetute E si verifichi esattamente x volte (deve necessariamente risultare $0 \le x \le N$).

L'evento casuale costituito dal presentarsi di E per x volte (e quindi dal presentarsi di $\overline{E}$ per le restanti $N - x$) è un evento complesso che può verificarsi in diverse maniere, corrispondenti a tutte le diverse possibili sequenze di successi e fallimenti; queste sono ovviamente mutuamente esclusive, ed in numero pari a quello delle possibili combinazioni di N oggetti a x a x, che vale

$$C_x^N = \binom{N}{x} = \frac{N!}{x!\,(N-x)!} \ .$$

Essendo poi ognuna delle prove statisticamente indipendente dalle altre (infatti la probabilità di E non cambia di prova in prova), ognuna delle possibili sequenze di x successi ed $N - x$ fallimenti ha una probabilità di presentarsi che vale $p^x q^{N-x}$; in definitiva

$$P(x;N) = \frac{N!}{x!\,(N-x)!} \, p^x q^{N-x} \ . \tag{8.7}$$

Questa distribuzione di probabilità $P(x;N)$ per una variabile casuale discreta x si chiama *distribuzione binomiale* o *di Bernoulli*[7]; vogliamo ora determinarne alcune costanti caratteristiche.

Verifichiamo per prima cosa che vale la condizione di normalizzazione: sfruttando la formula per lo sviluppo delle potenze del binomio, risulta

$$\sum_{x=0}^{N} P(x;N) = \sum_{x=0}^{N} \frac{N!}{x!\,(N-x)!} \, p^x q^{N-x} = (p+q)^N \equiv 1 \ .$$

Vogliamo ora calcolare la speranza matematica della variabile x (ossia il numero di successi attesi, in media, in N prove): per questo useremo la stessa variabile casuale ausiliaria già considerata nel paragrafo 5.6.3, y, che rappresenta il numero di successi nella generica delle N prove eseguite.

Avevamo a suo tempo già calcolato, sempre nel paragrafo 5.6.3, la speranza matematica della y

$$E(y) = 1 \cdot p + 0 \cdot q = p \ ;$$

e, osservando che anche y^2 può assumere i due soli valori 1 e 0, sempre con le probabilità rispettive p e q,

$$E(y^2) = 1 \cdot p + 0 \cdot q = p$$

e quindi la varianza della y esiste e vale

$$\mathrm{Var}(y) = E(y^2) - \left[E(y)\right]^2 = p - p^2 = p\,(1-p) = pq \ . \tag{8.8}$$

Il numero totale x di successi nelle N prove è legato ai valori y_i della y in ognuna di esse dalla

$$x = \sum_{i=1}^{N} y_i$$

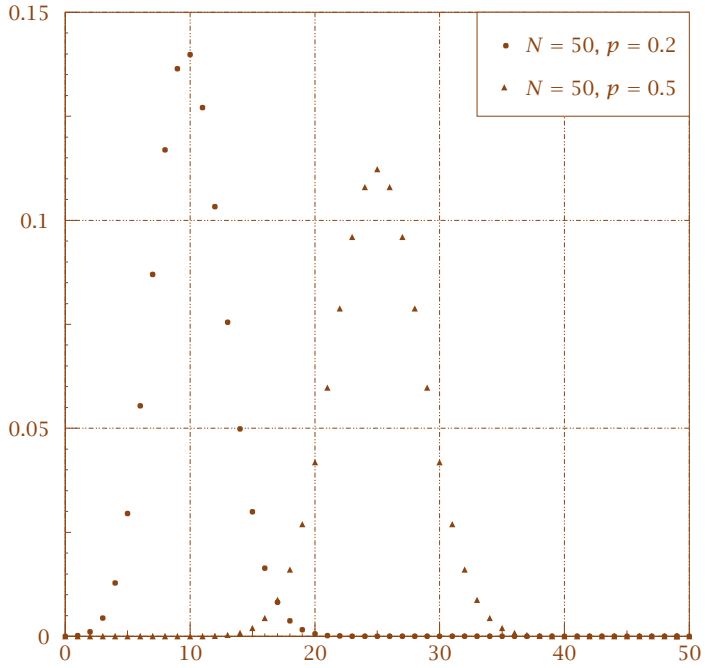
[7] I Bernoulli furono una famiglia originaria di Anversa poi trasferitasi a Basilea, numerosi membri della quale ebbero importanza per le scienze del diciassettesimo e del diciottesimo secolo; quello cui vanno attribuiti gli studi di statistica ebbe nome Jacob (o Jacques), visse dal 1654 al 1705, e fu zio del più noto Daniel (cui si deve il teorema di Bernoulli della dinamica dei fluidi).

e risulta quindi, per speranza matematica e varianza della distribuzione binomiale,

$$E(x) = E\left(\sum_{i=1}^{N} y_i\right) = \sum_{i=1}^{N} E(y_i) = Np$$

$$\mathrm{Var}(x) = \sigma_x{}^2 = \mathrm{Var}\left(\sum_{i=1}^{N} y_i\right) = \sum_{i=1}^{N} \mathrm{Var}(y_i) = Npq \ .$$

FIGURA 8e - La distribuzione binomiale, per un numero di prove $N = 50$ e due differenti valori della probabilità p.



Come è evidente dalla figura 8e, la forma della distribuzione binomiale è molto simile a quella di una curva di Gauss; si può in effetti dimostrare che

quando N tende all'infinito la distribuzione di probabilità dei possibili valori tende ad una distribuzione normale avente la stessa media Np e la stessa varianza Npq. Infatti la funzione generatrice dei momenti della distribuzione di Bernoulli è

$$\begin{aligned} M_x(t) &= E(e^{tx}) \\ &= \sum_{x=0}^{N} e^{tx} \binom{N}{x} p^x q^{N-x} \\ &= \sum_{x=0}^{N} \binom{N}{x} (pe^t)^x q^{N-x} \\ &= (pe^t + q)^N \end{aligned}$$

o anche, ricordando che $q = 1 - p$,

$$M_x(t) = \left[1 + p(e^t - 1)\right]^N$$

e se, per semplificare i calcoli, ci riferiamo alla *variabile standardizzata*

$$z = \frac{x - E(x)}{\sigma_x} = \frac{x - Np}{\sqrt{Npq}} = ax + b$$

ove si è posto

$$a = \frac{1}{\sqrt{Npq}} \qquad \text{e} \qquad b = -\frac{Np}{\sqrt{Npq}}$$

applicando la (6.16) si trova

$$\begin{aligned} M_z(t) &= e^{tb} M_x(at) \\ &= e^{-\frac{Np}{\sqrt{Npq}}t} \left[1 + p\left(e^{\frac{t}{\sqrt{Npq}}} - 1\right)\right]^N \end{aligned}$$

da cui, passando ai logaritmi naturali,

$$
\begin{aligned}
\ln M_z(t) &= -\frac{Np}{\sqrt{Npq}}\,t + N\ln\left[1 + p\left(e^{\frac{t}{\sqrt{Npq}}} - 1\right)\right] \\
&= -\frac{Np}{\sqrt{Npq}}\,t + N\left[p\left(e^{\frac{t}{\sqrt{Npq}}} - 1\right) - \frac{p^2}{2}\left(e^{\frac{t}{\sqrt{Npq}}} - 1\right)^2 + \right. \\
&\quad \left. + \frac{p^3}{3}\left(e^{\frac{t}{\sqrt{Npq}}} - 1\right)^3 + \cdots\right] \\
&= -\frac{Np}{\sqrt{Npq}}\,t + N\left\{p\left[\frac{t}{\sqrt{Npq}} + \frac{1}{2}\frac{t^2}{Npq} + \frac{1}{6}\frac{t^3}{(Npq)^{\frac{3}{2}}} + \cdots\right] - \right. \\
&\quad - \frac{p^2}{2}\left[\frac{t}{\sqrt{Npq}} + \frac{1}{2}\frac{t^2}{Npq} + \frac{1}{6}\frac{t^3}{(Npq)^{\frac{3}{2}}} + \cdots\right]^2 + \\
&\quad \left. + \frac{p^3}{3}\left[\frac{t}{\sqrt{Npq}} + \frac{1}{2}\frac{t^2}{Npq} + \frac{1}{6}\frac{t^3}{(Npq)^{\frac{3}{2}}} + \cdots\right]^3 + \cdots\right\} \\
&= N\left\{p\left[\frac{1}{2}\frac{t^2}{Npq} + \frac{1}{6}\frac{t^3}{(Npq)^{\frac{3}{2}}} + \cdots\right] - \right. \\
&\quad \left. - \frac{p^2}{2}\left[\frac{t^2}{Npq} + \frac{t^3}{(Npq)^{\frac{3}{2}}} + \cdots\right] + \frac{p^3}{3}\left[\frac{t^3}{(Npq)^{\frac{3}{2}}} + \cdots\right]\right\} \\
&= N\left[\frac{1}{2}\frac{t^2}{Npq}\,p(1-p) + \frac{t^3}{(Npq)^{\frac{3}{2}}}\left(\frac{p}{6} - \frac{p^2}{2} + \frac{p^3}{3}\right) + \mathcal{O}(t^4N^{-2})\right] \\
&= \frac{1}{2}t^2 + \mathcal{O}(t^3N^{-\frac{1}{2}})
\end{aligned}
$$

ove si è sviluppato in serie di McLaurin prima

$$\ln(1+x) \;=\; x - \frac{x^2}{2} + \frac{x^3}{3} + \cdots$$

e poi

$$e^x \;=\; 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

e si sono svolti i prodotti tenendo solo i termini dei primi ordini.

