

# Appunti di probabilità

Pisa – Anno Accademico 1973/74



---

## Premessa

Questi Appunti sono la trascrizione dei fogli ciclostilati degli appunti di probabilità del corso di Esperimentazioni di fisica I (soprannominato dagli studenti con il nomignolo di "Fisichetta 1") – Pisa – A.A. 1973–74.



---

## Indice

<b><u>Introduzione</u></b> .....	1
<b><u>CAPITOLO I</u></b> .....	3
1 <u>Definizione di probabilità</u> .....	3
2 <u>Proprietà e leggi della probabilità</u> .....	4
3 <u>Variabili casuali e funzioni di distribuzione</u> .....	6
4 <u>Caratteristiche comuni alle distribuzioni</u> .....	9
5 <u>Misure di dispersione intorno alla media</u> .....	10
<b><u>CAPITOLO II:</u></b>	
<b><u>PROPAGAZIONE DEGLI ERRORI - Parte I</u></b> .....	15
1 <u>Errore massimo</u> .....	15
2 <u>Errore relativo</u> .....	17
3 <u>Nozione di errore sistematico</u> .....	18
4 <u>Nozione di precisione ed accuratezza</u> .....	18
5 <u>Cifre significative e convenzioni di scrittura dei risultati delle misure</u>	18
<b><u>CAPITOLO III:</u></b>	
<b><u>DISTRIBUZIONI</u></b> .....	25
1 <u>Distribuzione binomiale</u> .....	25
2 <u>Distribuzione di Poisson</u> .....	28
3 <u>Distribuzione di Gauss</u> .....	30
4 <u>Distribuzione del <math>\chi^2</math></u> .....	39
<b><u>CAPITOLO IV</u></b> .....	41
1 <u>Propagazione degli errori</u> .....	41
2 <u>Teoria dei campioni</u> .....	43
<b><u>CAPITOLO V</u></b> .....	47
1 <u>Problemi di fit</u> .....	47
2 <u>Metodo del minimo <math>\chi^2</math></u> .....	50

VIII    Indice

3	<u>Fit di tipo generale</u> .....	51
4	<u>Test del <math>\chi^2</math></u> .....	53
<b><u>CAPITOLO VI</u></b> .....		59
1	<u>Formule approssimate per la derivata di una funzione</u> <u><math>f(x)</math> data per punti</u> .....	59
2	<u>Formule approssimate per gli integrali di una funzione</u> <u><math>f(x)</math>, data per punti</u> .....	60
3	<u>Interpolazione ed estrapolazione</u> .....	62

---

## 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, ciò 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.

Ci sono due situazioni tipiche che limitano la precisione delle misure:

- 1) Lo strumento ha dei limiti. Si parla allora di errore strumentale ed il risultato delle misure in generale 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. È chiaro che nella fisica moderna ha molta più importanza la 2<sup>a</sup> situazione perché quasi sempre si lavora al limite delle tecniche sperimentali.

Si comincerà quindi a studiare come ci si deve comportare nel caso in cui i risultati delle misure sono fluttuanti in modo imprevedibile e casuale: gli strumenti matematici e necessari sono la probabilità e la statistica.





---

## CAPITOLO I

### 1 Definizione di probabilità

1ªdefinizione: Probabilità di un evento E, indicata con P(E), è uguale al rapporto tra il numero di casi favorevoli e quello dei casi possibili.

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

Questa definizione è sufficiente per i casi più elementari.

Esempio: a) Lancio di una moneta: qual è la probabilità che esca testa (o croce)? Ci sono: un caso favorevole e due casi possibili:

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

b) Lancio di un dado: qual è la probabilità che esca un certo numero fissato, il 3 per esempio. Ci sono un caso favorevole e 6 casi possibili:

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

c) Lancio di 2 monete: qual è la probabilità che esca almeno una testa. I casi possibili sono: TT, TC, CT, CC. I casi favorevoli sono: TT, TC, CT

$$P(1T) = \frac{3}{4}$$

A questa definizione può essere ostata la seguente criticità: non appena la situazione si complica leggermente diventa necessario tenere conto che alcuni casi possono essere più o meno "favorevoli". Quindi si aggiunge alla definizione la clausola: "Purché siano ugualmente possibili" che ci riporta al problema di definire la probabilità.

2ªdefinizione: Statistica - Sperimentale.

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

dove  $n$  = numero di volte in cui si verifica l'evento  $E$ , risultato di un certo esperimento;  $N$  = numero totale di volte in cui si ripete l'esperimento;  $\frac{n}{N}$  è la frequenza relativa statistica dell'evento  $E$ .

Questo limite significa che se si compiono più serie di prove, con  $N$  sempre più grande, il rapporto  $\frac{n}{N}$  tende a stabilizzarsi intorno ad un certo valore, con oscillazioni sempre più piccole man mano  $N$  cresce.

Si propone di verificarlo sperimentalmente, per esempio per l'evento "testa" nel lancio di una moneta.

Questa definizione è buona per le applicazioni, ma non per una rigorosa costruzione matematica.

3ª definizione: Usando la teoria della misura.

Sia dato un insieme  $S$  :  $S = \{\text{insieme degli eventi } E\}$ . Si considerino tanti sottoinsiemi  $s$  di  $S$  ( $s \subset S$ ). Sia  $\mathcal{E}$  l'insieme dei sottoinsiemi  $s$ .

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

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

$$P(S) = 1 \tag{1.1}$$

Significa che la probabilità di osservare un evento qualsiasi tra quelli che possono essere osservati è  $= 1$ , od anche c'è la certezza di osservare uno qualsiasi degli eventi che possono essere osservati.

$$\forall A, B \in \mathcal{E} \quad (\text{cioè } A, B \subset S)$$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

## 2 Proprietà e leggi della probabilità

a) Addizione: Se  $A$  e  $B$  sono due eventi disgiunti, mutuamente esclusivi, la probabilità che si verifichi l'evento  $A$  o l'evento  $B$  (probabilità che  $A$  o  $B = C$ ) è

$$P(C) = P(A) + P(B) \quad \text{Legge dell'addizione}$$

Esempio: Si lanci un dado, la probabilità che esca il 2 o il 3 è:

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

Contro esempio: Estrarre una carta da un mazzo di 52 carte.

$$A = \text{fiori} \qquad P(A) = \frac{13}{52}$$

$$B = \text{re} \qquad P(B) = \frac{4}{52}$$

$C = \text{un re o una fiori}$

$$P(C) = \frac{16}{52} = \frac{\text{casi favorevoli}}{\text{casi possibili}}$$

(i casi favorevoli sono 16, perché ci sono 13 carte a fiori che comprendono il re di fiori più gli altri tre re).

$$P(C) \neq P(A) + P(B)$$

Bisogna sottrarre la probabilità che esca il re di fiori (altrimenti verrebbe contata due volte). Quindi:

$$P(C) = \frac{13}{52} + \frac{4}{52} - \frac{1}{52}$$

Segue quindi la generalizzazione della legge dell'addizione:

$$P(A \text{ o } B) = P(A) + P(B) - P(A \text{ e } B)$$

Esercizio: Sia data una scatola con 6 palline rosse, 3 blu e 5 bianche. Cercare la probabilità che esca una pallina blu o bianca.<sup>1</sup>

b) Moltiplicazione: Consideriamo due eventi A, B e l'evento  $C = A \text{ e } B$ . Quanto vale  $P(C)$ ?

Supponiamo prima che il fatto che si verifichi A non influenzi in nessun modo il verificarsi di B.

La legge della moltiplicazione quando A e B sono indipendenti è:

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

Infatti

$$P(A) = \frac{n_A}{N_A} \qquad P(B) = \frac{n_B}{N_B}$$

I casi possibili dell'evento A e B sono  $N_A \times N_B$ , i casi favorevoli sono  $n_A \cdot n_B$

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

Esempio: Lancio di un dado ed una moneta. L'evento C sia l'uscita di una testa e del 3,  $C = T \text{ e } 3$

$$P(C) = \frac{1}{6} \cdot \frac{1}{2} = \frac{1}{12}$$

---

<sup>1</sup> Svolgimento: Evidentemente in questo caso  $P(\text{blu e bianca}) = 0$ , quindi  $P(\text{blu o bianca}) = P(\text{blu}) + P(\text{bianca}) = \frac{3}{14} + \frac{5}{14} = \frac{8}{14}$  [ndc].

Lancio di 1 dado,  $C = 2$  e  $3$ :  $P(C) = 0$

Lancio di 2 dadi,  $C = 2$  e  $3$ :  $P(C) = \frac{1}{36}$

52 carte:  $C = \text{re di fiori}$

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

Contro esempio: Consideriamo 52 carte più un jolly.  $C = \text{re di fiori}$ .

$$P(\text{fiori}) = \frac{13}{53}$$

$$P(\text{re}) = \frac{4}{53}$$

$$P(\text{re di fiori}) = \frac{1}{53} \neq \frac{13}{53} \cdot \frac{4}{53}$$

Infatti gli eventi non sono indipendenti.

(Se esce una fiori questo esclude che sia un jolly quindi modifica la possibilità che sia un re). Generalizzazione della legge della moltiplicazione quando A e B non sono indipendenti.

Definiamo la probabilità condizionale:  $P(A/B) = \text{Probabilità di A una volta che se è verificato B}$ .

$$P(A \text{ e } B) = P(B) \cdot P\left(\frac{A}{B}\right)$$

Nell'esempio sopra:  $P(\text{re di fiori}) = \frac{4}{53} \cdot \frac{1}{4} = \frac{1}{53}$

Per eventi indipendenti  $P\left(\frac{A}{B}\right) = P(A)$ .

### 3 Variabili casuali e funzioni di distribuzione

Lanciamo due dadi e calcoliamo la somma dei risultati S.

S è un numero che è una funzione del lancio dei dadi, è quindi una variabile casuale e per ogni valore di S si può calcolare la probabilità che si ottenga proprio quel valore (tavola 1).

Si può costruire un istogramma con  $P(S)$  ed S: fig. 1.

L'insieme dei valori  $P(S)$  costituisce la funzione di distribuzione della variabile

S. Quindi conoscere la funzione di distribuzione di una variabile casuale significa conoscere la probabilità associata a ciascun valore della variabile casuale.

Il caso ora considerato è quello di una variabile casuale discreta (cioè può assumere soltanto valori discreti) in un intervallo finito.

Esistono casi di variabili discrete in un intervallo infinito, per esempio il numero di gocce di pioggia che cadono in un tempo determinato su un'area determinata.

Esistono casi di variabili continue su intervalli finiti o infiniti.

Per esempio supponiamo di sparare con una carabina contro un bersaglio. La

tav. 1

S	casi favorevoli	P(S)
2	1	1/36
3	2	2/36
4	3	3/36
5	4	4/36
6	5	5/36
7	6	6/36
8	5	5/36
9	4	4/36
10	3	3/36
11	2	2/36
12	1	1/36

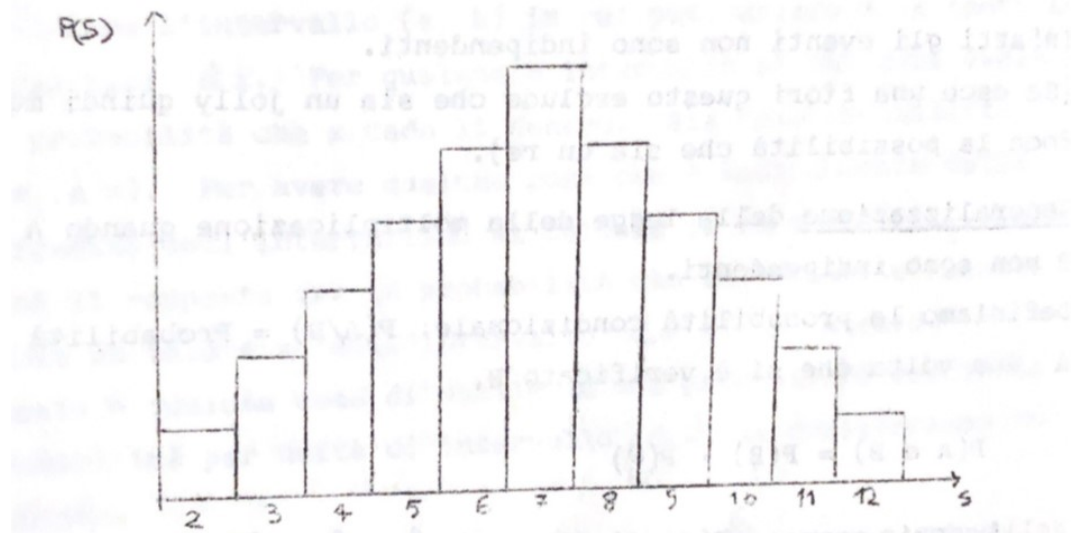


fig. 1

fig. 1.

distanza dal centro a cui arriva il colpo è una variabile casuale continua su un intervallo finito.

Come si può allora definire la distribuzione di una variabile casuale  $x$ : per i casi discreti la distribuzione è una funzione che associa a qualunque valore della variabile casuale  $x$  la sua probabilità.

Nel caso di una variazione continua di  $x$  questa definizione non è applicabile: è chiaro che per un valore esattamente definito la probabilità sarà zero. Bisogna quindi chiedersi quale è la probabilità che si abbia un valore in un intervallo assegnato. Dividiamo l'intervallo  $(a, b)$  in cui può variare  $x$  in tanti intervallini  $\Delta x$ . per qualunque intervallo si può dire quale è la probabilità che  $x$  cada lì dentro. Sia tale probabilità  $P(x_i, \Delta x)$ . Per avere qualche cosa che è indipendente dalla larghezza dell'intervallino si calcola il rapporto  $P(x_i, \Delta x)/\Delta x$ , cioè il rapporto tra la probabilità che la variabile casuale  $x$  abbia un valore  $x_i$  nell'intervallo  $\Delta x$  e  $\Delta x$  stesso.

Questa è qualche cosa di simile ad una probabilità specifica o probabilità per unità di intervallo  $\Delta x$ ; se consideriamo il limite:

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

si ottiene la densità di probabilità che per definizione è la distribuzione della variabile casuale  $x$ .

Premettiamo alcune definizioni.

Si definisce  $\sum_{i=1}^n a_i$  la somma di  $n$  elementi  $a_1, a_2, \dots, a_i, \dots, a_n$ .

Proprietà formali:

a)  $\sum_i a_i = \sum_j a_j = \sum_k a_k$  cioè l'indice è un indice muto

b)  $\sum_i C a_i = C \sum_i a_i$

c)  $\sum_i (a_i + b_i) = \sum_i a_i + \sum_i b_i$

d)  $\left(\sum_i a_i\right)^2 = \sum_{i,j} a_i a_j$

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

$$a_1 \cdot a_1 + a_1 \cdot a_2 + a_1 \cdot a_3 + a_2 \cdot a_1 + a_2 \cdot a_2 + a_2 \cdot a_3 + a_3 \cdot a_1 + a_3 \cdot a_2 + a_3 \cdot a_3 = \sum_{i,j} a_i a_j$$

Esercizi:<sup>2</sup>

<sup>2</sup> Svolgimento:

Indichiamo  $\sum_{i=1}^n 1 = 1 + 1 + \dots + 1$  ( $n$  volte) con  $S_0$ . Evidentemente  $S_0 = n$ . Ne segue per la proprietà b):  $\sum_{i=1}^n k = k S_0 = kn$ .