Chiaramente quando N viene fatto tendere all'infinito tutti i termini eccetto il primo tendono a zero, per cui

$$\lim_{N\to\infty} M_z(t) \;=\; e^{\frac{t^2}{2}}$$

e $M_z(t)$ tende quindi alla funzione generatrice dei momenti di una distribuzione normale standardizzata; in conseguenza si è effettivamente provato, visto uno dei teoremi citati nel paragrafo 6.4, che la distribuzione binomiale tende ad una distribuzione normale.

In pratica, quando il numero di prove N è elevato e la probabilità p non è troppo vicina ai valori estremi 0 ed 1, la distribuzione binomiale è bene approssimata da una distribuzione normale; in generale si ritiene che l'approssimazione sia accettabile quando entrambi i prodotti Np e Nq hanno valore non inferiore a 5.

8.4.1 Applicazione: decadimenti radioattivi

Se la probabilità Λ_t per un singolo nucleo instabile di decadere in un intervallo di tempo t è costante, la probabilità di avere un numero prefissato di decadimenti nel tempo t in un insieme di N nuclei è data dalla distribuzione binomiale; in particolare, il numero medio di decadimenti in N nuclei e nel tempo t è $N\Lambda_t$.

Se si ammette poi che Λ_t sia proporzionale al tempo[8] t, indicando con λ la probabilità di decadimento nell'unità di tempo avremo $\Lambda_t = \lambda t$; in un tempo infinitesimo $\mathrm{d}t$, il numero di atomi $N(t)$ presente al tempo t varia mediamente di

$$\mathrm{d}N \;=\; -N\lambda\,\mathrm{d}t \;\;.$$

Separando le variabili ed integrando, il numero *medio* di atomi presenti al tempo t, dopo il decadimento di una parte di quelli N_0 presenti all'istante iniziale $t = 0$, è dato dalla

$$N(t) \;=\; N_0\,e^{-\lambda t} \tag{8.9}$$

ed il numero medio di decadimenti dalla

$$N_0 - N(t) \;=\; N_0(1 - e^{-\lambda t}) \;\;.$$

La *vita media* τ di una sostanza radioattiva si può definire come il tempo necessario perché il numero originario di nuclei si riduca mediamente di un fattore $1/e$; quindi $\tau = 1/\lambda$, e la (8.9) si può riscrivere

$$N(t) \;=\; N_0\,e^{-\frac{t}{\tau}} \;\;. \tag{8.10}$$

[8] Questa ipotesi può evidentemente essere soddisfatta solo in prima approssimazione: basta pensare al fatto che Λ_t deve raggiungere l'unità solo dopo un tempo infinito. In particolare, la probabilità per un nucleo di decadere in un tempo $2t$ vale $\Lambda_{2t} = \Lambda_t + (1-\Lambda_t)\cdot\Lambda_t = 2\Lambda_t - {\Lambda_t}^2$; e l'ipotesi fatta è in effetti valida solo se Λ_t è infinitesimo, o (in pratica) se l'osservazione riguarda un lasso di tempo trascurabile rispetto alla vita media.

8.4.2 Applicazione: il rapporto di asimmetria

Frequentemente, nella fisica, si devono considerare esperimenti in cui si cerca di mettere in evidenza delle *asimmetrie*; ovvero, la non invarianza della funzione di frequenza di una qualche variabile casuale per riflessione rispetto ad un piano. Supponiamo, come esempio, di considerare la possibilità che un dato fenomeno abbia una diversa probabilità di presentarsi *in avanti* o *all'indietro* rispetto ad una opportuna superficie di riferimento; e di raccogliere N eventi sperimentali dividendoli in due sottoinsiemi (mutuamente esclusivi ed esaurienti) collegati a queste due categorie, indicando con F e B (iniziali delle due parole inglesi *forward* e *backward*) il loro numero: ovviamente dovrà risultare $N = F + B$.

Il cosiddetto *rapporto di asimmetria*, R, si definisce come

$$R = \frac{F-B}{F+B} = \frac{F-B}{N} = \frac{2F}{N} - 1 \quad : \tag{8.11}$$

è ovvio sia che $-1 \le R \le 1$, sia che soltanto due dei quattro valori N, F, B ed R sono indipendenti; e, volendo, dall'ultima forma della (8.11) si possono ricavare le espressioni di F e B in funzione di N ed R, ovvero

$$F = \frac{N(1+R)}{2} \qquad \text{e} \qquad B = \frac{N(1-R)}{2} \quad .$$

Se indichiamo con p la probabilità di un evento in avanti (e con $q = 1 - p$ quella di uno all'indietro), il trovare esattamente F eventi in avanti su un totale di N ha probabilità data dalla distribuzione binomiale: ovvero

$$\Pr(F) = \binom{N}{F} p^F (1-p)^{N-F}$$

con, inoltre,

$$E(F) = Np \qquad \text{e} \qquad \mathrm{Var}(F) = Np\,(1-p) \quad .$$

Ma, per quanto detto, c'è una corrispondenza biunivoca tra i valori di N ed F da una parte, e quello di R; così che

$$\Pr(R) = \binom{N}{\frac{N(1+R)}{2}} p^{\frac{N(1+R)}{2}} (1-p)^{\frac{N(1-R)}{2}} \quad ,$$

$$E(R) = \frac{2}{N} E(F) - 1 = 2p - 1$$

e

$$\mathrm{Var}(R) = \frac{4}{N^2}\,\mathrm{Var}(F) = \frac{4p(1-p)}{N} \quad .$$

Se il numero di eventi nel campione è elevato e p lontano dai valori estremi, così da potere sia sfruttare l'approssimazione normale alla distribuzione di Bernoulli, sia pensare che risulti

$$p \simeq \frac{F}{N} \qquad \text{che} \qquad q \simeq \frac{B}{N} \quad ,$$

come conseguenza anche la distribuzione di R sarà approssimativamente normale; e con i primi due momenti dati da

$$E(R) \simeq 2\,\frac{F}{N} - 1 \qquad \text{e} \qquad \mathrm{Var}(R) \simeq 4\,\frac{FB}{N^3} \quad .$$

Del rapporto di asimmetria parleremo ancora più avanti, nel corso di questo stesso capitolo: più esattamente nel paragrafo 8.5.2.

8.4.3 La distribuzione binomiale negativa

Consideriamo ancora un evento casuale E ripetibile, avente probabilità costante p di presentarsi (e quindi probabilità $q = 1 - p$ di non presentarsi) in una singola prova; in più prove successive l'evento seguirà dunque la statistica di Bernoulli. Vogliamo ora calcolare la probabilità $f(x; N, p)$ che, prima che si verifichi l'N-esimo successo, si siano avuti esattamente x insuccessi; o, se si preferisce, la probabilità che l'N-simo successo si presenti nella $(N + x)$-sima prova.

L'evento casuale considerato si realizza se e solo se nelle prime $N + x - 1$ prove si è presentata una delle possibili sequenze di $N - 1$ successi e x fallimenti; e se poi, nella prova successiva, si ha un ulteriore successo. La prima condizione, applicando la (8.7), ha probabilità

$$\binom{N+x-1}{N-1} p^{N-1} q^x \quad ;$$

e, vista l'indipendenza statistica delle prove tra loro, risulta dunque

$$f(x; N, p) = \binom{N+x-1}{N-1} p^N q^x = \binom{N+x-1}{x} p^N q^x \quad . \tag{8.12}$$

(nell'ultimo passaggio si è sfruttata la proprietà dei coefficienti binomiali $\binom{N}{K} \equiv \binom{N}{N-K}$; vedi in proposito il paragrafo A.6).

Questa distribuzione di probabilità prende il nome di *distribuzione binomiale negativa*[9]; il motivo di tale nome è che l'equazione (8.12) può essere riscritta in forma compatta sfruttando una "estensione" dei coefficienti binomiali che permette di definirli anche per valori negativi di N. La funzione generatrice dei momenti è

$$M_x(t) \;=\; E(e^{tx}) \;=\; \left(\frac{p}{1-qe^t}\right)^N \quad ;$$

da questa si possono poi ricavare la speranza matematica

$$E(x) = \frac{Nq}{p} \quad ,$$

e la varianza

$$\mathrm{Var}(x) = N\,\frac{q}{p^2} \quad .$$

La distribuzione binomiale negativa con $N = 1$ prende il nome di *distribuzione geometrica*; la probabilità (8.12) diventa

$$f(x;p) = pq^x \quad ,$$

ed è quella dell'evento casuale consistente nell'ottenere il primo successo dopo esattamente x insuccessi; ponendo $N = 1$ nelle formule precedenti, speranza matematica e varianza della distribuzione geometrica sono rispettivamente

$$E(x) = \frac{q}{p} \qquad \text{e} \qquad \mathrm{Var}(x) = \frac{q}{p^2} \quad .$$

8.5 La distribuzione di Poisson

Sia E un evento casuale che avvenga rispettando le seguenti ipotesi:

1. La probabilità del verificarsi dell'evento E in un intervallo di tempo[10] molto piccolo (al limite infinitesimo) $\mathrm{d}t$ è proporzionale alla durata di tale intervallo;

[9] La distribuzione binomiale negativa è talvolta chiamata anche *distribuzione di Pascal* o *di Pólya*.

[10] Sebbene ci si riferisca, per esemplificare le nostre considerazioni, ad un processo *temporale* (e si faccia poi l'esempio del numero di decadimenti in un intervallo costante di tempo per una sostanza radioattiva come quello di una variabile casuale che segue la distribuzione di Poisson), gli stessi ragionamenti naturalmente si applicano anche a fenomeni fisici riferiti ad intervalli di differente natura, per esempio di spazio.

Così anche il numero di urti per unità di lunghezza delle molecole dei gas segue la distribuzione di Poisson (se si ammette che la probabilità di un urto nel percorrere un intervallo infinitesimo di spazio sia proporzionale alla sua lunghezza, ed analogamente per le altre ipotesi).

2. Il verificarsi o meno dell'evento in un certo intervallo temporale è indipendente dal verificarsi o meno dell'evento prima o dopo di esso;

3. La probabilità che più di un evento si verifichi in un tempo infinitesimo $\mathrm{d}t$ è infinitesima di ordine superiore rispetto a $\mathrm{d}t$.

vogliamo ora ricavare la probabilità $P(x;t)$ che in un intervallo di tempo finito, di durata t, si verifichi esattamente un numero prefissato x di eventi E. Usando questa simbologia, la prima ipotesi fatta sul processo casuale in esame si scrive

$$P(1;\mathrm{d}t) \;=\; \lambda\,\mathrm{d}t \qquad\qquad \Longrightarrow \qquad\qquad P(0;\mathrm{d}t) \;=\; 1-\lambda\,\mathrm{d}t$$

e, viste le altre ipotesi ed applicando in conseguenza i teoremi delle probabilità totali e composte, la probabilità di avere x eventi in un intervallo di tempo lungo $t+\mathrm{d}t$ è data, a meno di infinitesimi di ordine superiore, da

$$\begin{aligned} P(x;t+\mathrm{d}t) &= P(x-1;t)\cdot P(1;\mathrm{d}t) + P(x;t)\cdot P(0;\mathrm{d}t) \\ &= P(x-1;t)\,\lambda\,\mathrm{d}t + P(x;t)\,(1-\lambda\,\mathrm{d}t) \end{aligned}$$

cioè

$$\frac{P(x;t+\mathrm{d}t)-P(x;t)}{\mathrm{d}t} \;\equiv\; \frac{\mathrm{d}}{\mathrm{d}t}P(x;t) \;=\; -\lambda\,P(x;t) \;+\; \lambda\,P(x-1;t) \quad .$$

Ora, quando $x = 0$, essendo chiaramente nulla la probabilità di avere un numero negativo di eventi E in un tempo qualsiasi, risulta in particolare

$$\frac{\mathrm{d}}{\mathrm{d}t}P(0;t) = -\lambda\,P(0;t)$$

da cui

$$P(0;t) = e^{-\lambda t}$$

(la costante di integrazione si determina imponendo che $P(0;0) = 1$). Da questa relazione si può ricavare $P(1;t)$ e, con una serie di integrazioni successive, $P(x;t)$: risulta

$$P(x;t) = \frac{(\lambda t)^x}{x!}\,e^{-\lambda t} \quad . \tag{8.13}$$

In questa espressione x è l'unica variabile casuale, e t funge da parametro: se introduciamo la nuova grandezza $\alpha = \lambda t$, possiamo scrivere

$$P(x;\alpha) = \frac{\alpha^x}{x!}\, e^{-\alpha} \quad . \tag{8.14}$$

Questa distribuzione di probabilità per una variabile casuale (discreta) x prende il nome di *distribuzione di Poisson*[11]; da essa si può ottenere, ad esempio, la probabilità di ottenere x decadimenti in una massa nota di sostanza radioattiva nel tempo t: infatti per questo tipo di processi fisici risultano soddisfatte le tre ipotesi di partenza.

Più esattamente, la probabilità di avere precisamente x decadimenti radioattivi nel tempo t è data dalla distribuzione binomiale; la distribuzione di Poisson è una approssimazione alla distribuzione binomiale che si può ritenere valida qualora si considerino eventi casuali di probabilità estremamente piccola, e che ci è possibile vedere solo perché si compiono osservazioni su un numero molto elevato di essi: in formula, quando

$$p^2 \ll p \qquad \text{e} \qquad p \ll Np \ll N \tag{8.15}$$

(*eventi rari su larga base statistica*).

Anche se la distribuzione di Poisson è, come nel caso dei decadimenti radioattivi, una approssimazione di quella binomiale, si preferisce però sempre usarla nella pratica al posto di quest'ultima quando le (8.15) siano approssimativamente verificate: infatti se N è grande i fattoriali e le potenze presenti nella (8.7) rendono generalmente l'espressione difficile da calcolare. Verifichiamo ora la condizione di normalizzazione:

$$\sum_{x=0}^{+\infty} P(x) \;=\; e^{-\alpha} \sum_{x=0}^{+\infty} \frac{\alpha^x}{x!} \;=\; e^{-\alpha} e^{\alpha} \;\equiv\; 1$$

(riconoscendo nella sommatoria l'espressione di uno sviluppo in serie di McLaurin della funzione esponenziale). Calcoliamo poi la speranza matematica di x:

[11] Siméon Denis Poisson visse in Francia dal 1781 al 1840; matematico e fisico di valore, si occupò della teoria degli integrali e delle serie, di meccanica, elettricità, magnetismo ed astronomia. Gli studi sulla distribuzione che porta il suo nome compaiono nel trattato del 1837 "Recherches sur la probabilité des jugements…".

$$\begin{aligned}
E(x) &= \sum_{x=0}^{+\infty} x\, \frac{\alpha^x}{x!}\, e^{-\alpha} \\
&= \sum_{x=1}^{+\infty} x\, \frac{\alpha^x}{x!}\, e^{-\alpha} \\
&= \alpha\, e^{-\alpha} \sum_{x=1}^{+\infty} \frac{\alpha^{x-1}}{(x-1)!} \\
&= \alpha\, e^{-\alpha} \sum_{y=0}^{+\infty} \frac{\alpha^{y}}{y!} \\
&= \alpha\, e^{-\alpha} e^{\alpha} \\
&= \alpha \quad .
\end{aligned}$$

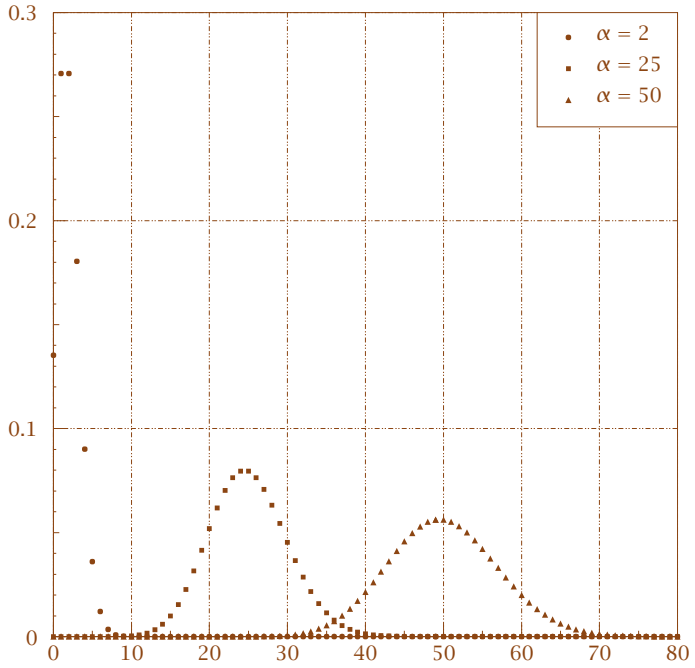
Nei passaggi si è prima osservato che il primo termine della sommatoria ($x = 0$) è nullo, e si è poi introdotta la nuova variabile $y = x - 1$. Troviamo ora la speranza matematica di x^2: con passaggi analoghi, si ottiene

$$\begin{aligned}
E(x^2) &= \sum_{x=0}^{+\infty} x^2 \frac{\alpha^x}{x!}\, e^{-\alpha} \\
&= \alpha \sum_{x=1}^{+\infty} x\, \frac{\alpha^{x-1}}{(x-1)!}\, e^{-\alpha} \\
&= \alpha \sum_{x=1}^{+\infty} \big[(x-1)+1\big] \frac{\alpha^{x-1}}{(x-1)!}\, e^{-\alpha} \\
&= \alpha \left[\sum_{y=0}^{+\infty} y\, \frac{\alpha^{y}}{y!}\, e^{-\alpha} \;+\; \sum_{y=0}^{+\infty} \frac{\alpha^{y}}{y!}\, e^{-\alpha} \right] \\
&= \alpha \left[\sum_{y=0}^{+\infty} y\, P(y) \;+\; \sum_{y=0}^{+\infty} P(y) \right] \\
&= \alpha\,(\alpha + 1)
\end{aligned}$$

e la varianza di x risulta allora anch'essa data da

$$\mathrm{Var}(x) \;=\; E(x^2) - \big[E(x)\big]^2 \;=\; \alpha\,(\alpha+1) - \alpha^2 \;=\; \alpha \quad .$$

FIGURA 8f - La distribuzione di Poisson, per tre diversi valori del parametro α.