Indichiamo  $\sum_{i=1}^n i$  con  $S_1$ . Distinguiamo 2 casi: 1)  $n$  pari

$$S_1 = 1+2+\dots+(n-1)+n = (1+n)+(2+n-1)+\dots \left(\frac{n}{2} \text{ volte}\right) \dots + \left(\frac{n}{2} + \frac{n}{2} + 1\right) = \frac{n}{2} \cdot (n+1)$$

2)  $n$  dispari

$$S_1 = (1+n-1) + (2+n-2) + \dots \left(\frac{n-1}{2} \text{ volte}\right) \dots + \left(\frac{n-1}{2} + \frac{n-1}{2} + 1\right) + n = \frac{n-1}{2} \cdot n + n = n \cdot \left(\frac{n+1}{2}\right). \text{ Ne segue in ogni caso } \sum_{i=1}^n i = S_1 = \frac{n(n+1)}{2}.$$

Indichiamo  $\sum_{i=1}^n i^2$  con  $S_2$  e consideriamo quanto segue:

$(i+1)^3 = i^3 + 3i^2 + 3i + 1$ ; sommando membro a membro per  $i$  da 1 a  $n$  ed eliminando i termini uguali in entrambi i membri,

$$(n+1)^3 = 1+3(1^2+2^2+\dots+n^2)+3(1+2+\dots+n)+(1+1+\dots(n \text{ volte})\dots+1) =$$

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

#### 4 Caratteristiche comuni alle distribuzioni

a) Normalizzazione: Data una variabile  $x_i$  discreta e la sua funzione di distribuzione  $P(x)$ , la (1.1) dice che la somma su tutti i casi possibili delle probabilità è uguale all'unità:

$$\sum_{x_i} P(x_i) = 1$$

Questa è la condizione di normalizzazione per la variabile discreta. Nel caso di variabile continua definita nell'intervallo  $(a, b)$  la condizione di normalizzazione diventa:

$$\int_a^b p(x) dx = 1$$

dove  $p(x)$  è la densità di probabilità.

Naturalmente  $\int_{x_1}^{x_2} p(x) dx = P(x_1 \leq x \leq x_2)$

b) Media: Sia  $x_i$  una variabile casuale discreta, la sua funzione di distribuzione sia  $P(x_i)$ . La media  $\mu$  della variabile  $x$  è definita in questo modo:

$$\mu = \sum_{x_i} x_i P(x_i)$$

Possiamo giustificare questa definizione tenendo presente che media nel significato comune è la media aritmetica:

$$\mu_a = \frac{1}{n} \sum_1^n x_i \quad (1.2)$$

e  $\mu_a$  e  $\mu$  coincidono nel caso semplice che gli  $x_i$  siano equiprobabili. Infatti

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

$$\sum_i P(x_i) = 1$$

ne segue che  $P(x_i) = \frac{1}{n}$ .

La (1.2)<sup>3</sup> dice che nel caso in cui gli eventi non sono equiprobabili ogni  $x_i$

---


$$S_2 = \frac{1 + 3S_2 + 3S_1 + S_0}{(n+1)^3 - 3S_1 - S_0 - 1} = \frac{(n+1)^3 - 3n(n+1)/2 - n - 1}{3} = \frac{n+1}{3} \left( (n+1)^2 - 3n/2 - 1 \right) =$$

$$\frac{n+1}{3} \cdot \frac{2n^2+n}{2} = \frac{n(n+1)(2n+1)}{6} \text{ [n.d.c.]}$$

<sup>3</sup> Più precisamente la definizione di  $\mu$  [n.d.c.].

viene pesato in modo diverso a seconda della sua probabilità.

c) Mediana di una distribuzione è definita come quel valore della variabile  $\mu_{\frac{1}{2}}$  tale che

$$P(x_i \leq \mu_{\frac{1}{2}}) = P(x_i \geq \mu_{\frac{1}{2}})$$

Il valore più probabile  $\mu_{max}$  è il valore della variabile per il quale la distribuzione ha un massimo

$$\mu_{max} : P(x_i) \leq P(\mu_{max})$$

Se la distribuzione è simmetrica  $\mu$ ,  $\mu_{\frac{1}{2}}$  e  $\mu_{max}$  coincidono.

Nel caso continuo le definizioni di media, mediana, valore più probabile diventano:

$$\mu = \int_a^b xP(x)dx$$

$$\mu_{\frac{1}{2}} : \int_a^{\mu_{\frac{1}{2}}} p(x)dx = \int_{\mu_{\frac{1}{2}}}^b p(x)dx = \frac{1}{2}$$

$\mu_{max}$  è il valore che rende massimo  $p(x)$ , fig. 2.

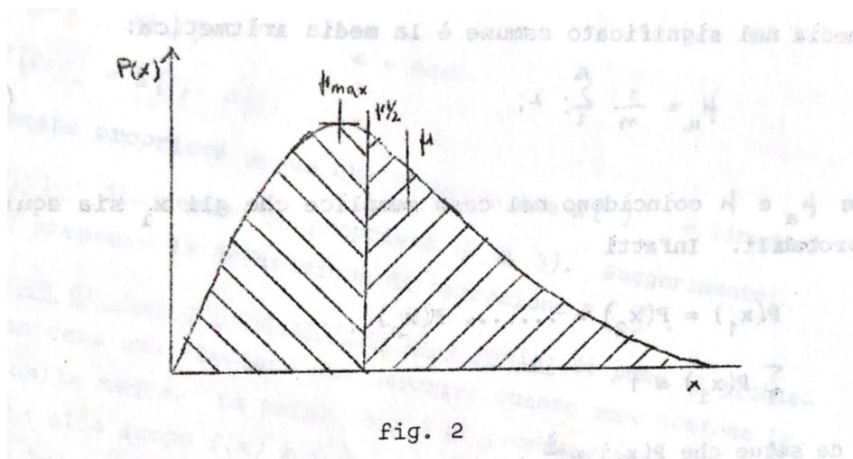


fig. 2.

Se la funzione di distribuzione è simmetrica  $\mu$ ,  $\mu_{\frac{1}{2}}$  e  $\mu_{max}$  coincidono.

## 5 Misure di dispersione intorno alla media

a) Definizione del valore di aspettazione o di previsione.

Consideriamo una funzione  $f(x)$  della variabile casuale  $x$ :  $E[f]$  si chiama valore di aspettazione o di previsione della  $f$  ed è:



$$E[f] = \sum_{x_i} f(x_i)P(x_i)$$

se  $x$  è una variabile discreta si chiama valore di aspettazione o di previsione della  $f$  ed è:

$$E[f] = \int_a^b f(x)p(x)dx$$

se  $x$  è una variabile continua.

È chiaro che se  $f(x) = x$

$$E[f] = E[x] = \mu$$

Per questo  $E[f]$  si chiama anche il valor medio di  $f$ .

Proprietà formali dell'operazione  $E[ ]$

- 1)  $E[\text{cost}] = \text{cost}$
- 2)  $E[af] = aE[f] \quad a = \text{cost}$
- 3)  $E[f + g] = E[f] + E[g]$

Da queste proprietà segue che l'operazione  $E[ ]$  è lineare.

Esercizio:<sup>4</sup> dimostrare le proprietà 1), 2), 3). Suggerimento: tenere presente la definizione di operazione  $E[ ]$ .

b) Misure di dispersione intorno alla media: Si pone il problema di cercare una funzione che descriva quanto sono lontane le misure dalla media. La prima idea è di prendere quale funzione adatta allo scopo  $f(x) = x - \mu$ , cioè la funzione che rappresenta lo scarto dalla media e calcolare quindi  $E[x - \mu]$ , ma  $E[x - \mu] = E[x] - E[\mu] = \mu - \mu = 0$ .

Quindi questa funzione non ci dà nessuna informazione come era logico aspettarsi perché gli errori per difetto compensano statisticamente gli errori per eccesso. Allora si può pensare di prendere  $f(x) = |x - \mu|$ , che rappresenta la deviazione dalla media, ma questa funzione è scomoda per i calcoli. Conviene quindi prendere  $f(x) = (x - \mu)^2$ , lo scarto quadratico e calcolare  $E[(x - \mu)^2]$ . Si definisce varianza, e si indica con  $\sigma^2$  il valore di aspettazione dello scarto quadratico:

<sup>4</sup> Svolgimento:

1)  $E[C] = \sum_{x_i} CP(x_i) = C \sum_{x_i} P(x_i) = C \cdot 1 = C$ ;  $E[C] = \int_a^b Cp(x)dx = C \int_a^b p(x)dx = C \cdot 1 = C$   
 2)  $E[af] = \sum_{x_i} af(x_i)P(x_i) = a \sum_{x_i} f(x_i)P(x_i) = aE[f]$ ;  $E[af] = \int_a^b af(x)p(x)dx = a \int_a^b f(x)p(x)dx = aE[f]$   
 3)  $E[f + g] = \sum_{x_i} [f(x_i)P(x_i) + g(x_i)P(x_i)] = \sum_{x_i} f(x_i)P(x_i) + \sum_{x_i} g(x_i)P(x_i) = E[f] + E[g]$ ;  $E[f + g] = \int_a^b [f(x)p(x) + g(x)p(x)]dx = \int_a^b f(x)p(x)dx + \int_a^b g(x)p(x)dx = E[f] + E[g]$  [n.d.c.]

$$E[(x - \mu)^2] \equiv \sigma^2 \equiv \begin{cases} \sum_{x_i} (x_i - \mu)^2 P(x_i) \\ \int_a^b (x - \mu)^2 p(x) dx \end{cases}$$

Si definisce deviazione standard ( $\sigma$ ), la radice quadrata della varianza.

$$\sqrt{E[(x - \mu)^2]} \equiv \sigma$$

Si può vedere che

$$\sigma^2 = E[x^2] - \mu^2.$$

Infatti:

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

c) Discussione e cenno al Teorema di Tschebyscheff: Vogliamo chiarire lo scopo per cui si fanno questi discorsi.

Quando si hanno misure i cui risultati fluttuano, il risultato della misura è una variabile casuale funzione delle condizioni sperimentali. 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 un grosso teorema di statistica dovuto a Tschebyscheff:

Sia  $x$  una variabile casuale con una distribuzione tale che  $\mu, \sigma^2$  esistono e sono finiti. Detto  $k$  un numero positivo

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

Cioè è poco probabile avere grandi deviazioni dalla media; non solo, ma dice anche assolutamente in generale qual è il limite superiore a questa probabilità.

d) Cenno alla teoria dei campioni: Vediamo ora dal punto di vista logico quali sono le ipotesi che si fanno quando si applicano i concetti statistici fino ad ora trattati ai risultati delle misure.

- 1) L'insieme delle misure sperimentali 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 fattibili: "distribuzione generatrice" (parent distribution).
- 3) Dal campione dobbiamo approssimare il meglio possibile la media e la varianza della distribuzione genitrice.

Nota: quasi tutti i test postulano l'esistenza di un "valore vero" della grandezza misurata. Questo postulato non solo non è necessario, ma è sbagliato: il "valore vero" di una grandezza fisica non esiste.

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 delle misure 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.

#### Esercizio

Supponiamo di lanciare due dadi. Vogliamo trovare la funzione di distribuzione per la variabile  $S$  somma delle uscite dei due dadi. Esaminiamo perciò la tabella (1).

Si vede subito che la funzione  $f(S) = S - 1$  dà il numero dei casi favorevoli per  $S$  che varia da 2 a 7. Per  $S$  che varia da 8 a 12 si osserva che  $S + n^\circ$  casi favorevoli = 13. Quindi  $f(S) = 13 - S$  per  $S$  variabile da 8 a 12. Per avere la probabilità di ottenere un certo valore  $S$  bisogna dividere il numero dei casi favorevoli per il numero dei casi possibili, che sono in tutto 36. Quindi

$$P(S) = \begin{cases} \frac{S-1}{36} & 2 \leq S \leq 7 \\ \frac{13-S}{36} & 7 < S \leq 12 \end{cases}$$

Si osservi che il fattore  $\frac{1}{36}$  è la costante di normalizzazione per la funzione  $f(S)$ . Infatti se cerchiamo il fattore moltiplicativo  $A$  tale che  $\sum_S A f(S) = 1$  si ottiene:

$$\begin{aligned} A \sum_{S=2}^7 (S-1) + A \sum_{S=8}^{12} (13-S) &= 1 \\ A(1+2+3+4+5+6+5+4+3+2+1) &= 1 \\ A \cdot 36 &= 1 \\ A &= \frac{1}{36} \end{aligned}$$

Determinare il valor medio di  $S$  e la varianza di  $S$ .<sup>5</sup>

<sup>5</sup> Svolgimento:

$$\begin{aligned} \mu &= \sum_{S=2}^7 S \cdot \frac{S-1}{36} + \sum_{S=8}^{12} S \cdot \frac{13-S}{36} = \\ &= \frac{1}{36} \left\{ \sum_{S=2}^7 (S^2 - S) + \sum_{S=8}^{12} (13S - S^2) \right\} = \frac{1}{36} \left\{ \sum_2^7 S^2 - \sum_2^7 S + 13 \sum_8^{12} S - \sum_8^{12} S^2 \right\} = \\ &= \frac{1}{36} \left\{ \left( \sum_1^7 S^2 - 1 \right) - \left( \sum_1^{12} S^2 - \sum_1^7 S^2 \right) - \left( \sum_1^7 S - 1 \right) + 13 \left( \sum_1^{12} S - \sum_1^7 S \right) \right\} = \\ &= \frac{1}{36} \left\{ 2 \sum_1^7 S^2 - 1 - \sum_1^{12} S^2 + 13 \sum_1^{12} S - 14 \sum_1^7 S + 1 \right\} = \\ &= \frac{1}{36} \left\{ 2 \cdot \frac{7(7+1)(2 \cdot 7+1)}{6} - \frac{12(12+1)(2 \cdot 12+1)}{6} + 13 \cdot \frac{12(12+1)}{2} - 14 \cdot \frac{7(7+1)}{2} \right\} = \end{aligned}$$

---


$$\begin{aligned}
&= \frac{7 \cdot 8 \cdot 5}{36} - \frac{2 \cdot 13 \cdot 25}{36} + \frac{13 \cdot 6 \cdot 13}{36} - \frac{7 \cdot 7 \cdot 8}{36} = \frac{280 - 650 + 1014 - 392}{36} = \frac{252}{36} = 7 \\
\sigma^2 &= \sum_S (S - \mu)^2 P(S) = \sum_2^7 (S - 7)^2 \frac{S-1}{36} + \sum_8^{12} (S - 7)^2 \frac{13-S}{36} = \\
&= \frac{1}{36} (5^2 \cdot 1 + 4^2 \cdot 2 + 3^2 \cdot 3 + 2^2 \cdot 4 + 1^1 \cdot 5 + 1^2 \cdot 5 + 2^2 \cdot 4 + 3^2 \cdot 3 + 4^2 \cdot 2 + 5^2 \cdot 1) = \\
&= \frac{1}{36} (25 + 32 + 27 + 16 + 5 + 5 + 16 + 27 + 32 + 25) = \frac{210}{36} = \frac{35}{6} = 6 - \frac{1}{6} \text{ [n.d.c.]}.
\end{aligned}$$

---

## **CAPITOLO II:**

### **PROPAGAZIONE DEGLI ERRORI - Parte I**

#### **1 Errore massimo**

Si supponga di voler calcolare il volume  $V$  di un cilindro di raggio  $r$  ed altezza  $H$ ;  $r$  ed  $H$  sono quantità misurate con una incertezza  $\Delta r$  e  $\Delta H$  rispettivamente. Si pone il problema di vedere come tali incertezze influenzano il risultato  $V$ , cioè come  $\Delta r$  e  $\Delta H$  si propagano per dare  $\Delta V$ .

Vediamo quindi l'effetto delle quattro operazioni sulla propagazione degli errori:

a) Somma  $S = a + b$

$a$  e  $b$  sono due quantità con incertezza  $\Delta a$  e  $\Delta b$  rispettivamente, quindi i valori che  $a$  e  $b$  possono assumere sono entro  $a \pm \Delta a$ ,  $b \pm \Delta b$ . Costruiamo la somma  $S$  mettendoci nel caso peggiore in cui sia  $\Delta a$  che  $\Delta b$  contribuiscono nello stesso verso:

$$(a + \Delta a) + (b + \Delta b) = a + b + (\Delta a + \Delta b)$$

$$(a - \Delta a) + (b - \Delta b) = a + b - (\Delta a + \Delta b)$$

Ne segue che l'errore massimo che possiamo fare su  $S$  è:

$$\Delta S = \Delta a + \Delta b$$

b) Differenza  $D = a - b$

$$\left. \begin{array}{l} a \pm \Delta a \\ b \pm \Delta b \end{array} \right\} \Rightarrow D \pm \Delta D$$

Per determinare  $\Delta D$  si procede come per la somma; il caso peggiore è:

$$(a + \Delta a) - (b + \Delta b) = (a - b) + (\Delta a + \Delta b)$$

$$(a - \Delta a) - (b - \Delta b) = (a - b) - (\Delta a + \Delta b)$$

ne segue che:

$$\Delta D = \Delta a + \Delta b$$

e non  $\Delta a - \Delta b$ .

c) Prodotto       $P = ab$

$$\left. \begin{array}{l} a \pm \Delta a \\ b \pm \Delta b \end{array} \right\} \Rightarrow P \pm \Delta P$$

Caso peggiore:

$$(a + \Delta a)(b + \Delta b) = ab + a\Delta b + b\Delta a + \Delta a\Delta b$$

$$(a - \Delta a)(b - \Delta b) = ab - a\Delta b - b\Delta a + \Delta a\Delta b$$

Si trascura il prodotto  $\Delta a\Delta b$  rispetto agli altri termini perché è molto minore di essi.