La funzione generatrice dei momenti, come si potrebbe facilmente ottenere dalla definizione, è la

$$M_x(t) = e^{-\alpha} e^{\alpha e^t} = e^{\alpha(e^t-1)} \; ; \tag{8.16}$$

la funzione caratteristica di variabile reale

$$\phi_x(t) = e^{\alpha(e^{it}-1)} \; ,$$

e la funzione caratteristica di variabile complessa

$$\phi_x(z) = e^{\alpha(z-1)} \; . \tag{8.17}$$

Da esse potrebbero essere ricavati tutti i momenti successivi; i primi quattro valgono

$$\begin{cases} \lambda_1 \equiv E(x) = \alpha \\ \lambda_2 \equiv E(x^2) = \alpha(\alpha+1) \\ \lambda_3 = \alpha\left[(\alpha+1)^2+\alpha\right] \\ \lambda_4 = \alpha\left[\alpha^3+6\alpha^2+7\alpha+1\right] \end{cases} \qquad \begin{cases} \mu_1 \equiv 0 \\ \mu_2 \equiv \mathrm{Var}(x) = \alpha \\ \mu_3 = \alpha \\ \mu_4 = \alpha(3\alpha+1) \end{cases}$$

Un'altra conseguenza della (8.16) è che la somma $w = x + y$ di due variabili casuali indipendenti che seguano la distribuzione di Poisson (con valori medi ξ ed η) segue anch'essa tale distribuzione (con valore medio pari a $\xi + \eta$):

$$M_w(t) = e^{\xi(e^t-1)}\, e^{\eta(e^t-1)} = e^{(\xi+\eta)(e^t-1)} \; . \tag{8.18}$$

Anche la distribuzione di Poisson, come si vede dai grafici di figura 8f, è bene approssimata da una distribuzione normale quando α è abbastanza elevato; questo non deve stupire, visto lo stretto legame che esiste tra la distribuzione di Poisson e quella di Bernoulli — il cui limite per grandi N è appunto la funzione di Gauss. Volendo, si potrebbe ripetere per la funzione generatrice (8.16) una analisi analoga a quella a suo tempo compiuta per la distribuzione binomiale; in questo modo si proverebbe rigorosamente il fatto che anche la distribuzione di Poisson, per grandi α, tende a quella normale.

In genere si ritiene che, per valori medi $\alpha \gtrsim 8$, si possa ritenere soddisfacente l'approssimazione normale alla distribuzione di Poisson.

8.5.1 Applicazione: esperimenti "negativi"

Si osserva un numero N_0 di protoni per un tempo t, e non si registra alcun decadimento. Quale è il limite inferiore che si può dare sulla vita media del protone, τ, con una probabilità (*livello di confidenza*) del 95%?

L'evento casuale consistente nel non avere osservato alcun decadimento è somma logica di altri due eventi mutuamente esclusivi: o il protone è stabile (e non può quindi decadere); o il protone è instabile, e si sono inoltre verificati 0 decadimenti nel tempo di osservazione (supponiamo per semplicità che ognuno di essi abbia poi probabilità 1 di essere osservato).

In questa seconda eventualità, dalla (8.10) si può ricavare il numero medio di decadimenti attesi nel tempo t, che è

$$\alpha \;=\; N_0\left(1 - e^{-\frac{t}{\tau}}\right) \;\approx\; N_0\,\frac{t}{\tau}$$

(supponendo che τ sia molto maggiore del periodo di osservazione t); e da esso la probabilità di osservare 0 eventi sempre nel tempo t, che è data dalla statistica di Poisson e vale

$$P(0) \;=\; \frac{\alpha^0}{0!}\,e^{-\alpha} \;=\; e^{-\alpha} \;.$$

Quello che si domanda è di calcolare, assumendo come certa l'ipotesi che il protone sia instabile, il valore minimo che deve avere la sua vita media perché la probabilità di non osservare nulla sia almeno del 95%: e questo avviene quando

$$\begin{aligned} P(0) \;=\; e^{-\alpha} \;\approx\; e^{-N_0\frac{t}{\tau}} \;\geq\; 0.95 \\ -N_0\,\frac{t}{\tau} \;\geq\; \ln 0.95 \\ \tau \;\geq\; -\frac{N_0\,t}{\ln 0.95} \end{aligned}$$

(abbiamo invertito il segno della disuguaglianza nell'ultimo passaggio perché $\ln 0.95 \approx -0.0513$ è un numero negativo).

8.5.2 Applicazione: ancora il rapporto di asimmetria

Nel paragrafo 8.4.2 abbiamo supposto che il numero N di osservazioni effettuate sia noto a priori e costante: però questo non è in generale corretto; e, nella realtà, il numero di volte in cui un certo fenomeno fisico si presenterà è di norma esso stesso una variabile casuale. Continuiamo la nostra analisi

supponendo che si tratti di un fenomeno nel quale N segua la distribuzione di Poisson con valore medio ν.

Continuando ad usare gli stessi simboli del paragrafo 8.4.2, la probabilità congiunta di osservare N eventi dei quali F in avanti è in realtà

$$\begin{aligned} \Pr(F,N) &= \frac{\nu^N}{N!}\,e^{-\nu}\cdot\frac{N!}{F!\,(N-F)!}\,p^F q^{N-F} \\ &= \frac{p^F q^{N-F}}{F!\,(N-F)!}\,\nu^N e^{-\nu} \;; \end{aligned}$$

o anche, cambiando coppia di variabili casuali passando da $\{F,N\}$ a $\{F,B\}$:

$$\begin{aligned} \Pr(F,B) &= \frac{p^F q^B}{F!\,B!}\,\nu^{F+B} e^{-\nu} \\ &= \frac{(\nu p)^F(\nu q)^B}{F!\,B!}\,e^{-\nu} \\ &= \frac{(\nu p)^F}{F!}\,e^{-\nu p}\cdot\frac{(\nu q)^B}{B!}\,e^{-\nu q} \;. \end{aligned}$$

che è il *prodotto di due funzioni di frequenza di Poisson.*

In definitiva abbiamo scoperto che la composizione di un processo di Poisson e di un processo di Bernoulli equivale al prodotto di due Poissoniane: il numero N di eventi osservato segue la statistica di Poisson; la scelta dello stato finale F o B quella binomiale; ma tutto avviene come se i decadimenti dei due tipi, in avanti ed all'indietro, si verificassero *separatamente* ed *indipendentemente* secondo la statistica di Poisson.

Accettato questo fatto appena dimostrato (ancorché inaspettato), e pensando sia ad F che a B come variabili casuali statisticamente indipendenti tra loro e che seguono singolarmente la statistica di Poisson, per il rapporto di asimmetria *asintoticamente* (ovvero per grandi N) si ricava:

$$\begin{aligned} \mathrm{Var}(F) \;&=\; E(F) \;\simeq\; F \\ \mathrm{Var}(B) \;&=\; E(B) \;\simeq\; B \end{aligned}$$

e, per il rapporto di asimmetria R:

$$\begin{aligned} R &= \frac{F-B}{F+B} \\ &\simeq \frac{E(F)-E(B)}{E(F)+E(B)} + \frac{\partial R}{\partial F}\left[F-E(F)\right] + \frac{\partial R}{\partial B}\left[B-E(B)\right] \;; \end{aligned}$$

visto che la speranza matematica degli ultimi due termini è nulla,

$$\begin{aligned} E(R) &\simeq \frac{E(F)-E(B)}{E(F)+E(B)} \\ &= \frac{2F}{N} - 1 \ ; \\ \mathrm{Var}(R) &\simeq \left(\frac{\partial R}{\partial F}\right)^2 \mathrm{Var}(F) + \left(\frac{\partial R}{\partial B}\right)^2 \mathrm{Var}(B) \\ &= \frac{4}{(F+B)^4}\left[B^2\,\mathrm{Var}(F) + F^2\,\mathrm{Var}(B)\right] \\ &= \frac{4FB}{(F+B)^3} \\ &= 4\,\frac{FB}{N^3} \ , \end{aligned}$$

e le cose, di fatto, non cambiano (almeno nel limite dei grandi N) rispetto alla precedente analisi del paragrafo 8.4.2.

8.5.3 La distribuzione esponenziale

Alla distribuzione di Poisson ne è strettamente legata un'altra, quella *esponenziale*: sia infatti un fenomeno casuale qualsiasi che segua la distribuzione di Poisson, ovvero tale che la probabilità di osservare x eventi nell'intervallo finito di tempo t sia data dalla (8.13); definiamo una nuova variabile casuale, δ, come l'intervallo di tempo che intercorre tra due eventi successivi.