Ne segue che:

$$\Delta P = a\Delta b + b\Delta a$$

Esempi:

$$1c) a = b^2 \quad P = a^2 \quad \Delta P = 2a\Delta a$$

$$2c) P = a^3 = a^2 \cdot a \quad \Delta P = a^2\Delta a + a\Delta(a^2) = a^2\Delta a + 2a^2\Delta a = 3a^2\Delta a$$

$$3c) P = a^n \quad \Delta P = na^{n-1}\Delta a$$

$$4c) P = 2 \cdot a \quad \Delta P = 2\Delta a$$

$$d) \text{ Quoziente } \quad Q = \frac{a}{b}$$

$$\left. \begin{array}{l} a \pm \Delta a \\ b \pm \Delta b \end{array} \right\} \Rightarrow Q \pm \Delta Q$$

Caso peggiore:

$$\frac{a + \Delta a}{b - \Delta b} \left[ = \frac{\frac{a}{b} + \frac{\Delta a}{b}}{1 - \frac{\Delta b}{b}} \simeq \left( \frac{a}{b} + \frac{\Delta a}{b} \right) \left( 1 + \frac{\Delta b}{b} \right) \right] \simeq \frac{a}{b} + \frac{\Delta a}{b} + \frac{a\Delta b}{b^2} = \frac{a}{b} + \frac{a\Delta b + b\Delta a}{b^2}$$

$$\frac{a - \Delta a}{b + \Delta b} \left[ = \frac{\frac{a}{b} - \frac{\Delta a}{b}}{1 + \frac{\Delta b}{b}} \simeq \left( \frac{a}{b} - \frac{\Delta a}{b} \right) \left( 1 - \frac{\Delta b}{b} \right) \right] \simeq \frac{a}{b} - \frac{\Delta a}{b} - \frac{a\Delta b}{b^2} = \frac{a}{b} - \frac{a\Delta b + b\Delta a}{b^2}$$

sono trascurati i termini del tipo  $\Delta a\Delta b$ .

Ne segue che:

$$\Delta Q = \frac{a\Delta b + b\Delta a}{b^2}$$

In generale se una grandezza  $G$  è funzione di una grandezza  $a$ , cioè  $G = f(a)$ , l'incertezza su  $a$  si ripercuote su  $G$  in questo modo:

$$\Delta G = \left| \frac{df}{da} \right| \Delta a$$

Se le grandezze da cui dipende  $G$  sono più di una, cioè  $G = f(a, b, c, \dots)$ , l'errore massimo su  $G$  è dato da:

$$\Delta G = \left| \frac{\partial f}{\partial a} \right| \Delta a + \left| \frac{\partial f}{\partial b} \right| \Delta b + \left| \frac{\partial f}{\partial c} \right| \Delta c + \dots$$

Purché i termini che contengono  $\Delta a \cdot \Delta b \dots$ ,  $(\Delta a)^2$ ,  $(\Delta b)^2 \dots$  siano piccoli rispetto ai termini che contengono  $\Delta a$ ,  $\Delta b$  etc.

Esempio: ritorniamo al problema iniziale e calcoliamo l'errore sul volume  $V$  del cilindro:

$$V = \pi r^2 H, \quad r \pm \Delta r, \quad H \pm \Delta H$$

$$\Delta V = \pi r^2 \Delta H + 2\pi r H \cdot \Delta r$$

## 2 Errore relativo

Quasi sempre quando si misura qualche grandezza  $x$  è fisicamente più significativo il rapporto  $\frac{\Delta x}{x}$  piuttosto che la incertezza  $\Delta x$ .

$$\frac{\Delta x}{x} \quad \text{si chiama errore relativo.}$$

Esempio: si consideri l'errore si 1 mm su a) 1 metro, b) 10 mm.

È più grave il secondo, infatti

$$\text{a) } \frac{\Delta x}{x} = \frac{1\text{mm}}{10^3\text{mm}} = 10^{-3} = 0.1\%$$

$$\text{b) } \frac{\Delta x}{x} = \frac{1\text{mm}}{10\text{mm}} = 10^{-1} = 10\%$$

Propagazione degli errori in un caso abbastanza comune: sia data la grandezza

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

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

$$\frac{\Delta G}{G} = \left| n \frac{\Delta a}{a} \right| + \left| m \frac{\Delta b}{b} \right| + \left| p \frac{\Delta c}{c} \right| + \dots$$

gli errori relativi si sommano con un peso che è uguale al modulo dell'esponente. Da questo si deduce la seguente informazione importante: è bene che le grandezze che entrano in una misura siano valutate tutte con errori relativi dello stesso ordine di grandezza.

### 3 Nozione di errore sistematico

Gli errori sistematici sono quelli che derivano da errori di taratura degli strumenti o dell'apparato sperimentale o dalla mancanza di imparzialità dell'osservatore.

Esempio 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, il risultato che si legge non ha l'errore di  $\frac{1}{5}$  di secondo, ma l'errore può essere molto più grande. In ogni caso non è facile stabilire se ci sono errori sistematici a meno di non disporre di un orologio campione di cui fidarsi.

Esempio 2): Si supponga di voler verificare la legge di Boyle misurando pressione e volume dell'aria che si trova nel ramo chiuso di un tubo ad U contenente Hg. Le misure di pressione non sono torr, perché la temperatura non è a  $0^\circ$ . Quindi le stime della costante  $=PV$  sono tutte consistenti tra di loro, ma non servono a calcolare la costante dei gas  $R$ .

### 4 Nozione di precisione ed accuratezza

La precisione delle misure è la consistenza interna dell'insieme delle misure. Indica di quanto si discostano l'uno dall'altro i risultati sperimentali.

La precisione di una misura viene identificata con l'errore relativo soltanto quando l'errore sistematico è trascurabile (o si ritiene che lo sia). Allora il contributo dell'incertezza  $\Delta x$  di una grandezza  $x$  viene solo dalle fluttuazioni dei risultati o dalla risoluzione degli strumenti.

Nel caso in cui i risultati sono tutti uguali perché la misura è limitata dallo strumento, la precisione non è ben definita. Allora si assume come precisione la risoluzione dello strumento (es.  $\frac{1}{2}$  divisione di una scala graduata o se c'è un nonio come nel calibro  $\frac{1}{10}$  o  $\frac{1}{20}$  di divisione). Si noti che a rigore questa è la risoluzione dello strumento, non la precisione delle misure.

Accuratezza = valutazione dell'errore sistematico: una misura è tanto più accurata quanto più il risultato è vicino a quello che si otterrebbe in assenza di errori sistematici.

### 5 Cifre significative e convenzioni di scrittura dei risultati delle misure

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 (indicata nell'esempio con +).
2. Se non c'è la virgola decimale, la cifra più a destra diversa da zero è la



meno significativa (indicata nell'esempio con -). Questa è la convenzione più diffusa, però non è accettata da tutti. Per cui leggendo un numero come 3200 può rimanere ambiguità se le cifre significative sono due (come da questa convenzione) o quattro.

3. Se c'è la virgola decimale la cifra più a destra è la meno significativa anche se è zero.

4. Tutte le cifre tra la più significativa e la meno significativa comprese sono cifre significative. Esempio:

3215	3215.4	3200	0.032	3200.0	18.00	0.180
+	-	+	-	+-	+	-
+	-	+	-	+-	+	-

Per sapere quante cifre significative devono essere riportate nel risultato di un esperimento bisogna valutare gli errori commessi. Si riportano allora tutte le cifre fino alla prima influenzata dall'errore inclusa.

Quando le cifre non significative vengono tagliate da un numero, le rimanenti cifre devono essere arrotondate per una migliore accuratezza. Per arrotondare un numero lo si tronca fino al numero di cifre significative desiderato e le cifre in più vengono trattate come frazione decimale. Allora:

1. Se la frazione è più grande di  $\frac{1}{2}$ , si incrementa di una unità l'ultima cifra significativa.

2. Se la frazione è meno di  $\frac{1}{2}$  non si incrementa,

3. Se la frazione è uguale a  $\frac{1}{2}$  si incrementa l'ultima cifra soltanto se è dispari.

In questo modo il valore del risultato finale è sempre entro la metà dell'ultima cifra significativa del numero originale. La regola 3. viene usata per evitare un sistematico aumento del valore quando si fanno letture con una risoluzione di mezza divisione.

Quando i risultati di misure vengono usati per calcolare altre grandezze, bisogna fare attenzione agli errori che si possono commettere nei calcoli numerici. Per chiarire questo consideriamo i seguenti esempi

a) Prodotto  $P = a \cdot b$

Esempio 1):  $a = 1.25432$

$b = 9.35$

se "a" viene arrotondato ad 1.25 e "b" a 9.4 quale errore massimo si commette nel prodotto:

$$P = a \cdot b = 11.727892$$

$$P^* = a^* b^* = 1.25 \cdot 9.4 = 11.750$$

$$P^* - P = 0.022$$

L'errore è sulla 4<sup>a</sup> cifra significativa.

Quindi  $P^* = 11.75 \pm 0.02$  od anche  $11.750 \pm 0.022$ . Se non si riporta l'errore si scrive  $P^* = 11.7$ , perché solo tre cifre sono sicure. Con questo si intende  $P = 11.75 \div 11.65$ .

Esercizio 1:<sup>6</sup>

Verificare che per l'es. 1 si sarebbe ottenuto lo stesso errore usando la propagazione degli errori in un prodotto:

$$\Delta P = a\Delta P + b\Delta a \quad \text{dove} \quad \Delta a = 0.005 \quad \Delta b = 0.05.$$

Esempio 2):  $a = 3.658$

$$b = 24.763$$

Con quante cifre significative è ragionevole scrivere il prodotto.

Il risultato dell'operazione  $3.658 \times 24.763$  per come viene fatta è: 90.583054.

Per come si conoscono i numeri il risultato potrebbe avere un valore compreso tra

$$3.6585 \times 24.7635 = 90.59726475$$

e

$$3.6575 \times 24.7625 = 90.56884375$$

Allora il prodotto va scritto con 4 cifre significative.

L'errore è sulla 4<sup>a</sup> cifra significativa.

Verificarlo per esercizio<sup>7</sup> sfruttando la legge di propagazione degli errori in un prodotto con  $\Delta a = 5 \times 10^{-4} = \Delta b$ .

Tanto valeva fare il prodotto così:  $3.658 \times 24.76$ .

Esercizio:<sup>8</sup> nel prodotto  $a \cdot b$  con  $a = 1.25432$   $b = 9.3$ , a quante cifre significative conviene arrotondare "a". Suggerimento: eseguire i prodotti arrotondando "a" successivamente a 5, 4, 3, 2, ... cifre significative.

Una certa attenzione va fatta quando i prodotti sono più di due. Per esempio:

$$3.658 \times 24.763 \times 1.4345 \times 72.43$$

<sup>6</sup> Svolgimento: Per come si conoscono i numeri di partenza gli errori commessi negli arrotondamenti sono  $\Delta a = a^* - a = -0.00432$  e  $\Delta b = b^* - b = 0.05$ . Allora  $\Delta P = a\Delta b + b\Delta a = 1.25432 \cdot 0.005 + 9.35 \cdot (-0.00432) = 0.0627160 - 0.0403920 = 0.0223240$ . Si è così riottenuto  $P^* = 11.750$  e  $\Delta P = 0.022$ .

Se i numeri di partenza non sono noti con esattezza si può valutare l'errore massimo. Il risultato potrebbe avere un valore compreso tra  $1.255 \cdot 9.45 = 11.85975$  e  $1.245 \cdot 9.35 = 11.64075$ . In questo caso il prodotto va scritto con tre cifre significative. L'errore è sulla 3<sup>a</sup> cifra significativa  $P = 11.7 \pm 0.1$ .

Otteniamo lo stesso risultato usando la propagazione degli errori:  $\Delta P = a\Delta b + b\Delta a$  dove questa volta  $\Delta a = 0.005$  e  $\Delta b = 0.05$ . Infatti  $\Delta P = 1.25 \cdot 0.05 + 9.4 \cdot 0.005 = 0.0625 + 0.0470 = 0.1095$ . L'errore è sulla 3<sup>a</sup> cifra significativa. Riotteniamo  $P = 11.7 \pm 0.1$  od anche  $11.75 \pm 0.11$  [n.d.c.].

<sup>7</sup> Infatti  $\Delta P = a\Delta b + b\Delta a = (3.658 + 24.763) \cdot 5 \cdot 10^{-4} = 28.421 \cdot 5 \cdot 10^{-4} = 142.105 \cdot 10^{-4} = 0.0142105$  [n.d.c.].

<sup>8</sup> Svolgimento: L'errore sul prodotto è sulla 4<sup>a</sup> cifra significativa. Infatti  $\Delta P = a\Delta b + b\Delta a = 1.25432 \cdot 0.05 + 9.3 \cdot 5 \cdot 10^{-6} = 0.062716 + 0.0000465 = 0.0627625$ . Eseguiamo ora i prodotti arrotondando "a" successivamente a 5, 4, 3, 2, ... cifre significative:

Mostrare che potrebbe essere dannoso il procedimento di arrotondamento prima della moltiplicazione ciascun fattore a quattro cifre.<sup>9</sup> Conviene partire dai fattori con più cifre significative o da quelli con meno?<sup>10</sup>

b) Quoziente  $Q = \frac{a}{b}$   
 $a = 10.348$   
 $b = 6.24$

Con quante cifre significative deve essere preso il quoziente.  
 Suggerimento: confrontare<sup>11</sup>

$$\begin{array}{r} 10.348 : 6.24 \\ 10.3485 : 6.235 \\ 10.3475 : 6.245 \end{array}$$

Verificare<sup>12</sup> che  $\Delta Q = \frac{a\Delta b + b\Delta a}{b^2}$  c) Somma  $S = a + b$   
 $a = 56.434$   
 $b = 251.37$

Conviene scrivere i numeri in questo modo:

$$a = 0.056434 \times 10^3 \quad b = 0.25137 \times 10^3$$

Facendo la somma:

$$\begin{array}{r} 0.056434 \times 10^3 \\ 0.25137 \times 10^3 \\ \hline 0.308804 \end{array}$$

---

1.25432	$\times 9.3 = 11.665176$
1.2543	$\times 9.3 = 11.66499$
1.254	$\times 9.3 = 11.6622$
1.25	$\times 9.3 = 11.625$
1.2	$\times 9.3 = 11.16$

Poiché l'errore è sulla 4<sup>a</sup> cifra significativa conviene arrotondare "a" a 4 cifre significative e scrivere  $1.254 \cdot 9.3 = 11.66$  [n.d.c.].

<sup>9</sup> In questo caso l'errore è sulla 4<sup>a</sup> cifra significativa. Il procedimento di arrotondare prima della moltiplicazione ciascun fattore a quattro cifre è dannoso in quanto l'errore sul prodotto si sposterebbe sulla 3<sup>a</sup> [n.d.c.].

<sup>10</sup> Conviene partire dai fattori con più cifre significative e tenere nei prodotti intermedi una cifra in più e infine arrotondare a quattro cifre significative il risultato finale [n.d.c.].

<sup>11</sup>

$$\begin{array}{r} 10.348 : 6.24 = 1.6583333... \\ 10.3485 : 6.235 = 1.6597433... \\ 10.3475 : 6.245 = 1.6569255... \end{array}$$

Si vede che il quoziente deve essere preso con quattro cifre significative [n.d.c.].

<sup>12</sup>  $\Delta Q = \frac{10.348 \cdot 5 \cdot 10^{-3} + 6.25 \cdot 5 \cdot 10^{-4}}{(6.24)^2} = \frac{0.05174 + 0.00312}{38.9376} = \frac{0.05486}{38.9376} = 0.0014089$  [n.d.c.]

Risulta subito evidente che la 6<sup>a</sup> cifra è non significativa, come si ottiene pure da

$$\Delta S = \Delta a + \Delta b$$

$$a = 5 \times 10^{-6}$$

$$b = 5 \times 10^{-5}$$

$$\Delta S \simeq \Delta b$$

d) Differenza  $D = a - b$

$$a = 86.67$$

$$b = 7.0228$$

Convieni scrivere i numeri in questo modo:

$$a = 0.8667 \times 10^2 \quad b = 0.070228 \times 10^2$$

$$10^{-2}(a - b) = 0.8667 \quad -$$

$$\underline{0.070228}$$

$$0.796472$$

Il risultato è  $D = 0.7965 \times 10^2$ . La 5<sup>a</sup> e la 6<sup>a</sup> cifra non sono significative.

Ancora  $a = 0.738 \quad b = 0.736 \quad D = a - b = 0.002$  una sola cifra significativa!

La differenza di due numeri quasi uguali ha una precisione molto inferiore a quella dei due numeri stessi.

Ne segue che una certa attenzione va fatta nell'arrotondare numeri di cui poi si deve fare la differenza.

Esercizi:

1) A quante cifre significative va arrotondato  $\pi = 3.14159265$  perché le potenze  $\pi^2$ ,  $\pi^3$ ,  $\pi^4$ ,  $\pi^5$ ,  $\pi^6$  siano corrette all'1%, all'1‰.<sup>13</sup>

2)  $R = \frac{\sqrt{2} + \pi^2}{\sqrt{10} - \pi}$  con quante cifre significative bisogna prendere  $\sqrt{2}$ ,  $\sqrt{10}$ ,  $\pi$

<sup>13</sup> Svolgimento: Consideriamo il caso generale  $\pi^n$ :

$$\frac{\Delta \pi^n}{\pi^n} = n \frac{\Delta \pi}{\pi}$$

Se l'errore relativo è dell'1% si ha

$$n \frac{\Delta \pi}{\pi} = 10^{-2}$$

$$\Delta \pi = \frac{\pi}{n} \cdot 10^{-2}$$

e troviamo:

$n = 2$	0.01570796325...	3 <sup>a</sup> cifra
$n = 3$	0.0104719755...	"
$n = 4$	0.007853981625...	4 <sup>a</sup> cifra
$n = 5$	0.0062831853...	"
$n = 6$	0.00523598775...	"

Se l'errore relativo è dell'1‰ si ha

$$\Delta \pi = \frac{\pi}{n} \cdot 10^{-3}$$

e troviamo:

affinché  $\frac{\Delta R}{R} = 1\%, 1\%$ .<sup>14</sup>

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

Quando in una operazione sono implicate anche costanti che a priori hanno infinite cifre significative, il numero di cifre significative al quale si arrotondano tali costanti deve essere tale da rendere trascurabili gli errori di calcolo

---

$n = 2$	0.00157079...	4 <sup>a</sup> cifra
$n = 3$	0.00104719...	"
$n = 4$	0.00078539...	5 <sup>a</sup> cifra
$n = 5$	0.00062831...	"
$n = 6$	0.00052359...	"

[n.d.c.].

$$\begin{aligned}
 \text{14 Svolgimento: } \frac{\Delta R}{R} &= \frac{(\Delta\sqrt{2}+\Delta\pi^2)(\sqrt{10}-\pi)+(\Delta\sqrt{10}+\Delta\pi)(\sqrt{2}+\pi^2)}{(\sqrt{10}-\pi)^2} \cdot \frac{\sqrt{10}-\pi}{\sqrt{2}+\pi^2} = \\
 &= \frac{\Delta\sqrt{2}+2\pi\Delta\pi}{\sqrt{2}+\pi^2} + \frac{\Delta\sqrt{10}+\Delta\pi}{\sqrt{10}-\pi} = \\
 &= \frac{\Delta\sqrt{2}}{\sqrt{2}+\pi^2} + \left( \frac{2\pi}{\sqrt{2}+\pi^2} + \frac{1}{\sqrt{10}-\pi} \right) \Delta\pi + \frac{\Delta\sqrt{10}}{\sqrt{10}-\pi}
 \end{aligned}$$

Con i valori numerici

$$\sqrt{2} = 1.4142135\dots$$

$$\sqrt{10} = 3.1622776\dots$$

$$\pi = 3.1415926\dots$$

si vede che  $\frac{2\pi}{\sqrt{2}+\pi^2}$  è trascurabile rispetto a  $\frac{1}{\sqrt{10}-\pi}$  perciò

$$\frac{\Delta R}{R} \simeq \frac{\Delta\sqrt{2}}{\sqrt{2}+\pi^2} + \frac{\Delta\pi}{\sqrt{10}-\pi} + \frac{\Delta\sqrt{10}}{\sqrt{10}-\pi}$$

Affinché  $\frac{\Delta R}{R} = 10^{-2}$

$$\frac{\Delta\sqrt{2}}{\sqrt{2}+\pi^2} \simeq \frac{\Delta\pi}{\sqrt{10}-\pi} \simeq \frac{\Delta\sqrt{10}}{\sqrt{10}-\pi} \simeq 10^{-2}$$

$\Delta\sqrt{2}$	$\simeq (\sqrt{2} + \pi^2) \cdot 10^{-2}$	$\simeq 0.1128\dots$	2 cifre
$\Delta\pi$	$\simeq (\sqrt{10} - \pi) \cdot 10^{-2}$	$\simeq 0.0002\dots$	5 cifre
$\Delta\sqrt{10}$	$\simeq \Delta\pi$	$\simeq 0.0002\dots$	"

Affinché  $\frac{\Delta R}{R} = 10^{-3}$

$\Delta\sqrt{2}$	$\simeq (\sqrt{2} + \pi^2) \cdot 10^{-3}$	$\simeq 0.01128\dots$	3 cifre
$\Delta\pi$	$\simeq (\sqrt{10} - \pi) \cdot 10^{-3}$	$\simeq 0.00002\dots$	6 cifre
$\Delta\sqrt{10}$	$\simeq \Delta\pi$	$\simeq 0.00002\dots$	"

[n.d.c.].

numerico rispetto agli errori di misura. (Vedi esercizi precedenti).

---

## CAPITOLO III: DISTRIBUZIONI

### 1 Distribuzione binomiale

Consideriamo  $n$  possibili eventi  $E$  identici e indipendenti ciascuno dei quali si può realizzare in due modi  $e_1, e_2$ , con probabilità  $P(e_1) = p$  e  $P(e_2) = q = 1 - p$ .

La probabilità di ottenere  $k$  volte l'evento  $e_1$  + data dalla distribuzione binomiale\* \*\*

$$P(n, k) = \binom{n}{k} p^k q^{n-k}$$

Infatti supponiamo che l'evento  $E$  sia il risultato del lancio di una moneta. I modi nei quali l'evento può accadere sono due: testa o croce.

Sia  $P(T) = p$  e  $P(C) = q$ .

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

$$p^k q^{n-k} \quad (\text{legge della moltiplicazione})$$

però avere  $k$  Teste in questo ordine e  $n - k$  Croci è uno dei tanti possibili modi nei quali si possono avere  $k$  volte T e  $(n - k)$  volte C. Tutti i possibili modi sono le combinazioni di  $n$  elementi a  $k$  a  $k$ . Quindi segue che

$$P(n, k) = \binom{n}{k} p^k q^{n-k}$$

---

\* Si noti che questo è un termine dello sviluppo di un binomio elevato alla potenza  $n$ :

$$(p + q)^n = \sum_{k=0}^n \binom{n}{k} p^k q^{n-k}$$

\*\*  $\binom{n}{k}$  = combinazione di  $n$  elementi a  $k$  a  $k$  =  $\frac{n(n-1)(n-2)\dots(n-k+1)}{k!}$   
 $k! = k(k-1)(k-2)\dots 2 \cdot 1$

La distribuzione  $P(n, k)$  è normalizzata, cioè  $\sum_{k=0}^n P(n, k) = 1$ .

$$\sum_k P(n, k) = \sum_k \binom{n}{k} p^k q^{n-k} = (p + q)^n$$

ma  $p + q = 1$ , quindi  $\sum_k P(n, k) = 1$ .

La media  $\mu$  e la varianza  $\sigma^2$  sono:

$$\begin{aligned} \mu &= np & \sigma^2 &= npq \\ \sigma &= \sqrt{npq} \end{aligned}$$

Calcolo della media:

$$\begin{aligned} \mu &= \sum_{k=0}^n k P(n, k) = \sum_{k=1}^n k \binom{n}{k} p^k q^{n-k} = \sum_{k=1}^n \frac{k \cdot n \cdot (n-1) \dots (n-k+1)}{k \cdot (k-1) \dots 2 \cdot 1} p^k q^{n-k} = \\ &= np \sum_{k=1}^n \frac{(n-1) \cdot (n-2) \dots (n-k+1)}{(k-1) \cdot (k-2) \dots 2 \cdot 1} p^{k-1} q^{n-k} \end{aligned}$$

si ponga  $s = k - 1$      $m = n - 1$ . Si ottiene

$$\mu = np \sum_{s=0}^m \frac{m \cdot (m-1) \dots (m-s+1)}{s \cdot (s-1) \dots 2 \cdot 1} p^s q^{m-s} = np$$

Calcolo della varianza:

$$\begin{aligned} \sigma^2 &= \sum_{k=0}^n (k - \mu)^2 P(n, k) = \sum_{k=1}^n k^2 P(n, k) - \mu^2 = \\ \sum_{k=1}^n k^2 P(n, k) &= \sum_{k=1}^n \frac{k^2 \cdot n \cdot (n-1) \dots (n-k+1)}{k \cdot (k-1) \dots 2 \cdot 1} p^k q^{n-k} = \\ &= \sum_{k=1}^n \frac{k \cdot n \cdot (n-1) \dots (n-k+1)}{(k-1) \cdot (k-2) \dots 2 \cdot 1} p^k q^{n-k} = \\ &= \sum_{k=2}^n (k-1) \frac{n \cdot (n-1) \dots (n-k+1)}{(k-1) \cdot (k-2) \dots 2 \cdot 1} p^k q^{n-k} + \sum_{k=1}^n \frac{n \cdot (n-1) \dots (n-k+1)}{(k-1) \cdot (k-2) \dots 2 \cdot 1} p^k q^{n-k} = \\ &= \sum_{k=2}^n \frac{n \cdot (n-1) \dots (n-k+1)}{(k-2) \cdot (k-3) \dots 2 \cdot 1} p^k q^{n-k} + \mu = \\ &= n(n-1)p^2 \sum_{k=2}^n \frac{(n-2) \cdot (n-3) \dots (n-k+1)}{(k-2) \cdot (k-3) \dots 2 \cdot 1} p^{k-2} q^{n-k} + \mu \end{aligned}$$

si ponga  $s = k - 2$      $m = n - 2$ . Si ottiene



$$\begin{aligned}
\sigma^2 &= n(n-1)p^2 \sum_{s=0}^m \frac{m \cdot (m-1) \dots (m-s+1)}{s \cdot (s-1) \dots 2 \cdot 1} p^s q^{m-s} + \mu - \mu^2 = \\
&= n(n-1)p^2 + \mu - \mu^2 = n(n-1)p^2 + np - n^2p^2 = n^2p^2 + np - np^2 - n^2p^2 = \\
&= np - np^2 = np(1-p) = npq.
\end{aligned}$$

Esempio:

Sia  $E$  il lancio di un dado. Sia  $n = 4$  il numero di volte che si lancia un dado. La probabilità che un numero, per es. il 5 esca 0 volte, 1 volta, ..., 4 volte è data dalla distribuzione binomiale con  $n = 4$  e  $k = 0, 1, 2, 3, 4$  rispettivamente. Calcoliamo esplicitamente  $P(4, k)$  per ogni valore di  $k$ :

$$\begin{aligned}
P(4, 0) &= \binom{4}{0} \left(\frac{5}{6}\right)^4 = \frac{5^4}{6^4} = 625/6^4 \\
P(4, 1) &= \binom{4}{1} \frac{1}{6} \left(\frac{5}{6}\right)^3 = 4 \frac{5^3}{6^4} = 500/6^4 \\
P(4, 2) &= \binom{4}{2} \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^2 = \frac{4 \cdot 3 \cdot 5^2}{2 \cdot 6^4} = 150/6^4 \\
P(4, 3) &= \binom{4}{3} \left(\frac{1}{6}\right)^3 \frac{5}{6} = \frac{4 \cdot 3 \cdot 2}{3 \cdot 2} 5 \frac{1}{6^4} = 20/6^4 \\
P(4, 4) &= \binom{4}{4} \left(\frac{1}{6}\right)^4 = 1/6^4 \\
\mu &= np = 4 \frac{1}{6} = 0.667 \\
\sigma &= \sqrt{4 \frac{1}{6} \frac{5}{6}} \simeq 0.7
\end{aligned}$$

Costruiamo l'istogramma  $P(4, k)$  fig.3

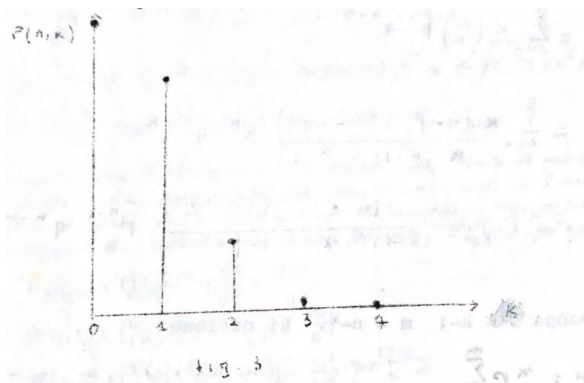


fig. 3.

Si propone di costruire sperimentalmente l'istogramma  $P(4, k)$  operando in questo modo.

- 1) Fare  $N$  prove, ogni prova sia un lancio di 4 dadi.
- 2) Contare quanti "cinque" escono in una prova.
- 3) Contare quante prove contengono  $k$  "cinque".
- 4) Dividere per il numero di prove, in modo da ottenere la frequenza relativa, che è una stima della probabilità.
- 5) Fare l'istogramma delle frequenze relative.

## 2 Distribuzione di Poisson

La probabilità che si verifichino  $k$  eventi in una situazione in cui in media se ne verificano  $a$  è:

$$P(k) = \frac{a^k}{k!} e^{-a}$$

questa è la distribuzione di Poisson.

Nota: il numero  $e$  è un numero reale definibile in tante maniere, ad es:

$$\lim_{s \rightarrow \infty} \left(1 + \frac{1}{s}\right)^s = e$$

vale circa 2,71828...

Inoltre

$$\lim_{s \rightarrow \infty} \left(1 + \frac{x}{s}\right)^s = e^x$$

Un'altra utile proprietà matematica è la seguente:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = \lim_{N \rightarrow \infty} \sum_{n=0}^N \frac{x^n}{n!}$$

In particolare questa formula dà un metodo per calcolare con la precisione che si vuole  $e$ :

$$e = 1 + \frac{1}{2} + \frac{1}{3!} + \frac{1}{4!} + \dots$$

La distribuzione di Poisson si ottiene come limite della distribuzione binomiale nel caso in cui

$$p \ll 1 \quad n \gg 1 \quad \text{ma} \quad np = \text{cost.} = a$$

allora  $p = \frac{a}{n}$

$$\begin{aligned} \lim_{n \rightarrow \infty} P(n, k) &= \lim_{n \rightarrow \infty} \binom{n}{k} p^k q^{n-k} = \lim_{n \rightarrow \infty} n(n-1) \dots (n-k+1) \frac{\left(\frac{a}{n}\right)^k}{k!} \left(1 - \frac{a}{n}\right)^{n-k} = \\ &= \lim_{n \rightarrow \infty} \frac{a^k}{k!} \frac{n}{n} \frac{n-1}{n} \dots \frac{n-k+1}{n} \frac{\left(1 - \frac{a}{n}\right)^n}{\left(1 - \frac{a}{n}\right)^k} = \frac{a^k}{k!} e^{-a} \end{aligned}$$

Questo significa che ci si deve aspettare una poissoniana tutte le volte che il numero degli eventi possibili è enorme, ma quelli che in media si verificano sono abbastanza pochi. Notare che in generale non è noto il numero  $n$  di eventi possibili, né la probabilità di un singolo evento, ma quello che si conosce e che serve conoscere è il numero medio degli eventi osservati o la sua stima.

Oltre che come limite di una distribuzione binomiale la distribuzione di Poisson si può ottenere in un'altra maniera.