Visto che in un tempo δ nessun evento deve venire osservato, la probabilità che δ risulti maggiore di un valore predeterminato d coincide con la probabilità di osservare zero eventi nel tempo d:

$$\Pr(\delta > d) \ \equiv \ \Pr(0; d) \ = \ e^{-\lambda d}$$

e quindi la *funzione di distribuzione* di δ è la

$$F(d) \ = \ \Pr(\delta \le d) \ = \ 1 - e^{-\lambda d}$$

Come conseguenza, la funzione di frequenza è *esponenziale*:

$$f(\delta) \ = \ \frac{\mathrm{d}\,F(\delta)}{\mathrm{d}\delta} \ = \ \lambda\, e^{-\lambda\delta} \ ; \tag{8.19}$$

e, volendo, da essa si può ricavare la funzione caratteristica — che vale

$$\phi_\delta(t) = \frac{\lambda}{\lambda - it} \ .$$

I momenti successivi della distribuzione esponenziale si possono ottenere o integrando direttamente la funzione densità di probabilità (moltiplicata per potenze opportune di δ) o derivando successivamente la funzione caratteristica; troviamo i primi due momenti, speranza matematica e varianza, usando questo secondo metodo:

$$\frac{\mathrm{d}\,\phi_\delta(t)}{\mathrm{d}t} \ = \ \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\lambda}{\lambda - it}\right) \ = \ \frac{i\lambda}{(\lambda - it)^2}$$

$$\left.\frac{\mathrm{d}\,\phi_\delta(t)}{\mathrm{d}t}\right|_{t=0} \ = \ \frac{i}{\lambda} \ \equiv \ i \cdot E(\delta)$$

per cui la speranza matematica di δ vale

$$E(\delta) = \frac{1}{\lambda} \ ;$$

poi

$$\frac{\mathrm{d}^2\,\phi_\delta(t)}{\mathrm{d}t^2} \ = \ \frac{-i\lambda \cdot 2(\lambda - it)(-i)}{(\lambda - it)^4} \ = \ -\frac{2\lambda}{(\lambda - it)^3}$$

$$\left.\frac{\mathrm{d}^2\,\phi_\delta(t)}{\mathrm{d}t^2}\right|_{t=0} \ = \ -\frac{2}{\lambda^2} \ \equiv \ i^2\,E(\delta^2) \ = \ -E(\delta^2) \ ,$$

ed infine la varianza è

$$\mathrm{Var}(\delta) \ = \ E(\delta^2) - \left[E(\delta)\right]^2 \ = \ \frac{1}{\lambda^2} \ .$$

Se una variabile casuale t rappresenta il tempo trascorso tra due eventi casuali successivi che seguono una distribuzione di Poisson, t necessariamente ha una distribuzione di probabilità di tipo esponenziale data dalla (8.19); vogliamo ora calcolare la probabilità che t sia maggiore di una quantità $t_0 + \Delta t$, *condizionata* però dal sapere in anticipo che t è sicuramente

maggiore di t_0. Sfruttando la (3.3), abbiamo:

$$\begin{aligned}
\Pr(t > t_0 + \Delta t \mid t > t_0) &= \frac{\Pr(t > t_0 + \Delta t)}{\Pr(t > t_0)} \\
&= \frac{\int_{t_0+\Delta t}^{+\infty} e^{-\lambda t}\,\mathrm{d}t}{\int_{t_0}^{+\infty} e^{-\lambda t}\,\mathrm{d}t} \\
&= \frac{\left[-e^{-\lambda t}\right]_{t_0+\Delta t}^{+\infty}}{\left[-e^{-\lambda t}\right]_{t_0}^{+\infty}} \\
&= \frac{e^{-\lambda(t_0+\Delta t)}}{e^{-\lambda t_0}} \\
&= e^{-\lambda\,\Delta t} \\
&\equiv \Pr\left(t > \Delta t\right) \quad .
\end{aligned}$$

In conclusione, la distribuzione esponenziale (ovvero la cadenza temporale di eventi casuali che seguono la statistica di Poisson) *non ricorda la storia precedente*: il presentarsi o meno di uno di tali eventi in un tempo Δt non dipende in alcun modo da quello che è accaduto nell'arbitrario intervallo di tempo t_0 precedente; così come ci dovevamo aspettare, vista l'ipotesi numero 2 formulata a pagina 117.

8.5.4 La distribuzione di Erlang

La funzione di frequenza esponenziale (8.19) si può considerare come un caso particolare di un'altra funzione di frequenza, detta *di Erlang*[12]. Supponiamo di voler trovare la densità di probabilità $f_n(t;\lambda)$ dell'evento casuale consistente nel presentarsi, dopo un tempo t, dell'n-esimo di una serie di altri eventi che seguano la statistica di Poisson con costante di tempo λ; la (8.19) è ovviamente la prima di esse, $f_1(t;\lambda) = \lambda\, e^{-\lambda t}$.

Il secondo evento si manifesta dopo un tempo t con densità di probabilità

[12] Agner Krarup Erlang fu un matematico danese vissuto dal 1878 al 1929; si occupò di analisi e di fisica oltre che di statistica. Dette notevoli contributi alla tabulazione di varie funzioni, ed applicò in particolare la statistica a numerosi problemi relativi al traffico telefonico.

data da

$$\begin{aligned}
f_2(t;\lambda) &= \int_0^t f_1(x;\lambda)\, f_1(t-x;\lambda)\,\mathrm{d}x \\
&= \lambda^2 \int_0^t e^{-\lambda x}\, e^{-\lambda(t-x)}\,\mathrm{d}x \\
&= \lambda^2\, e^{-\lambda t} \int_0^t \mathrm{d}x \\
&= \lambda^2\, t\, e^{-\lambda t} \quad ;
\end{aligned}$$

si è infatti supposto che il primo dei due eventi si sia presentato dopo un tempo x (con $0 < x < t$), si è sfruttata l'indipendenza statistica degli eventi casuali tra loro ed infine si è sommato su tutti i possibili valori di x. Allo stesso modo

$$\begin{aligned}
f_3(t;\lambda) &= \int_0^t f_2(x;\lambda)\, f_1(t-x;\lambda)\,\mathrm{d}x \\
&= \lambda^3 \int_0^t x\, e^{-\lambda x}\, e^{-\lambda(t-x)}\mathrm{d}x \\
&= \lambda^3\, e^{-\lambda t} \int_0^t x\,\mathrm{d}x \\
&= \frac{t^2}{2}\,\lambda^3\, e^{-\lambda t} \quad ;
\end{aligned}$$

la formula generale (appunto la *funzione di frequenza di Erlang*) è la

$$f_n(t;\lambda) = \frac{t^{n-1}}{(n-1)!}\,\lambda^n\, e^{-\lambda t} \quad ,$$

con speranza matematica

$$E(t) = \frac{n}{\lambda}$$

e varianza

$$\mathrm{Var}(t) = \frac{n}{\lambda^2} \quad .$$

8.5.5 La distribuzione composta di Poisson

La *distribuzione composta di Poisson* è quella seguita da una variabile che sia somma *di un numero casuale* N di valori di un'altra variabile casuale x, quando sia N che x seguono singolarmente delle distribuzioni di Poisson. Indichiamo con ν e ξ i valori medi delle popolazioni delle variabili N e x rispettivamente; le funzioni caratteristiche (di variabile complessa) ad esse associate sono date, come sappiamo, dalla (8.17):

$$\phi_N(z) = e^{\nu(z-1)} \qquad \text{e} \qquad \phi_x(z) = e^{\xi(z-1)} \ ;$$

ricordando la (6.14), la variabile casuale

$$S = \sum_{i=1}^{N} x_i$$

ha funzione caratteristica di variabile complessa

$$\phi_S(z) \ = \ \phi_N[\phi_x(z)] \ = \ e^{\nu\left[e^{\xi(z-1)}-1\right]}$$

e funzione caratteristica di variabile reale

$$\phi_S(t) = e^{\nu\left[e^{\xi(e^{it}-1)}-1\right]} \ ;$$

da quest'ultima poi si ricava

$$\frac{\mathrm{d}\,\phi_S(t)}{\mathrm{d}t} = \phi_S(t) \cdot \nu\, e^{\xi(e^{it}-1)} \cdot \xi\, e^{it} \cdot i$$

ed infine

$$\left.\frac{\mathrm{d}\,\phi_S(t)}{\mathrm{d}t}\right|_{t=0} = i\nu\xi \ .$$

La speranza matematica di una variabile che segua la distribuzione composta di Poisson vale quindi

$$E(S) = \nu\xi \ ,$$

e, similmente, si potrebbero ottenere

$$\mathrm{Var}(S) = \nu\xi(1+\xi)$$

per la varianza, e

$$\Pr(S) = \sum_{N=0}^{\infty} \left[\frac{(N\xi)^S}{S!}\, e^{-N\xi} \cdot \frac{\nu^N}{N!}\, e^{-\nu} \right]$$

per la funzione di frequenza.