Date le ipotesi:

- a) Il verificarsi di un evento non è influenzato in nessun modo dal verificarsi o meno degli altri eventi
- b) Per piccoli intervalli di tempo la probabilità che si verifichino  $k$  eventi è proporzionale al tempo: si può impostare un'equazione differenziale che permette di ricavare la poissoniana <sup>†</sup>).

L'importanza di questo discorso è la seguente: dalle ipotesi necessarie alla dimostrazione si può notare che la poissoniana ha un campo di applicabilità estremamente vasto <sup>††</sup>).

Ecco alcuni esempi di eventi casuali che possono essere descritti da una poissoniana:

- Decadimento di particelle radioattive: cioè il numero di particelle che decadono in un certo intervallo di tempo.
- 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 biblioteca ogni 10 minuti.
- Il numero di un particolare tipo di automobili che passano per un certo luogo in un intervallo di tempo fissato.

Verifichiamo che la distribuzione è normalizzata:

$$\sum_{\text{casi possibili}} P(k) = \sum_{k=0}^{\infty} \frac{a^k}{k!} e^{-a} = e^{-a} \sum_{k=0}^{\infty} \frac{a^k}{k!} = e^{-a} e^a = 1$$

Calcolo della media:

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

posto  $h = k - 1 \rightarrow k = h + 1$

$$\mu = e^{-a} \sum_{h=0}^{\infty} \frac{a^{h+1}}{h!} = a e^{-a} \sum_{h=0}^{\infty} \frac{a^h}{h!} = a e^{-a} e^a = a$$

Quindi d'ora in avanti scriviamo:

<sup>†</sup> Vedi per es.: P. R. Bevington, Data reduction and error analysis for the physical sciences, McGraw Hill Book Co.

<sup>††</sup> Physics Teacher, 10, 314 (1972)

$$p(k) = \frac{\mu^k}{k!} e^{-\mu}$$

Calcolo della varianza:

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

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

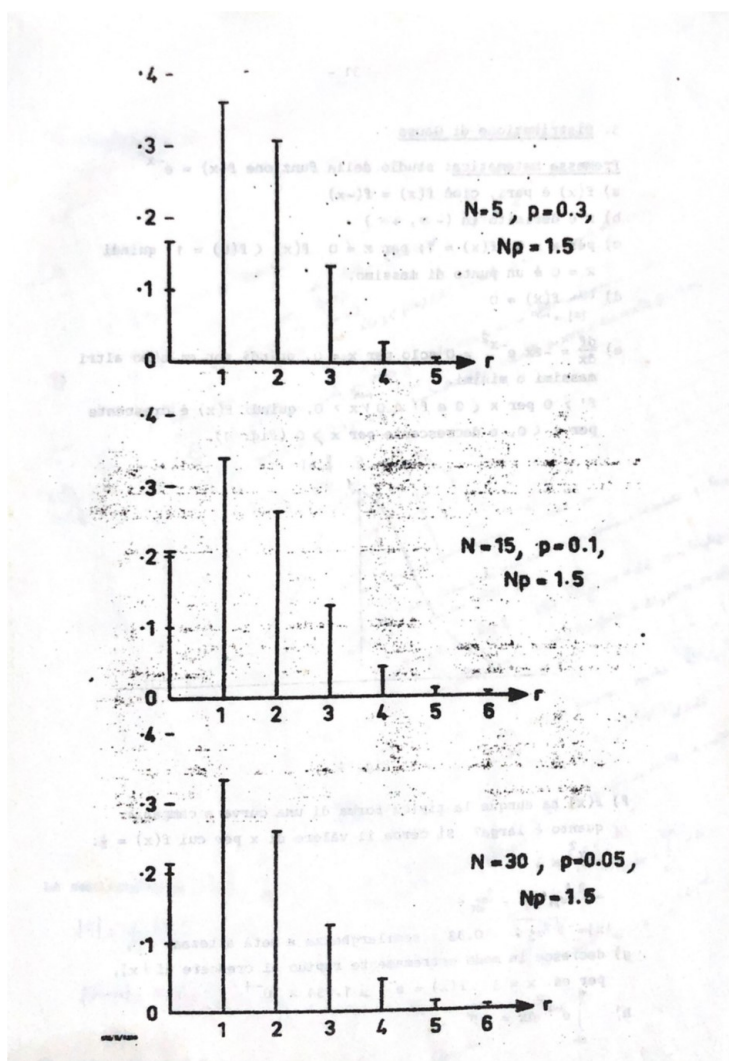
### 3 Distribuzione di Gauss

Premessa matematica: studio della funzione  $f(x) = e^{-x^2}$

- a)  $f(x)$  è pari, cioè  $f(x) = f(-x)$
- b)  $x$  è definito in  $(-\infty, +\infty)$
- c) per  $x = 0$   $f(x) = 1$ ; per  $x \neq 0$   $f(x) < f(0) = 1$  quindi  $x = 0$  è un punto di massimo.
- d)  $\lim_{|x| \rightarrow \infty} f(x) = 0$
- e)  $\frac{df}{dx} = -2xe^{-x^2} = 0$  solo per  $x = 0$ , quindi non ci sono altri minimi e massimi.
- $f' > 0$  per  $x < 0$  e  $f' < 0$  per  $x > 0$ , quindi  $f(x)$  è crescente per  $x < 0$ , e decrescente per  $x > 0$  (fig. 4).
- f)  $f(x)$  ha dunque la tipica forma di una curva a campana: quanto è larga? Si cerca il valore di  $x$  per cui  $f(x) = \frac{1}{2}$ :

$$\begin{aligned}e^{-x^2} &= \frac{1}{2} \\ -x^2 \ln e &= -\ln 2 \\ |x| &= \sqrt{\ln 2} \sim 0.83 \quad : \text{ semilarghezza a metà altezza}\end{aligned}$$

- g) decresce in modo estremamente rapido al crescere di  $|x|$ , per es.  $x = 3$   
 $f(x) = e^{-9} \simeq 1.234 \times 10^{-4}$



h) <sup>15</sup>  $\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$

Distribuzione di Gauss: La funzione di distribuzione di Gauss è:

<sup>15</sup> Posto  $I = \int_{-\infty}^{\infty} e^{-x^2} dx$  si può scrivere anche  $I = \int_{-\infty}^{\infty} e^{-y^2} dy$ .

Moltiplicando membro a membro

$$I^2 = \left( \int_{-\infty}^{\infty} e^{-x^2} dx \right) \cdot \left( \int_{-\infty}^{\infty} e^{-y^2} dy \right) = \iint_{(S)} e^{-(x^2+y^2)} dx dy$$

$I^2$  ha la forma di un integrale di superficie esteso a tutto il piano  $xy$ . Passando a coordinate polari piane:

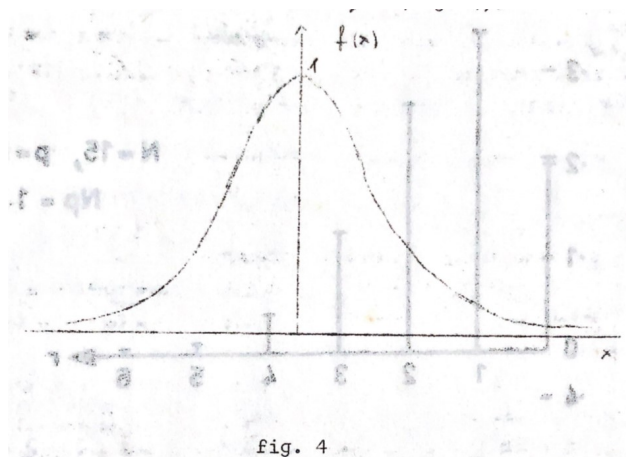


fig. 4.

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

dove  $\mu$  e  $\sigma$  sono due costanti di cui si chiarirà in seguito il significato. È una funzione di distribuzione continua, cioè la variabile casuale è una variabile continua e  $p(x)$  ne dà la probabilità, o meglio la densità di probabilità.

Assomiglia molto alla funzione  $e^{-x^2}$  con un cambiamento di variabile:

$$t = \frac{1}{\sqrt{2}} \frac{x - \mu}{\sigma} \quad \text{si vede che } p \sim e^{-t^2}.$$

Da questo si deduce che il massimo della distribuzione di Gauss è a  $t = 0$ , quindi per  $x = \mu$ .

La semilarghezza a metà altezza è:

$$|t| = \frac{1}{\sqrt{2}} \frac{|x - \mu|}{\sigma} = \sqrt{\ln 2}$$

$$|x - \mu| = \sigma\sqrt{2 \ln 2} \simeq 1.2 \sigma$$

Ne segue che  $\mu$  rappresenta il centro della campana e  $\sigma$  dà una misura della larghezza.

Come tutte le distribuzioni è normalizzata:

$$\begin{aligned} x^2 + y^2 &= r^2 & dS &= r dr d\theta \\ I^2 &= \iint_{(S)} e^{-r^2} r dr d\theta = 2\pi \int_0^\infty e^{-r^2} r dr = 2\pi \cdot \frac{1}{2} \int_0^\infty e^{-r^2} d(r^2) = -\pi [e^{-r^2}]_0^\infty = \\ &= -\pi(0 - 1) = \pi \end{aligned}$$

Quindi  $I = \sqrt{\pi}$  [n.d.c.].

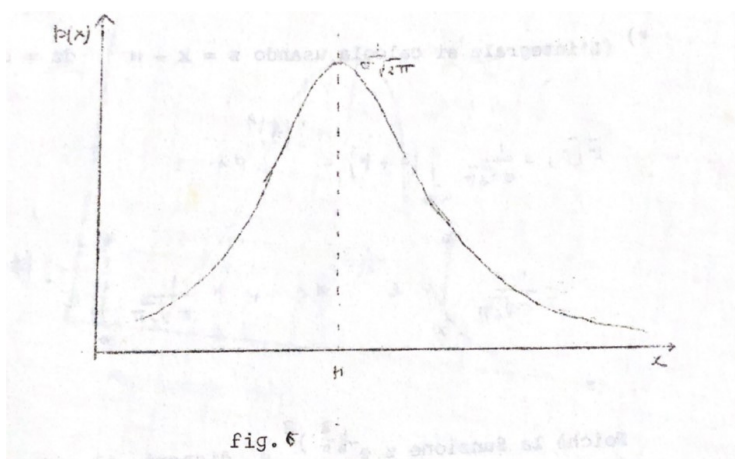


fig. 5.

$$\int_{-\infty}^{\infty} p(x) dx = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = 1$$

si dimostra cambiando variabile: si pone  $t = \frac{1}{\sqrt{2}} \frac{x-\mu}{\sigma}$  e si ricorda che  $\int_{-\infty}^{\infty} e^{-t^2} dt = \sqrt{\pi}$ .

La media della distribuzione è data da \*) Quindi la costante che compare nell'esponente della funzione di Gauss ed indicata con  $\mu$  è proprio la media della distribuzione.

La varianza della distribuzione è:

$$E[(x - \mu)^2] = \int_{-\infty}^{\infty} (x - \mu)^2 p(x) dx$$

---

\* (L'integrale si calcola usando  $z = x - \mu$   $dz = dx$ )

$$\begin{aligned} E[x] &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (z + \mu) e^{-\frac{1}{2}\left(\frac{z}{\sigma}\right)^2} dz = \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} z e^{-\frac{1}{2}\left(\frac{z}{\sigma}\right)^2} dz + \mu \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{z}{\sigma}\right)^2} dz \end{aligned}$$

Poiché la funzione  $z e^{-\frac{1}{2}\left(\frac{z}{\sigma}\right)^2}$  è dispari, il primo di questi integrali è nullo. Rimane

$$E[x] = \mu \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{z}{\sigma}\right)^2} dz$$

posto  $t = \frac{z}{\sigma} \cdot \frac{1}{\sqrt{2}}$  si ottiene immediatamente

$$E[x] = \mu$$

Il calcolo di questo integrale porta a \*)

$$E[(x - \mu)^2] = \sigma^2$$

Quindi la costante indicata con  $\sigma$  nella distribuzione di Gauss è proprio la deviazione standard.

In generale la probabilità che  $-a \leq x \leq a$  è:

$$P(|x| \leq a) = \int_{-a}^a p(x) dx$$

Questo integrale per la Gaussiana non ha espressione analitica: esistono però delle tabelle che permettono di sapere rapidamente i valori che interessano (Tav. 1). Le tavole sono di solito date per la variabile in forma standard:

$$z = \frac{x - \mu}{\sigma}$$

La sua funzione di distribuzione è

$$p(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

$z$  ha media = 0 e varianza = 1.

La funzione tabulata nella Tav. 1 è:

$$P(x) = \int_0^x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz$$

---

\*

$$\begin{aligned} E[(x - \mu)^2] &= E[x^2] - \mu^2 \\ E[x^2] &= \int_{-\infty}^{\infty} x^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx \\ z &= x - \mu \quad x = z + \mu \quad dx = dz \\ E[x^2] &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (z^2 + \mu^2 + 2z\mu) e^{-\frac{1}{2}\left(\frac{z}{\sigma}\right)^2} dz = \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} z^2 e^{-\frac{1}{2}\left(\frac{z}{\sigma}\right)^2} dz + \mu^2 \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{z}{\sigma}\right)^2} dz + 0 \\ t &= \frac{z}{\sqrt{2}\sigma} \quad dt = \frac{dz}{\sqrt{2}\sigma} \\ E[x^2] &= \frac{2\sigma^2}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} t^2 e^{-t^2} dt \sqrt{2}\sigma + \frac{\mu^2 \sqrt{2}\sigma}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-t^2} dt = \\ &= \frac{2\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} t^2 e^{-t^2} dt + \frac{\mu^2}{\sqrt{\pi}} \sqrt{\pi} = \frac{2\sigma^2}{\sqrt{\pi}} \cdot \frac{\sqrt{\pi}}{2} + \mu^2 \\ E[(x - \mu)^2] &= \sigma^2 + \mu^2 - \mu^2 = \sigma^2 \end{aligned}$$



Alcuni esempi sull'uso pratico delle tavole.

Si supponga di avere una variabile casuale gaussiana  $x$  di media  $\mu = 20$  e deviazione standard  $\sigma = 4$ .

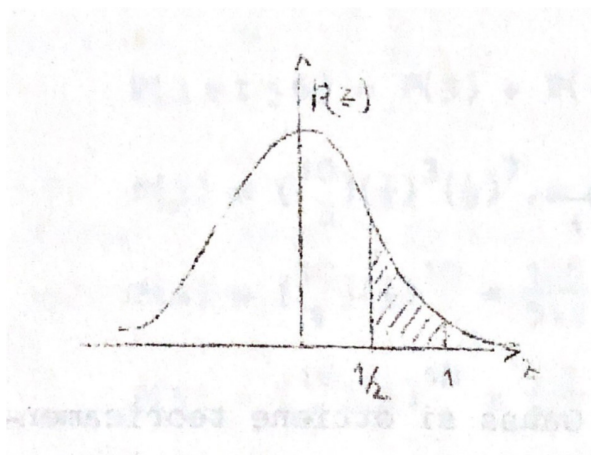
a) Quale è la probabilità che  $22 \leq x \leq 24$

$$z = \frac{x - 20}{4} \Rightarrow z_1 = \frac{22 - 20}{4} = \frac{1}{2} \quad z_2 = \frac{24 - 20}{4} = 1$$

$$P(22 \leq x \leq 24) = P\left(\frac{1}{2} \leq z \leq 1\right) = \int_{\frac{1}{2}}^1 p(z) dz =$$

$$= \int_0^1 p(z) dz - \int_0^{\frac{1}{2}} p(z) dz = 0.3413 - 0.1915 = 0.1498$$

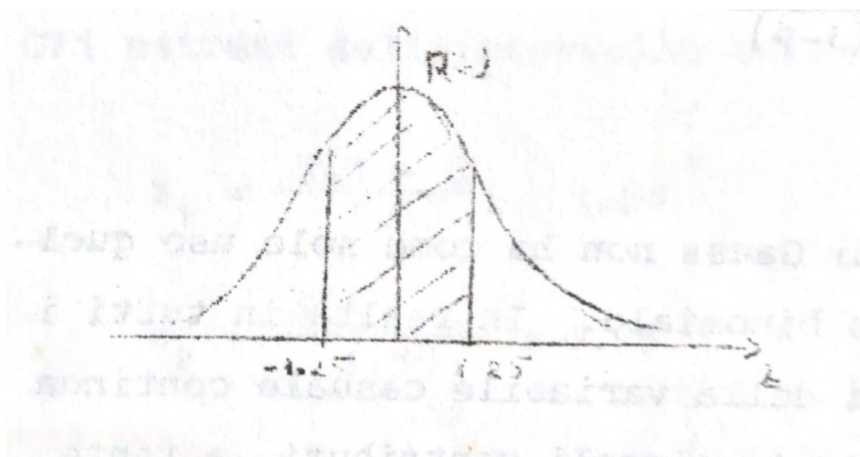
b) Quale è la probabilità che  $15 \leq x \leq 25$



$$z_1 = \frac{15 - 20}{4} = -1.25 \quad z_2 = \frac{25 - 20}{4} = 1.25$$

$$P(15 \leq x \leq 25) = P(-1.25 \leq z \leq 1.25) = \int_{-1.25}^{1.25} p(z) dz =$$

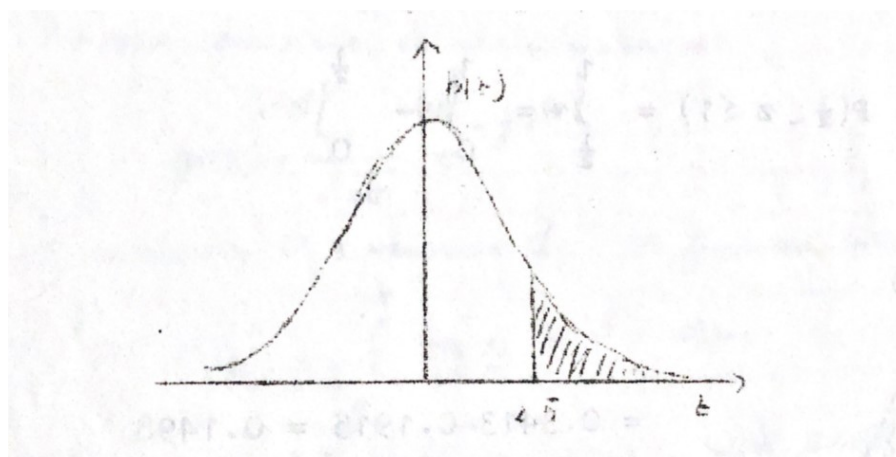
$$= 2 \int_0^{1.25} p(z) dz = 2 \cdot 0.3944 = 0.7888$$



c) Quale è la probabilità che  $x > 30$ :

$$z = \frac{30 - 20}{4} = \frac{10}{4} = 2.5$$

$$\begin{aligned} P(x > 30) &= P(z > 2.5) = \int_{2.5}^{\infty} p(z) dz = \\ &= \int_0^{\infty} p(z) dz - \int_0^{2.5} p(z) dz = \frac{1}{2} - 0.4946 = 0.0054 \end{aligned}$$



Discussione. La distribuzione di Gauss si ottiene teoricamente come caso limite della binomiale quando  $n \rightarrow \infty$  e  $p$  resta costante. In pratica quando sia

$np$  che  $n(1-p)$  sono  $\geq 5$  la gaussiana è già una buona approssimazione della binomiale.

La dimostrazione che la binomiale tende alla gaussiana quando  $np, nq \rightarrow \infty$  si può trovare per esempio su: Brownlee - Statistical theory and methodology etc. Cap. III (Wiley publications in statistics).<sup>16</sup>

Poiché la binomiale ha media  $np$  e deviazione standard  $\sqrt{np(1-p)}$  la corrispondente gaussiana è:

<sup>16</sup> Per dimostrarlo introduciamo la funzione  $g_X(t) = E[e^{tX}]$  che viene detta generatrice di  $X$ . Per la binomiale

$$g_X^{(B)}(t) = E[e^{tX}] = \sum_{X=0}^n e^{tX} \binom{n}{X} p^X q^{n-X} = \sum_{X=0}^n \binom{n}{X} (pe^t)^X q^{n-X} = (q + pe^t)^n$$

Per la gaussiana standardizzata ( $\mu = 0, \sigma = 1$ )  $P(X; \mu, \sigma) = \frac{1}{\sqrt{2\pi}} e^{-\frac{X^2}{2}}$

$$\begin{aligned} g^{(G)}(t; 0, 1) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{tX} e^{-\frac{X^2}{2}} dX = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{X^2}{2} + tX - \frac{t^2}{2} + \frac{t^2}{2}} dX = \\ &= e^{\frac{t^2}{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\left(\frac{X-t}{\sqrt{2}}\right)^2} dX = e^{\frac{t^2}{2}} \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-\left(\frac{X-t}{\sqrt{2}}\right)^2} d\left(\frac{X-t}{\sqrt{2}}\right) = \\ &= e^{\frac{t^2}{2}} \frac{1}{\sqrt{\pi}} \sqrt{\pi} = e^{\frac{t^2}{2}} \end{aligned}$$

Introducendo la variabile  $Z = \frac{X-\mu}{\sigma}$ ,  $g_Z(t) = E[e^{tZ}] = E[e^{-t\mu/\sigma} \cdot e^{Xt/\sigma}] = e^{-t\mu/\sigma} \cdot g_X(t/\sigma)$ . Per la binomiale ( $\mu = np$ ,  $\sigma = \sqrt{npq}$ )

$$g_Z^{(B)}(t) = e^{-npt/\sqrt{npq}} (q + pe^{t/\sqrt{npq}})^n = (qe^{-pt/\sqrt{npq}} + pe^{qt/\sqrt{npq}})$$

Sviluppando in serie:

$$\begin{aligned} g_Z^{(B)}(t) &= \left\{ q \left[ 1 - \frac{p}{\sqrt{npq}} t + \frac{1}{2} \frac{p^2}{npq} t^2 - \frac{1}{6} \frac{p^3}{(npq)^{3/2}} t^3 + \dots \right] + \right. \\ &\quad \left. + p \left[ 1 + \frac{q}{\sqrt{npq}} t + \frac{1}{2} \frac{q^2}{npq} t^2 + \frac{1}{6} \frac{q^3}{(npq)^{3/2}} t^3 + \dots \right] \right\}^n = \\ &= \left[ (q+p) + \frac{qp-pq}{\sqrt{npq}} t + \frac{1}{2} \frac{pq(p+q)}{npq} t^2 + \frac{1}{6} \frac{pq(q^2-p^2)}{npq\sqrt{npq}} t^3 + \dots \right]^n = \\ &= \left[ 1 + \frac{t^2}{2n} \left( 1 + \frac{1}{3} \frac{q-p}{\sqrt{npq}} + \dots \right) \right]^n = \left\{ 1 + \frac{t^2}{2n} \left[ 1 + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right) \right] \right\}^n \end{aligned}$$

Quindi  $\lim_{n \rightarrow \infty} g_Z^{(B)}(t) = \lim_{n \rightarrow \infty} \left\{ 1 + \frac{t^2}{2n} \left[ 1 + \mathcal{O}\left(\frac{1}{\sqrt{n}}\right) \right] \right\}^n = e^{\frac{t^2}{2}} = g^{(G)}(t; 0, 1)$ .

Se  $X$  è binomiale con  $\mu = np$  e  $\sigma = \sqrt{npq}$  la generatrice di  $Z = \frac{X-np}{\sqrt{npq}}$  tende per  $n \rightarrow \infty$  alla generatrice di una gaussiana generalizzata [n.d.c.].

$$\frac{1}{\sqrt{2\pi np(1-p)}} \cdot e^{-\frac{1}{2} \frac{(x-np)^2}{np(1-p)}}$$

Tuttavia la distribuzione di Gauss non ha come solo uso quello di essere il limite della binomiale. In realtà in tutti i casi in cui le fluttuazioni della variabile casuale continua sono dovute alla somma di tanti piccoli contributi, 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 loro, cioè non indipendenti, la distribuzione dei risultati di una misura è molto ben descritta dalla gaussiana.

Esempio sull'uso della gaussiana come approssimazione della binomiale.

Si cerchi la probabilità che lanciando 10 monete il numero di teste  $t$  sia  $3 \leq t \leq 6$ .

Usando la distribuzione binomiale  $P(t) = \binom{10}{t} \left(\frac{1}{2}\right)^t \left(1 - \frac{1}{2}\right)^{10-t}$  si ottiene:

$$P(3 \leq t \leq 6) = P(3) + P(4) + P(5) + P(6)$$

$$P(3) = \binom{10}{3} \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^7 = \frac{15}{128}$$

$$P(4) = \binom{10}{4} \left(\frac{1}{2}\right)^4 \left(\frac{1}{2}\right)^6 = \frac{105}{512}$$

$$P(5) = \binom{10}{5} \left(\frac{1}{2}\right)^5 \left(\frac{1}{2}\right)^5 = \frac{63}{256}$$

$$P(6) = \binom{10}{6} \left(\frac{1}{2}\right)^6 \left(\frac{1}{2}\right)^4 = \frac{105}{512}$$

$$P(3 \leq t \leq 6) = 0.7734$$

Per usare l'approssimazione gaussiana  $t$  deve essere considerata come una variabile continua, non discreta. Allora  $3 \leq t \leq 6$  diventa  $2.5 \leq x \leq 6.5$

$$\mu = np = 10 \cdot \frac{1}{2} = 5$$

$$\sigma = \sqrt{np(1-p)} = \sqrt{5 \cdot \frac{1}{2}} \simeq 1.58$$

Gli estremi dell'intervallo delle  $x$  in unità standard diventano:

$$z_1 = \frac{2.5 - 5}{1.58} = -1.58 \quad z_2 = \frac{6.5 - 5}{1.58} = 0.95$$

$$\begin{aligned} P(2.5 \leq x \leq 6.5) &= P(-1.58 \leq z \leq 0.95) = \int_{-1.58}^{0.95} p(z) dz = \\ &= \int_0^{1.58} p(z) dz + \int_0^{0.95} p(z) dz = 0.4429 + 0.3289 = 0.7718 \end{aligned}$$

che è molto vicino al valore 0.7734.

In questo caso il lavoro numerico per calcolare le due cose è all'incirca lo stesso,

ma se  $n = 100$  e si cerca la probabilità che  $30 \leq t \leq 60$ , l'approssimazione gaussiana è molto più comoda.

## 4 Distribuzione del $\chi^2$

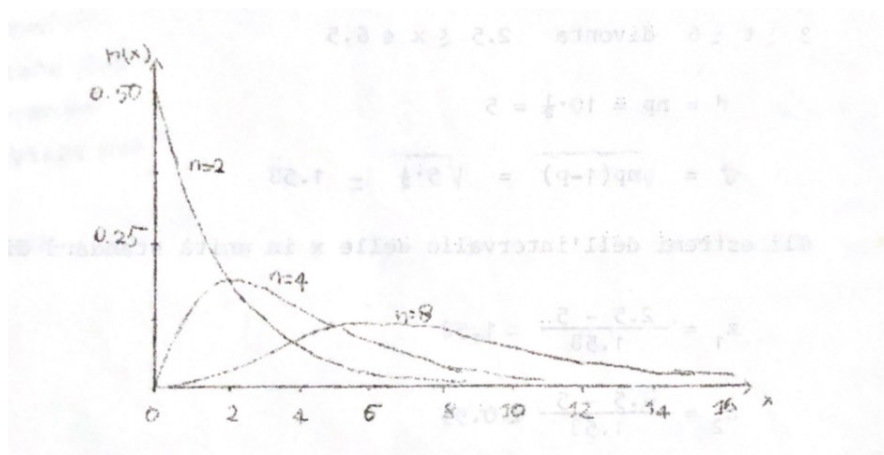
Sia  $x$  la somma dei quadrati di  $n$  variabili gaussiane standard ( $\mu = 0$ ,  $\sigma = 1$ ) indipendenti:

$$x = \sum_{i=1}^n z_i^2 \quad z_i \text{ normale } (0, 1)$$

Questa quantità si chiama  $\chi^2$  ad  $n$  gradi di libertà, è definito nell'intervallo  $(0, +\infty)$ , la sua funzione di distribuzione è:

$$p(x) = C_n x^{\frac{n-2}{2}} e^{-\frac{x}{2}}$$

$C_n$  costante di normalizzazione ( $C_n = \frac{1}{2^{n/2} \Gamma(\frac{n}{2})}$ )



Per grandi valori di  $n$ ,  $n \geq 30$ ,  $\sqrt{2\chi^2} - \sqrt{2n-1}$  è quasi normalmente distribuita con media zero e deviazione standard 1.

Il valor medio della distribuzione del  $\chi^2$  è uguale al numero di gradi di libertà, la varianza è uguale a due volte il numero dei gradi di libertà:

$$\begin{aligned} \mu &= n \\ \sigma^2 &= 2n \quad \sigma = \sqrt{2n} \end{aligned}$$

Anche per la distribuzione del  $\chi^2$  si trovano delle tavole che danno sia i valori della densità  $p(x)$  sia i valori della probabilità  $P = \int_0^x p(x') dx'$  (tav. 2).



---

## CAPITOLO IV

### 1 Propagazione degli errori

In questo capitolo si riprende la propagazione degli errori introdotta al Cap. II, dove si era trattato soltanto l'errore massimo che si commette nell'eseguire operazioni con grandezze  $a, b, c$ , affette da errori  $\Delta a, \Delta b, \Delta c \dots$

I risultati di quel capitolo si applicano quando conosciamo gli errori  $\Delta a, \Delta b, \Delta c \dots$  nel senso che certamente il valore della grandezza misurata è in  $a \pm \Delta a, b \pm \Delta b$ , etc.

Una situazione di questo tipo si incontra quando gli strumenti usati non sono molto sensibili, non si hanno fluttuazioni casuali e  $\Delta a, \Delta b \dots$  sono la risoluzione degli strumenti stessi.

Quando invece si ha a che fare con misure più raffinate e si hanno fluttuazioni casuali, in generale non si conoscono gli errori di una grandezza misurata perché questo implicherebbe conoscere il valore "vero" della grandezza, ma si conosce una stima dell'errore. Abbiamo visto che tale stima è data dalla deviazione standard della distribuzione che descrive la probabilità di determinare vari valori di quella grandezza.

Il problema della propagazione degli errori 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 del risultato.

a) Consideriamo dapprima il caso semplice:  $y = f(x)$ . Sia  $\bar{x}$  il valor medio di  $x$  e  $\sigma_x$  sia la sua deviazione standard. Anche per  $y$  possiamo ipotizzare l'esistenza di una funzione di distribuzione che dia la probabilità di determinare i vari valori della  $y$ , e quindi definire

$$\begin{aligned}\bar{y} &= E[y] \\ \sigma_y^2 &= E[(y - \bar{y})^2] \\ \sigma_y &= \sqrt{E[(y - \bar{y})^2]}\end{aligned}$$

Il problema della propagazione degli errori è ricondotto a vedere se  $\bar{y}$  e  $\sigma_y^2$  possono essere legate a  $\bar{x}$  e  $\sigma_x^2$ , e sotto quali ipotesi lo sono.

Scriviamo  $y$  mediante lo sviluppo in serie di Taylor intorno al punto  $\bar{x}$ :

$$y = f(\bar{x}) + (x - \bar{x}) \left( \frac{df}{dx} \right) \Big|_{\bar{x}} + \frac{1}{2} (x - \bar{x})^2 \left( \frac{d^2 f}{dx^2} \right) \Big|_{\bar{x}} + \dots$$

$$E[y] = E[f(\bar{x})] + E \left[ (x - \bar{x}) \left( \frac{df}{dx} \right) \Big|_{\bar{x}} \right] + E \left[ \frac{1}{2} (x - \bar{x})^2 \left( \frac{d^2 f}{dx^2} \right) \Big|_{\bar{x}} \right] + \dots$$

Essendo  $f(\bar{x})$ ,  $\left( \frac{df}{dx} \right)_{\bar{x}}$ ,  $\left( \frac{d^2 f}{dx^2} \right)_{\bar{x}}$  costanti ne segue che:

$$E[y] = f(\bar{x}) + 0 + \frac{1}{2} \left( \frac{d^2 f}{dx^2} \right)_{\bar{x}} \sigma_x^2 + \dots$$

Se ne deduce che

$$\bar{y} = f(\bar{x})$$

quando la funzione  $f(x)$  è lineare o comunque quando si possono trascurare i termini del 2° ordine e successivi.