Quest'ultima formula non sorprende: è la somma (su tutti i valori ammissibili) della probabilità di ottenere un determinato N, moltiplicata per la probabilità di ottenere il valore di S condizionato da quello di N; infatti la somma di N variabili indipendenti distribuite secondo Poisson con valore medio ξ è ancora, in base a quanto dedotto dall'equazione (8.18), una variabile distribuita secondo Poisson e con valore medio $N\xi$.

8.5.6 Esempio: l'osservazione di un quark isolato

Un esempio classico di applicazione delle formule precedenti è la discussione di un articolo del 1969[13] in cui veniva annunciata l'osservazione di un quark isolato; l'esperienza ivi descritta consisteva nell'analizzare foto esposte in una camera a nebbia, attraversata da un fascio di particelle (aventi carica unitaria) che generavano tracce con un numero medio di gocce per unità di lunghezza $\alpha = 229$: su 55'000 tracce osservate ce ne era una con un numero di gocce per unità di lunghezza $n = 110$.

Questa traccia apparteneva indiscutibilmente al fascio di particelle; la probabilità che venisse osservato un numero di gocce per unità di lunghezza pari (o inferiore) a 110, se il fenomeno è descritto da una distribuzione di Poisson con media 229, è data da

$$\Pr(n \le 110) \ = \ \sum_{k=0}^{110} \frac{229^k}{k!}\, e^{-229} \ \approx \ 1.6 \times 10^{-18}$$

e risulta ben inferiore (per 13 ordini di grandezza!) alla frequenza osservata $f = 1/55'000 \approx 2 \times 10^{-5}$. Per questo motivo gli autori sostenevano di avere osservato una particella con carica frazionaria (un quark), e che causava in conseguenza una ionizzazione assai inferiore alle altre.

Una prima obiezione[14] fu che, in ogni urto elementare tra le particelle del fascio e le molecole di gas della camera, vengono generati in media $\nu = 4$ prodotti ionizzati indipendenti: e quindi 4 gocce. Il numero medio effettivo

[13] McCusker e Cairns: Evidence of quarks in air-shower cores; Phys. Rev. Lett. **23** (1969), pagg. 658–659.

[14] Adair e Kasha: Analysis of some results of quark searches; Phys. Rev. Lett. **23** (1969), pagg. 1355–1358.

di urti per unità di lunghezza era $229/4 = 57.25$, mentre la traccia osservata ne aveva invece $\lambda = 110/4 = 27.5$; la probabilità di ottenere, per motivi puramente casuali, una fluttuazione almeno pari a quella osservata doveva essere quindi calcolata come probabilità di avere meno di 28 eventi da una distribuzione di Poisson con valore medio 57.25, che vale

$$\Pr(n \le 110) \;=\; \sum_{k=0}^{27} \frac{57.25^k}{k!}\, e^{-57.25} \;\approx\; 6.7 \times 10^{-6}$$

ed è quindi assai maggiore di quanto venisse ipotizzato all'inizio dai due autori (pur mantenendosi più di 33 volte superiore alla frequenza osservata).

L'analisi del fenomeno è però ancora più complessa[15]: il numero u di urti elementari per unità di lunghezza segue la distribuzione di Poisson con valore medio $\lambda = 57.25$, ed ogni urto genera un numero di gocce che non è costante, ma segue anch'esso la distribuzione di Poisson con valore medio $\nu = 4$; quindi il numero complessivo di gocce segue una legge di distribuzione che è quella *composta di Poisson*.

La probabilità di osservare k gocce per unità di lunghezza è quindi

$$\Pr(k) \;=\; \sum_{u=0}^{\infty} \left[\frac{(u\nu)^k}{k!}\, e^{-u\nu} \cdot \frac{\lambda^u}{u!} e^{-\lambda} \right] \;,$$

e la probabilità cercata vale

$$\Pr(k \le 110) \;=\; \sum_{k=0}^{110} \Pr(k) \;\approx\; 4.7 \times 10^{-5}$$

(ben compatibile quindi con quanto osservato).

8.5.7 Applicazione: segnale e fondo

Supponiamo di osservare sperimentalmente un processo fisico, per il quale il numero di eventi s che potrebbero presentarsi in un intervallo temporale prefissato (eventi di *segnale*) segua una distribuzione di Poisson con valore medio S, e che indicheremo col simbolo $P(s;S)$;

$$\Pr(s) \;=\; P(s;S) \;=\; \frac{S^s}{s!}\, e^{-S} \quad :$$

in generale S è ignoto, e ci si propone appunto di determinarlo dall'esperimento. Questo problema verrà poi trattato anche nel paragrafo 11.2.1 a

[15] Eadie, Drijard, James, Roos e Sadoulet: Statistical Methods in Experimental Physics; North-Holland Publishing Co. (1971), pag. 53.

pagina 173, usando altri metodi; qui vogliamo solo vedere come, partendo dal dato di fatto consistente nell'osservazione effettiva di N eventi in un singolo esperimento, si possa ricavare un *limite superiore* sui possibili valori di S.

Fissato arbitrariamente un valore della probabilità ϵ, si tratta di trovare il valore S_u per il quale la probabilità di osservare un numero di eventi non superiore ad N vale proprio ϵ: vale a dire, risolvere rispetto a S_u l'equazione

$$\sum_{s=0}^{N} P(s;S_u) = \epsilon \;\;;$$

e diremo poi di poter affermare che $S \le S_u$ con un *livello di confidenza* ϵ. Il significato esatto della frase è che, se risultasse realmente $S \le S_u$, in una frazione pari almeno ad ϵ di esperimenti analoghi a quello i cui risultati stiamo esaminando ci aspetteremmo di ottenere al massimo N eventi come in esso.

Le cose si complicano in presenza di processi fisici che possono produrre risultati che simulano il segnale: processi fisici che indicheremo col nome complessivo di *fondo*. Se il fondo è presente, se è inoltre indipendente dal segnale e se segue la distribuzione di Poisson con valore medio *noto* F, già sappiamo che la probabilità di osservare N eventi in tutto *tra fondo e segnale* segue ancora la distribuzione di Poisson, con valore medio $F + S$:

$$\Pr(N) \;\equiv\; P(N;F+S) \;=\; \frac{(F+S)^N}{N!}\, e^{-(F+S)} \;\;.$$

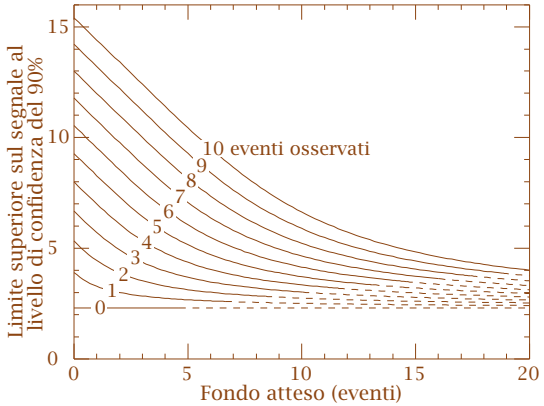
Se sono stati realmente osservati N eventi, si può ancora determinare un *limite superiore per* S; questo calcolando il valore S_u per il quale la probabilità di osservare un numero di eventi (tra fondo e segnale) non superiore a N *e condizionato all'avere ottenuto un numero di eventi di fondo che non può superare* N vale una quantità predeterminata ϵ. Insomma, si tratta di risolvere, rispetto a S_u, l'equazione

$$\frac{\sum_{n=0}^{N} P(n;F+S_u)}{\sum_{f=0}^{N} P(f;F)} = \epsilon \;\;; \tag{8.20}$$

e, con lo steso significato della frase prima evidenziato, potremo allora affermare che risulta $S \le S_u$ con un libello di confidenza ϵ.

Nella figura 8g, che è tratta dal libretto pubblicato annualmente dal "Particle Data Group" (vedi la bibliografia), si possono trovare già calcolate e rappresentate da curve continue le soluzioni dell'equazione (8.20) relative ad un livello di confidenza fisso del 90%.

FIGURA 8g - I limiti superiori sul segnale, ad un livello di confidenza fisso $\epsilon = 90\%$, in presenza di fondo noto.