Assunto  $\bar{y} = f(\bar{x})$  si ha:

$$\begin{aligned} \sigma_y^2 &= E[(f(x) - f(\bar{x}))^2] = E\left[(x - \bar{x})^2 \left( \frac{df}{dx} \right)_{\bar{x}}^2\right] = \\ &= \left( \frac{df}{dx} \right)_{\bar{x}}^2 E[(x - \bar{x})^2] = \left( \frac{df}{dx} \right)_{\bar{x}}^2 \sigma_x^2 \\ \sigma_y &= \left( \frac{df}{dx} \right)_{\bar{x}} \sigma_x \end{aligned}$$

Questo risultato è simile a quello che si ottiene per l'errore massimo ( $\Delta y = \left| \frac{df}{dx} \right|_{\bar{x}} \Delta x$ ) con la deviazione standard al posto dell'errore massimo.

b) Le variabili siano più di una per es. sia  $f = f(x, y)$ .

Nelle stesse approssimazioni del punto a) si può scrivere

$$\begin{aligned} \sigma_f^2 &= E[(f(x, y) - f(\bar{x}, \bar{y}))^2] = E \left[ \left\{ \left( \frac{df}{dx} \right)_{\bar{x}, \bar{y}} (x - \bar{x}) + \left( \frac{df}{dy} \right)_{\bar{x}, \bar{y}} (y - \bar{y}) \right\}^2 \right] = \\ &= E \left[ f_x^2 (x - \bar{x})^2 + f_y^2 (y - \bar{y})^2 + 2f_x f_y (x - \bar{x})(y - \bar{y}) \right] = \\ & \quad (\text{Si è usata la notazione } f_x = \left( \frac{df}{dx} \right)_{\bar{x}, \bar{y}}, \quad f_y = \left( \frac{df}{dy} \right)_{\bar{x}, \bar{y}}) \\ &= f_x^2 E[(x - \bar{x})^2] + f_y^2 E[(y - \bar{y})^2] + 2f_x f_y E[(x - \bar{x})(y - \bar{y})] = \\ &= f_x^2 \sigma_x^2 + f_y^2 \sigma_y^2 + 2f_x f_y \sigma_{xy} \end{aligned}$$

Si definisce covarianza la quantità

$$\sigma_{xy} = E[(x - \bar{x})(y - \bar{y})]$$



$\sigma_{xy}$  è una misura di quanto le fluttuazioni di  $x$  e di  $y$  sono legate tra di loro. Se sono indipendenti  $\sigma_{xy} = 0$ . Allora per grandezze indipendenti

$$\sigma_f^2 = f_x^2 \sigma_x^2 + f_y^2 \sigma_y^2$$

Più in generale

$$\sigma_f^2 = \sum_{i=1}^n f_{x_i}^2 \sigma_{x_i}^2$$

$$\sigma_f = \sqrt{\sum_{i=1}^n f_{x_i}^2 \sigma_{x_i}^2} \quad (*)$$

C'è da osservare che anche nel caso in cui si prende come errore la risoluzione dello strumento, quando si calcola una grandezza che è funzione di molte altre conviene usare la (\*) perché è pessimistico pensare che tutti gli errori contribuiscano nello stesso verso; è più significativo l'errore statistico.

Per esempio in una somma  $S = \sum_{i=1}^n \Delta s_i$  conviene usare  $\Delta S = \sqrt{\sum_i (\Delta s_i)^2}$ , come si deduce dalla (\*).

## 2 Teoria dei campioni

È rimasto il problema di vedere come trattare i risultati delle misure quando questi risultati non sono costanti. Riprendiamo perciò alcuni concetti introdotti al cap. I.

Era stato osservato che l'insieme delle misure sperimentali è una piccola parte di tutte le misure effettuabili, è quindi un campione di tutte le possibili misure, distribuite secondo la "parent distribution".

Da questo campione bisogna stimare la media  $\mu$  e la varianza  $\sigma^2$  della "parent distribution".

Come miglior stima della media  $\mu$  si prende la media aritmetica  $m$ :

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

dove  $n$  è il numero delle misure fatte, cioè la dimensione del campione.

$m$  è la "media campione", cioè la media della distribuzione campione e tende a  $\mu$  quando  $n$  tende all'infinito.

Come stima della varianza della distribuzione campione, tenuto conto del significato di varianza, si può pensare di prendere la media degli scarti al quadrato, cioè la quantità

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

Questo andrebbe benissimo se le quantità  $(x_i - m)$  fossero indipendenti tra di loro. In realtà non lo sono. Infatti:

$$\sum_{i=1}^n (x_i - m) = \sum_i x_i - \sum_i m = \sum_i x_i - nm = \sum_i x_i - n \sum_i \frac{x_i}{n} = 0$$

Fisicamente dipende dal fatto che il valore di  $m$  è stato ricavato proprio dagli  $x_i$ , quindi le quantità indipendenti sono  $n - 1$ . Allora si prende come stima della varianza della parent distribution la quantità:

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

$s^2$  è la varianza campione, cioè la varianza della distribuzione campione delle  $x_i$ . Si può dimostrare che

$$\lim_{n \rightarrow \infty} s^2 = \sigma_x^2$$

Quasi sempre interessa sapere qual è la varianza della media  $m$  delle  $x_i$ . Teniamo presente che anche la media campione è una variabile casuale che ha una sua funzione di distribuzione. La stima della varianza  $\sigma_m^2$  della distribuzione della media campione si può calcolare in questo modo:<sup>17</sup>

$$\begin{aligned} m &= \frac{1}{n} \sum_{i=1}^n x_i & s_m^2 &= \sum_{i=1}^n \left( \frac{\partial m}{\partial x_i} \right)^2 s_{x_i}^2 = s^2 \sum_{i=1}^n \left( \frac{\partial m}{\partial x_i} \right)^2 \\ \frac{\partial m}{\partial x_i} &= \frac{1}{n} \Rightarrow s_m^2 = s^2 \sum_{i=1}^n \frac{1}{n^2} = \frac{s^2}{n} \Rightarrow s_m^2 = \frac{1}{n(n-1)} \sum_{i=1}^n (x_i - m)^2 \end{aligned}$$

Questa è la varianza della distribuzione campione della media  $m$ . La deviazione standard è

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

Allora quando si hanno misure i cui risultati fluttuano, una volta verificato che tutto è stato fatto con la massima cura possibile e le variazioni non sono dovute ad errori dell'osservatore, il risultato della misura si scrive:

$$m \pm s_m \quad \text{od anche} \quad m \pm 2s_m$$

Qual è il significato da attribuire a questa scrittura?

Per poterlo discutere bisognerebbe conoscere la distribuzione della media campione. Ci aiuta a dire qualche cosa di questa distribuzione il teorema del limite

<sup>17</sup> Da (\*). Si può portare fuori  $s^2$  perché le  $x_i$  hanno la stessa funzione di distribuzione [n.d.c.].

centrale, del quale diamo soltanto l'enunciato:

qualunque sia la distribuzione delle  $x_i$  con  $\mu$  e  $\sigma$  finiti, la media campione  $m = \frac{1}{n} \sum_{i=1}^n x_i$  tende ad avere una distribuzione di Gauss col tendere all'infinito 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 la distribuzione è praticamente indistinguibile da una gaussiana in quasi tutti i casi sperimentati.

In una gaussiana la probabilità che

$$\begin{array}{ll} |x - \mu| \leq \sigma & \text{è} = 68.26\% \\ |x - \mu| \leq 2\sigma & \text{è} = 95.44\% \end{array}$$

Quando si dice che il risultato delle misure fatte è  $m \pm 2s_m$  questo vuol dire che il valore "giusto" (la media  $\mu$  della parent distribution) ha il 95% di probabilità di stare nell'intervallo  $(m - 2s_m, m + 2s_m)$ .

È chiaro che questi discorsi sono precisi 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 (cioè al numero delle grandezze indipendenti  $= n - 1$ ).

Esempio

Sia  $n = 5$ . Qual è l'intervallo intorno ad  $m$  tale che ci sia il 95% di probabilità di trovarci la  $\mu$ ? Questo intervallo è

$$m \pm t_{95, n-1} s$$

Esistono delle tabelle per i fattori correttivi  $t$  in funzione della probabilità e di  $n$ . Eccone alcuni valori tipici: \*

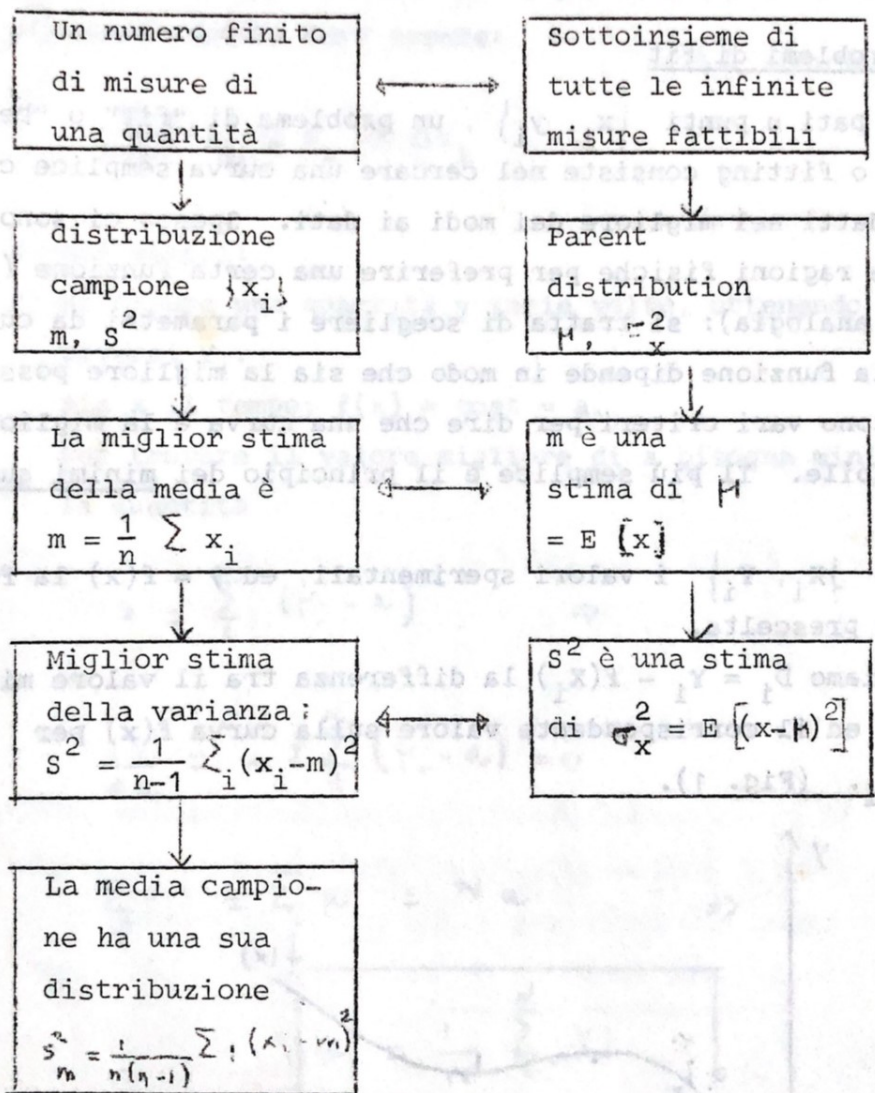
**Tabella 3.**

$n - 1$	$t_{95, n-1}$	$t_{99, n-1}$
1	12,7	63,7
4	2,78	4,60
8	2,31	3,36
20	2,09	2,85
$\infty$	1,96	2,58

---

\* Handbook of tables for probability and statistics by W.H. Beyer. The Chemical Rubber Co. Pag. 226.

## Schema riassuntivo della teoria dei campioni



---

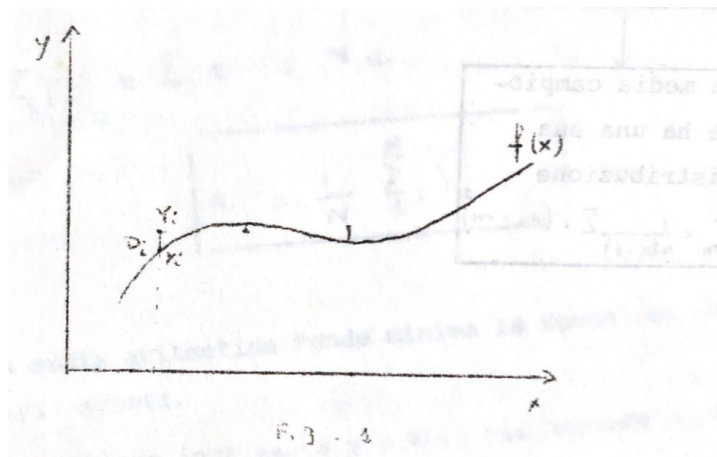
## CAPITOLO V

### 1 Problemi di fit

Dati  $n$  punti  $\{x_i, y_i\}$ , un problema di "fit" o "best fit" o fitting consiste nel cercare una curva semplice che si adatti nel migliore dei modi ai dati. Spesso ci sono delle ragioni fisiche per preferire una certa funzione (teoria, analogia): si tratta di scegliere i parametri da cui questa funzione dipende in modo che sia la migliore possibile. Esistono vari criteri per dire che una curva è la migliore possibile. il più semplice è il principio del minimi quadrati:

Siano  $\{X_i, Y_i\}$  i valori sperimentali, ed  $y = f(x)$  la funzione prescelta.

Chiamiamo  $D_i = Y_i - f(X_i)$  la differenza tra il valore misurato  $Y_i$  ed il corrispondente valore sulla curva  $f(x)$  per  $x = X_i$ . (Fig. 1).



$D_i$  può essere  $\geq 0$ . I parametri nella funzione  $f(x)$  vanno scelti in modo che

$$\sum_{i=1}^n D_i^2 = \text{minimo.}$$

Questo è il principio dei minimi quadrati.

Osservazione: il principio che sta alla base del metodo dei minimi quadrati tratta tutte le misure sullo stesso piano, cosa che è accettabile solo se le misure sono egualmente precise, cioè hanno lo stesso errore. Inoltre l'errore sulle  $X_i$  deve essere molto piccolo rispetto a quello sulle  $Y_i$ , più precisamente deve essere:

$$\left( \frac{df(x)}{dx} \right) \bigg|_{x_i} \Delta x_i \ll \Delta Y_i$$

Alcuni esempi:

1) Si misura una quantità  $y$  varie volte, ottenendo  $N$  valori diversi  $Y_i$ .

Sia  $x$  il tempo;  $f(x) = \text{cost} = a$ .

Per trovare il valore migliore di  $a$  bisogna minimizzare la quantità

$$S = \sum_{i=1}^N (Y_i - a)^2 \Rightarrow \frac{dS}{da} = -2 \sum_{i=1}^N (Y_i - a) = 0$$

$$\sum_i Y_i = \sum_i a = Na \Rightarrow a = \frac{1}{N} \sum_{i=1}^N Y_i$$

La media aritmetica rende minima la somma dei quadrati degli scarti.

2) La funzione ipotizzata  $y = f(x)$  sia lineare:  $f(x) = a + bx$

$\{X_i, Y_i, i = 1, \dots, N\}$  misure.