8.6 La distribuzione log-normale

Sia una variabile casuale y avente distribuzione normale con media μ e varianza σ^2: la sua densità di probabilità sia insomma data dalla funzione

$$g(y) \;=\; N(y;\mu,\sigma) \;=\; \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{1}{2\sigma^2}(y-\mu)^2} \quad ;$$

definiamo poi una nuova variabile casuale x attraverso la relazione

$$x = e^y \qquad \Longleftrightarrow \qquad y = \ln x$$

(la corrispondenza tra x ed y è ovviamente biunivoca). Il dominio di definizione della x è il semiasse $x > 0$ e, in base alla (6.15), la sua densità di probabilità $f(x)$ sarà data da

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\,\frac{1}{x}\, e^{-\frac{1}{2\sigma^2}(\ln x-\mu)^2} \quad .$$

Questa funzione di frequenza si chiama *log-normale*; sfruttando l'identità

$$\begin{aligned}
\frac{(y-\mu-k\sigma^2)^2}{2\sigma^2} - k\mu - \frac{1}{2}k^2\sigma^2 &= \frac{1}{2\sigma^2}\Big(y^2+\mu^2+k^2\sigma^4-2\mu y \\
&\qquad -2k\sigma^2 y + 2k\mu\sigma^2 - 2k\mu\sigma^2 - k^2\sigma^4\Big) \\
&= \frac{1}{2\sigma^2}\Big[\big(y^2-2\mu y+\mu^2\big) - 2k\sigma^2 y\Big] \\
&= \frac{(y-\mu)^2}{2\sigma^2} - ky
\end{aligned}$$

se ne possono facilmente calcolare i momenti rispetto all'origine, che valgono

$$\begin{aligned}
\lambda_k &= \int_0^{+\infty} x^k f(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{+\infty} e^{ky}\, g(y)\,\mathrm{d}y \\
&= \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\left[\frac{(y-\mu)^2}{2\sigma^2}-ky\right]}\,\mathrm{d}y \\
&= e^{\left(k\mu+\frac{1}{2}k^2\sigma^2\right)} \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{\left(y-\mu-k\sigma^2\right)^2}{2\sigma^2}}\,\mathrm{d}y \\
&= e^{\left(k\mu+\frac{1}{2}k^2\sigma^2\right)}
\end{aligned}$$

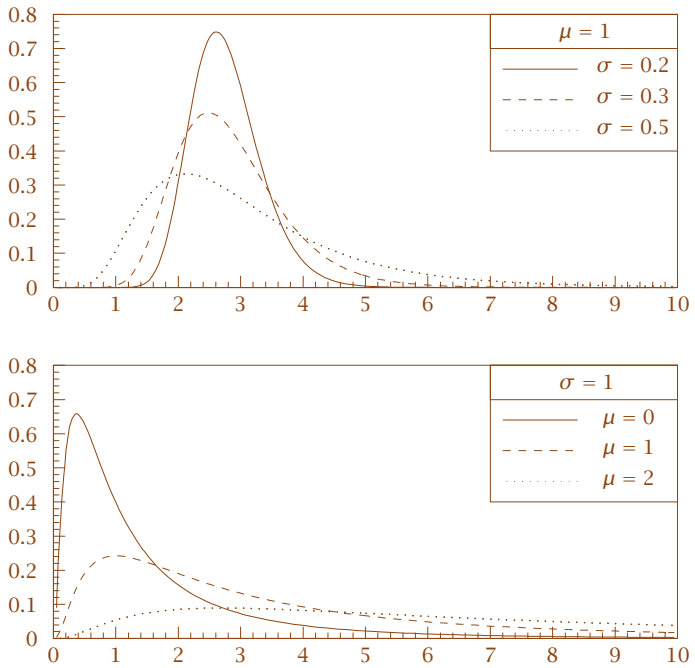
(infatti l'integrale è quello di una distribuzione normale avente $\mu + k\sigma^2$ come valore medio e σ come varianza — e vale dunque uno).

In particolare

$$\begin{aligned}
E(x) \equiv \lambda_1 &= e^{\left(\mu+\frac{\sigma^2}{2}\right)} \\
E(x^2) \equiv \lambda_2 &= e^{(2\mu+2\sigma^2)} \\
\mathrm{Var}(x) = \lambda_2 - {\lambda_1}^2 &= e^{(2\mu+\sigma^2)}\left(e^{\sigma^2}-1\right)
\end{aligned}$$

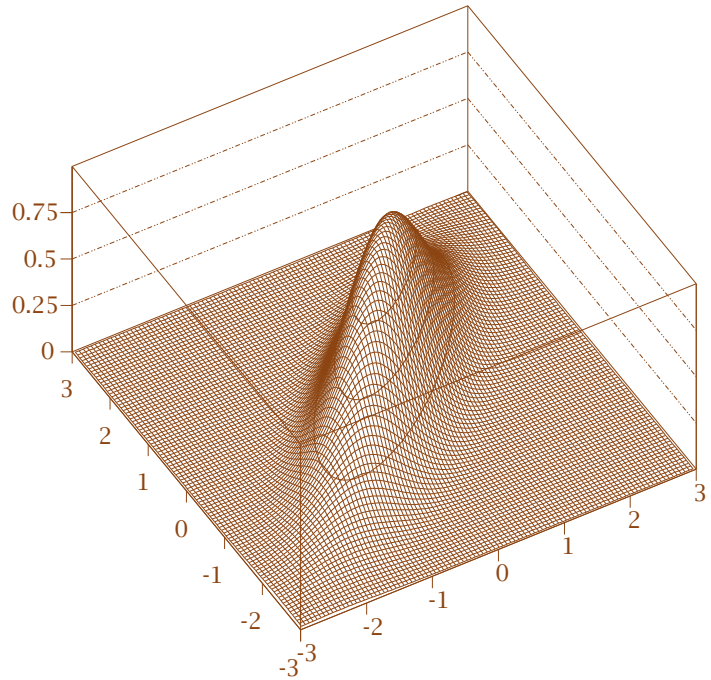
Nella figura 8h ci sono i grafici di alcune distribuzioni log-normali corrispondenti a vari valori dei parametri μ e σ della funzione normale di partenza (*non* della funzione di frequenza esaminata); per finire notiamo che, analogamente a quanto ricavato nel teorema di pagina 103 per quanto attiene alle somme, si può dimostrare che il *prodotto* di variabili casuali *log-normali ed indipendenti* debba seguire *una distribuzione log-normale*.

FIGURA 8h - La distribuzione log-normale, per vari valori dei parametri (μ e σ) della funzione normale di partenza.



8.7 La distribuzione normale in più dimensioni

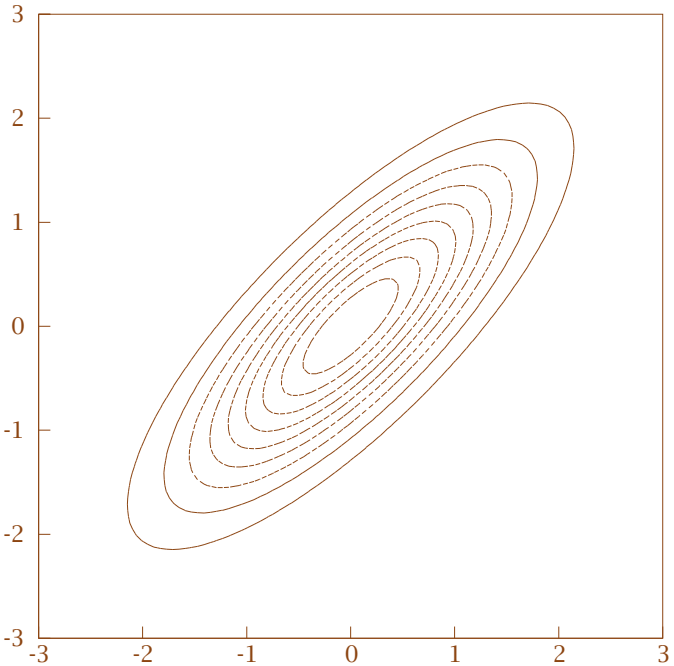
FIGURA 8i - La funzione normale in due dimensioni, nel piano $\{u, v\}$ delle variabili standardizzate e per un coefficiente di correlazione $r = 0.8$.



Accenniamo solo brevemente alla cosiddetta *distribuzione normale bidimensionale*: per essa la densità di probabilità congiunta delle due variabili x ed y è, nel caso generale, data dalla

$$f(x,y) = \frac{e^{\left\{-\frac{1}{2(1-r^2)}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2 - 2r\frac{(x-\mu_x)(y-\mu_y)}{\sigma_x\,\sigma_y} + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right]\right\}}}{2\pi\,\sigma_x\,\sigma_y\sqrt{1-r^2}}$$

FIGURA 8j - La sole curve di livello della stessa funzione rappresentata nella figura 8i.



$r = 0.8$

o, espressa in funzione delle *variabili standardizzate*

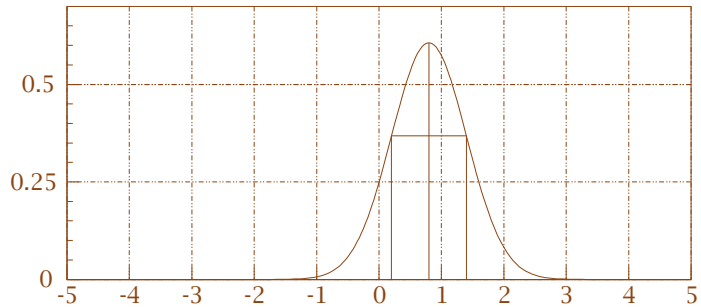
$$u = \frac{x - E(x)}{\sigma_x} \qquad \text{e} \qquad v = \frac{y - E(y)}{\sigma_y} \ ,$$

dalla

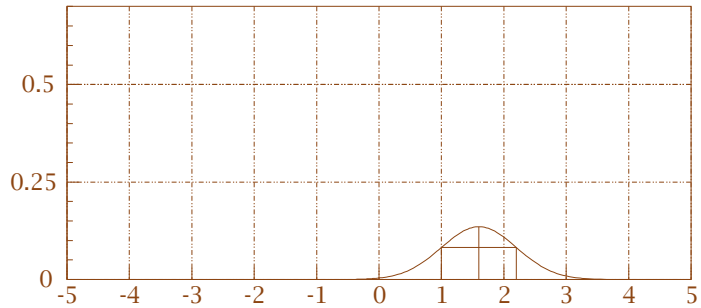
$$f(u, v) = \frac{e^{-\frac{1}{2}\frac{u^2 - 2ruv + v^2}{1 - r^2}}}{2\pi\sqrt{1 - r^2}} \ ;$$

un esempio è dato dalle figure 8i e 8j.

FIGURA 8k - Le sezioni della figura 8i con due piani di equazione $y = 1$ e $y = 2$ rispettivamente: ovvero (a parte la normalizzazione) le densità di probabilità condizionate $\pi(x|y = 1)$ e $\pi(x|y = 2)$.



Sezione a $y = 1$ - $f(x, 1) \sim N(0.8, 0.6)$



Sezione a $y = 2$ - $f(x, 2) \sim N(1.6, 0.6)$

r è il coefficiente di correlazione lineare tra x ed y; $r = 0$ è condizione *necessaria e sufficiente* perché le due variabili siano statisticamente indipendenti tra loro. Le due distribuzioni marginali $g(x)$ e $h(y)$ sono due funzioni normali, aventi speranza matematica e varianza μ_x e ${\sigma_x}^2$ la prima, e μ_y e ${\sigma_y}^2$ la seconda:

$$g(x) = N(x; \mu_x, \sigma_x) \qquad \text{e} \qquad h(y) = N(y; \mu_y, \sigma_y) \ .$$

Le densità di probabilità condizionate sono anch'esse sempre delle funzioni normali; come esempio, nella figura 8k si mostrano due di queste funzioni per la stessa distribuzione di figura 8i.

Nel caso più generale, la densità di probabilità di una distribuzione di Gauss N-dimensionale è del tipo

$$f(x_1, x_2, \ldots, x_N) = K e^{-\frac{1}{2} H(x_1, x_2, \ldots, x_N)} \ , \tag{8.21}$$

ove H è una *forma quadratica* nelle variabili standardizzate

$$t_i = \frac{x_i - E(x_i)}{\sigma_i}$$

nella quale però i coefficienti dei termini quadratici sono *tutti uguali*; le t_i non sono generalmente indipendenti tra loro, e quindi H contiene anche i termini rettangolari del tipo $t_i \cdot t_j$.

K, nella (8.21), è un fattore di normalizzazione che vale

$$K = \sqrt{\frac{\Delta}{(2\pi)^N}} \ ;$$

a sua volta, Δ è il determinante della matrice (*simmetrica*) dei coefficienti della forma quadratica H. La condizione, poi, che la f di equazione (8.21) debba essere integrabile implica che le ipersuperfici di equazione $H =$ cost. siano tutte al finito, e siano quindi *iperellissi* nello spazio N-dimensionale dei parametri; le funzioni di distribuzione marginali e condizionate di *qualsiasi* sottoinsieme delle variabili sono ancora *sempre normali*.

Si può inoltre dimostrare che *esiste sempre* un cambiamento di variabili $x_i \to y_k$ che muta H nella cosiddetta *forma canonica* (senza termini rettangolari); in tal caso

$$\Delta = \left[\prod_{k=1}^{N} \mathrm{Var}(y_k) \right]^{-1}$$

e

$$f(y_1, \ldots, y_k) = \frac{1}{\sqrt{\prod_{k=1}^{N} [2\pi \, \mathrm{Var}(y_k)]}} e^{-\frac{1}{2} \sum_{k=1}^{N} \frac{[y_k - E(y_k)]^2}{\mathrm{Var}(y_k)}} \ :$$

e le y_k sono quindi tutte *statisticamente indipendenti* (oltre che normali).

Capitolo 9

La legge di Gauss

Vogliamo ora investigare sulla distribuzione dei risultati delle misure ripetute di una grandezza fisica, nell'ipotesi che esse siano affette da errori esclusivamente casuali[1].

9.1 La funzione di Gauss

Dall'esame di molte distribuzioni sperimentali di valori ottenuti per misure ripetute in condizioni omogenee, si possono astrarre due proprietà generali degli errori casuali:

- La probabilità di ottenere un certo scarto dal valore vero deve essere funzione del modulo di tale scarto e non del suo segno, se valori in difetto ed in eccesso rispetto a quello vero si presentano con uguale probabilità; in definitiva la distribuzione degli scarti deve essere *simmetrica rispetto allo zero*.
- La probabilità di ottenere un certo scarto dal valore vero (in modulo) deve essere *decrescente* al crescere di tale scarto e *tendere a zero* quando esso tende all'infinito; questo perché deve essere più probabile commettere errori piccoli che errori grandi, ed infinitamente improbabile commettere errori infinitamente grandi.

[1] Il primo ad intuire la forma e l'equazione della distribuzione normale fu Abraham de Moivre nel 1733, che la derivò dalla distribuzione binomiale facendo uso della formula di Stirling per il fattoriale; fu poi studiata da Laplace, ma la teoria completa è dovuta a Gauss.

A queste due ipotesi sulla distribuzione delle misure affette da errori puramente casuali se ne può aggiungere una terza, valida per *tutte* le distribuzioni di probabilità; la condizione di normalizzazione, ossia l'equazione (6.1) di cui abbiamo già parlato prima:

- L'area compresa tra la curva densità di probabilità dello scarto e l'asse delle ascisse, da $-\infty$ a $+\infty$, deve valere 1.

Da queste tre ipotesi e dal principio della media aritmetica, Gauss[2] dimostrò in modo euristico che la distribuzione degli scarti z delle misure affette da errori casuali è data dalla funzione

$$\boxed{f(z) = \frac{h}{\sqrt{\pi}} e^{-h^2 z^2}} \tag{9.1}$$

che da lui prese il nome (*funzione di Gauss* o *legge normale* di distribuzione degli errori). Si deve originalmente a Laplace una prova più rigorosa ed indipendente dall'assunto della media aritmetica; una versione semplificata di questa dimostrazione è data nell'appendice D.

La funzione di Gauss ha una caratteristica forma a campana: simmetrica rispetto all'asse delle ordinate (di equazione $z = 0$), decrescente man mano che ci si allontana da esso sia nel senso delle z positive che negative, e tendente a 0 per z che tende a $\pm\infty$; così come richiesto dalle ipotesi di partenza.

Essa dipende da un parametro $h > 0$ che prende il nome di *modulo di precisione* della misura: infatti quando h è piccolo la funzione è sensibilmente diversa da zero in una zona estesa dell'asse delle ascisse; mentre al crescere di h l'ampiezza di tale intervallo diminuisce, e la curva si stringe sull'asse delle ordinate (come si può vedere nella figura 9a).

9.2 Proprietà della legge normale

Possiamo ora, come già anticipato nei paragrafi 4.2.6 e 6.2, determinare il valore medio di una qualsiasi grandezza $W(z)$ legata alle misure, nel limite

[2] Karl Friedrich Gauss fu senza dubbio la maggiore personalità del primo 800 nel campo della fisica e della matematica; si occupò di teoria dei numeri, analisi, geometria analitica e differenziale, statistica e teoria dei giochi, geodesia, elettricità e magnetismo, astronomia e ottica. Visse a Göttingen dal 1777 al 1855 e, nel campo di cui ci stiamo occupando, teorizzò (tra le altre cose) la funzione normale ed il metodo dei minimi quadrati, quest'ultimo (studiato anche da Laplace) all'età di soli 18 anni.

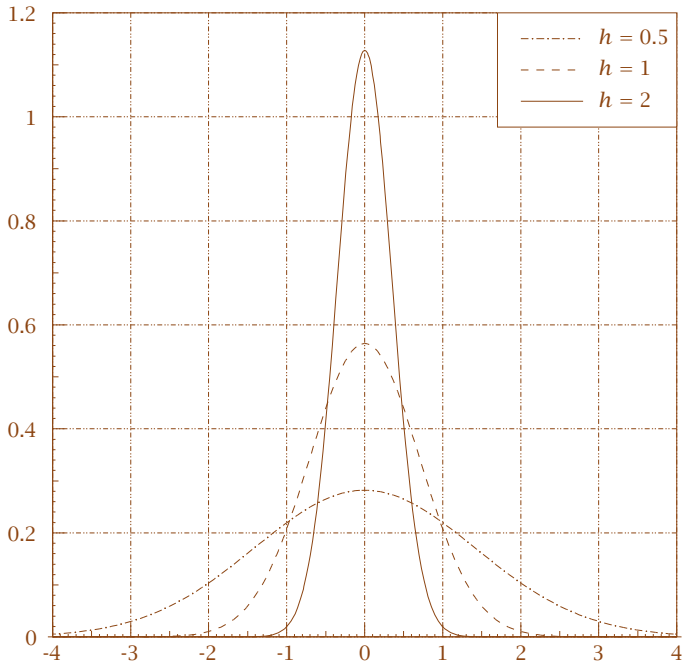



FIGURA 9a - La funzione di Gauss per tre diversi valori di h.