Per trovare i migliori valori di  $a$  e  $b$  perché la retta si adatti il meglio possibile ai dati sperimentali costruiamo la funzione

$$S = \sum_{i=1}^N (Y_i - f(X_i))^2 = \sum_{i=1}^N (Y_i - a - bX_i)^2$$

Bisogna cercare il minimo di  $S$  al variare di  $a$  e  $b$ .

$$\begin{cases} \frac{\partial S}{\partial a} = -2 \sum_{i=1}^N (Y_i - a - bX_i) = 0 \\ \frac{\partial S}{\partial b} = -2 \sum_{i=1}^N X_i (Y_i - a - bX_i) = 0 \end{cases}$$

$$\begin{cases} Na + b \sum_{i=1}^N X_i = \sum_{i=1}^N Y_i \\ a \sum_{i=1}^N X_i + b \sum_{i=1}^N X_i^2 = \sum_{i=1}^N X_i Y_i \end{cases}$$

$$a = \frac{\sum_i Y_i \sum_i X_i^2 - \sum_i X_i \sum_i X_i Y_i}{N \sum_i X_i^2 - (\sum_i X_i)^2}$$

$$b = \frac{N \sum_i X_i Y_i - \sum_i X_i \sum_i Y_i}{N \sum_i X_i^2 - (\sum_i X_i)^2}$$

$$\Delta a = \sqrt{\sum_j \left( \frac{\partial a}{\partial Y_j} \right)^2 \Delta Y_j^2} = \sqrt{\frac{\Delta Y^2 \sum_i X_i^2}{N \sum_i X_i^2 - (\sum_i X_i)^2}}$$

$$\Delta b = \sqrt{\sum_j \left( \frac{\partial b}{\partial Y_j} \right)^2 \Delta Y_j^2} = \sqrt{\frac{\Delta Y^2 N}{N \sum_i X_i^2 - (\sum_i X_i)^2}}$$

18

Esercizi:

Ricavare col metodo dei minimi quadrati il miglior valore di  $a$  per le seguenti funzioni:

- 1)  $f(x) = ax$
- 2)  $f(x) = ax^2$

18

$$\frac{\partial a}{\partial Y_j} = \frac{\sum_i X_i^2 - X_j \sum_i X_i}{N \sum_i X_i^2 - (\sum_i X_i)^2} \quad \frac{\partial a}{\partial Y_j} = \frac{NX_j - \sum_i X_i}{N \sum_i X_i^2 - (\sum_i X_i)^2}$$

$$\left( \frac{\partial a}{\partial Y_j} \right)^2 = \frac{(\sum_i X_i^2)^2 - 2X_j \sum_i X_i^2 \sum_i X_i + X_j^2 (\sum_i X_i)^2}{(N \sum_i X_i^2 - (\sum_i X_i)^2)^2}$$

$$\left( \frac{\partial b}{\partial Y_j} \right)^2 = \frac{N^2 X_j^2 - 2NX_j \sum_i X_i + (\sum_i X_i)^2}{(N \sum_i X_i^2 - (\sum_i X_i)^2)^2}$$

$$\sum_{j=1}^N \left( \frac{\partial a}{\partial Y_j} \right)^2 \Delta Y_j^2 = \frac{N(\sum_i X_i^2)^2 - 2\sum_i X_i^2 (\sum_i X_i)^2 + \sum_i X_i^2 (\sum_i X_i)^2}{(N \sum_i X_i^2 - (\sum_i X_i)^2)^2} \Delta Y^2 =$$

$$= \frac{N(\sum_i X_i^2)^2 - \sum_i X_i^2 (\sum_i X_i)^2}{(N \sum_i X_i^2 - (\sum_i X_i)^2)^2} \Delta Y^2 = \sum_i X_i^2 \frac{N \sum_i X_i^2 - (\sum_i X_i)^2}{(N \sum_i X_i^2 - (\sum_i X_i)^2)^2} \Delta Y^2 =$$

$$= \frac{\sum_i X_i^2}{N \sum_i X_i^2 - (\sum_i X_i)^2} \Delta Y^2 \text{ essendo } \Delta Y_j = \Delta Y (\text{errore strumentale})$$

$$\sum_{j=1}^N \left( \frac{\partial b}{\partial Y_j} \right)^2 \Delta Y_j^2 = \frac{N^2 \sum_i X_i^2 - 2N(\sum_i X_i)^2 + N(\sum_i X_i)^2}{(N \sum_i X_i^2 - (\sum_i X_i)^2)^2} \Delta Y^2 =$$

$$= \frac{N^2 \sum_i X_i^2 - N(\sum_i X_i)^2}{(N \sum_i X_i^2 - (\sum_i X_i)^2)^2} \Delta Y^2 = N \frac{N \sum_i X_i^2 - (\sum_i X_i)^2}{(N \sum_i X_i^2 - (\sum_i X_i)^2)^2} \Delta Y^2 =$$

$$= \frac{N}{N \sum_i X_i^2 - (\sum_i X_i)^2} \Delta Y^2 \quad [n.d.c.].$$

3)  $f(x) = a e^{-x}$  <sup>19</sup>

Il metodo dei minimi quadrati dà dei risultati semplici in tutti i casi in cui la funzione  $f(x)$  dipende in modo lineare dai parametri: infatti in questo caso le condizioni di minimo conducono a sistemi di equazioni lineari nei parametri. In generale una forma della funzione del tipo

$$f(x) = a_1 g_1(x) + a_2 g_2(x) + \dots$$

dove le  $g_i$  non contengono ulteriori parametri, si può trattare coi metodi indicati sopra.

## 2 Metodo del minimo $\chi^2$ .

Supponiamo di avere delle misure  $\{X_i, Y_i\}$  tali che l'errore su  $Y_i$  sia  $\sigma_i$  mentre manteniamo per ora l'ipotesi che l'errore su  $X_i$  sia molto piccolo.

Invece di calcolare le differenze  $D_i = Y_i - f(X_i)$  si usa  $\frac{Y_i - f(X_i)}{\sigma_i}$  cioè ogni differenza viene misurata col relativo errore; si cercherà poi il minimo della funzione:

$$S = \sum_{i=1}^N \left( \frac{Y_i - f(X_i)}{\sigma_i} \right)^2$$

Questo metodo viene indicato come il metodo del minimo  $\chi^2$  infatti se si assume che le  $Y_i$  abbiano una distribuzione gaussiana  $\frac{Y_i - f(X_i)}{\sigma_i}$  sono variabili gaussiane con media zero e varianza 1 ed  $S$  è proprio un  $\chi^2$ . (Vedi cap. III).  
Alcuni esempi:

1)  $f(x) = \text{cost} = a$

$$\begin{aligned} S &= \sum_{i=1}^N \left( \frac{Y_i - a}{\sigma_i} \right)^2 = \min \Rightarrow \\ \frac{dS}{da} &= -2 \sum_i \left( \frac{Y_i - a}{\sigma_i} \right) \frac{1}{\sigma_i} = 0 \\ a &= \frac{\sum_i \frac{Y_i}{\sigma_i^2}}{\sum_i \frac{1}{\sigma_i^2}} \quad \Delta a = \sqrt{\sum_j \left( \frac{\partial a}{\partial Y_j} \right)^2 \sigma_j^2} = \sqrt{\frac{1}{\sum_i \frac{1}{\sigma_i^2}}} \end{aligned}$$

Questa è la media pesata delle  $Y_i$ : i valori delle  $Y_i$  sono pesati con l'inverso del quadrato dell'errore, e si usa tutte le volte che le misure hanno precisione

<sup>19</sup> Poniamo in generale  $f(x) = ag(x)$

$$S = \sum_i (Y_i - ag(X_i))^2 = \min \Rightarrow -2 \sum_i g(X_i) (Y_i - ag(X_i)) = 0$$

$$a = \frac{\sum_i g(X_i) Y_i}{\sum_i (g(X_i))^2} \Rightarrow 1) a = \frac{\sum_i X_i Y_i}{\sum_i X_i^2} \quad 2) a = \frac{\sum_i X_i^2 Y_i}{\sum_i (X_i^2)^2} \quad 3) a = \frac{\sum_i e^{-X_i} Y_i}{\sum_i e^{-2X_i}} \text{ [n.d.c.]}$$



differente.

2)  $f(x) = a - bx$

$$S = \sum_{i=1}^N \left( \frac{Y_i - a - bX_i}{\sigma_i} \right)^2 = \min \Rightarrow$$

$$\begin{cases} \frac{\partial S}{\partial a} = -2 \sum_{i=1}^N \frac{1}{\sigma_i} \left( \frac{Y_i - a - bX_i}{\sigma_i} \right) = 0 \\ \frac{\partial S}{\partial b} = -2 \sum_{i=1}^N \frac{X_i}{\sigma_i} \left( \frac{Y_i - a - bX_i}{\sigma_i} \right) = 0 \end{cases}$$

$$\begin{cases} a \sum_i \frac{1}{\sigma_i^2} - b \sum_i \frac{X_i}{\sigma_i^2} = \sum_i \frac{Y_i}{\sigma_i^2} \\ a \sum_i \frac{X_i}{\sigma_i^2} - b \sum_i \frac{X_i^2}{\sigma_i^2} = \sum_i \frac{X_i Y_i}{\sigma_i^2} \end{cases}$$

$$a = \frac{\sum_i \frac{Y_i}{\sigma_i^2} \sum_i \frac{X_i^2}{\sigma_i^2} - \sum_i \frac{X_i}{\sigma_i^2} \sum_i \frac{X_i Y_i}{\sigma_i^2}}{\sum_i \frac{1}{\sigma_i^2} \sum_i \frac{X_i^2}{\sigma_i^2} - \left( \sum_i \frac{X_i}{\sigma_i^2} \right)^2}$$

$$b = \frac{\sum_i \frac{1}{\sigma_i^2} \sum_i \frac{X_i Y_i}{\sigma_i^2} - \sum_i \frac{Y_i}{\sigma_i^2} \sum_i \frac{X_i}{\sigma_i^2}}{\sum_i \frac{1}{\sigma_i^2} \sum_i \frac{X_i^2}{\sigma_i^2} - \left( \sum_i \frac{X_i}{\sigma_i^2} \right)^2}$$

$$\Delta a = \sqrt{\sum_j \left( \frac{\partial a}{\partial Y_j} \right)^2} \Delta Y_j^2 = \sqrt{\frac{\sum_i \frac{X_i^2}{\sigma_i^2}}{\sum_i \frac{1}{\sigma_i^2} \sum_i \frac{X_i^2}{\sigma_i^2} - \left( \sum_i \frac{X_i}{\sigma_i^2} \right)^2}}$$

$$\Delta b = \sqrt{\sum_j \left( \frac{\partial b}{\partial Y_j} \right)^2} \Delta Y_j^2 = \sqrt{\frac{\sum \frac{1}{\sigma_i^2}}{\sum_i \frac{1}{\sigma_i^2} \sum_i \frac{X_i^2}{\sigma_i^2} - \left( \sum_i \frac{X_i}{\sigma_i^2} \right)^2}}$$

### 3 Fit di tipo generale

Si vuole studiare il problema di fitting nel caso più generale in cui non sono trascurabili né gli errori su  $y$ , né quelli su  $x$ .

$x$  ed  $y$  siano due variabili legate dalla relazione  $y = f(x, p)$ , dove  $p$  è un insieme di parametri  $p \equiv \{p_\mu; \mu = 1, \dots, k\}$  il cui valore migliore si vuole determinare in base ai dati sperimentali.

Siano  $\{X_i \pm \xi_i, Y_i \pm \eta_i, i = 1, \dots, N\}$  un insieme di risultati sperimentali,  $\xi_i$  e  $\eta_i$  sono stime degli errori sui singoli valori di  $X_i$  ed  $Y_i$  rispettivamente.

il primo problema che si incontra è che in una equazione del tipo  $y = f(x, p)$

non è chiaro qual è il valore da usare per  $x$  in modo da calcolare il valore teorico di  $y$  e quindi cercare i migliori valori dei parametri che soddisfano la richiesta di rendere minimo il  $\chi^2$ .

Non conoscendo  $x$  il problema è dunque concettualmente più complicato: bisogna determinare i valori  $(x_i, y_i)$  tali che

$$\begin{cases} \sum_{i=1}^N \left( \frac{x_i - X_i}{\xi_i} \right)^2 + \left( \frac{y_i - Y_i}{\eta_i} \right)^2 = \min \\ \text{con la condizione che} \\ F(x_i, y_i, p) = 0 \quad \forall i = 1, \dots, N \end{cases}$$

$F(x, y, p)$  è la relazione che si suppone legare la  $x$  alla  $y$  scritta in forma implicita.

Questo è un problema di minimo condizionato, usando il metodo dei moltiplicatori di Lagrange ci si riduce al problema seguente:

$$S = \sum_{i=1}^N \left\{ \left( \frac{x_i - X_i}{\xi_i} \right)^2 + \left( \frac{y_i - Y_i}{\eta_i} \right)^2 + \lambda_i F(x_i, y_i, p) \right\} = \min$$

Le cose da determinare sono  $x_i, y_i, \lambda_i, p$ , in tutto  $3n + k$ ; ci sono quindi derivate parziali rispetto a  $3n + k$  variabili. Il problema si può risolvere nel caso generale solo con metodi numerici approssimati.

Diamo alcuni cenni sulla tecnica della soluzione:

La cosa più importante è supporre di conoscere un valore approssimato dei parametri. Il secondo passo è di usare invece di  $F(x, y, p)$  un'approssimazione lineare sviluppando al 1° ordine.

Quindi c'è una certa quantità di algebra, sostituzioni ecc.; il risultato è un sistema di equazioni lineari che dicono di quanto devono essere corrette le stime iniziali ai parametri. A questo punto si procede con metodo iterativo, cioè si prendono i nuovi valori dei parametri si sostituiscono nella funzione e si ricomincia. Se il processo converge si ottengono alla fine i valori richiesti per i parametri.

Esempio: Sia  $F(x, y, p) = y - px = 0$

$$S = \sum_{i=1}^N \left\{ \left( \frac{x_i - X_i}{\xi_i} \right)^2 + \left( \frac{y_i - Y_i}{\eta_i} \right)^2 + \lambda_i (y_i - px_i) \right\} = \min$$

$$\frac{\partial S}{\partial x_j} = 2 \frac{(x_j - X_j)}{\xi_j^2} - \lambda_j p = 0 \quad 1)$$

$$\frac{\partial S}{\partial y_j} = 2 \frac{(y_j - Y_j)}{\eta_j^2} + \lambda_j = 0 \quad 2)$$

$$\frac{\partial S}{\partial \lambda_j} = y_j - px_j = 0 \quad 3)$$

$$\frac{\partial S}{\partial p} = \sum_i \lambda_i x_i = 0 \approx \sum_i \lambda_i X_i \quad 4)$$

(questa è la linearizzazione della funzione  $F$ ).

Si ricavano  $x_j$  e  $y_j$  dalle prime due e si sostituiscono nella terza equazione, quindi si ricava  $\lambda_j$  dalla terza:

$$\lambda_j = 2 \frac{Y_j - pX_j}{p^2\xi^2 + \eta^2}$$

questo si sostituisce nella quarta equazione e si trova

$$\sum_i X_i \frac{Y_j - pX_j}{p^2\xi^2 + \eta^2} = 0$$

da cui

$$p = \frac{\sum_{i=1}^N \frac{X_i Y_i}{p^2 \xi_i^2 + \eta_i^2}}{\sum_{i=1}^N \frac{X_i^2}{p^2 \xi_i^2 + \eta_i^2}} \quad (*)$$

a questo punto entra in funzione il metodo iterativo.

Si stima  $p$  (ad esempio dal grafico), si usa questa stima nel secondo membro della (\*) e si ottiene un nuovo  $p$  (chiamato  $p'$ ); se è molto diverso dalla stima si ripete il procedimento; cioè si sostituisce  $p'$  nel 2° membro della (\*), si ottiene un nuovo  $p$ :  $p''$  etc. Ci si ferma quando la differenza tra un valore di  $p$  ed un successivo è entro gli errori su  $p$ . Per stimare gli errori sui parametri si usa l'equazione (\*) e le tecniche della propagazione degli errori ed il valore stimato di  $p$ .

$$\sigma_p^2 = \Delta p^2 = \sum_j \left\{ \left( \frac{\partial p}{\partial x_j} \right) \xi_j^2 + \left( \frac{\partial p}{\partial y_j} \right) \eta_j^2 \right\}$$

#### 4 Test del $\chi^2$ .

Richiamiamo la definizione della variabile " $\chi^2$  a  $n$  gradi di libertà".

Date  $n$  variabili  $x_i$  gaussiane a media  $\mu_i$  e deviazione standard  $\sigma_i$ , costruite le variabili standard

$$z_i = \frac{x_i - \mu_i}{\sigma_i}$$

la quantità

$$S = \sum_{i=1}^n \left( \frac{x_i - \mu_i}{\sigma_i} \right)^2$$

si chiama  $\chi^2$  ad  $n$  gradi di libertà

$$P(S) = \frac{S^{\frac{n}{2}-1}}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} e^{-\frac{S}{2}}$$

$$E[S] = n$$

$$\sigma_S^2 = 2n$$

Le tavole riportate in app. danno qual è il valore  $\alpha$  di  $S$  tale che per un assegnato  $n$  ed una assegnata probabilità  $v$

$$P(S_n \leq \alpha) = v$$

Per esempio sia  $n = 10$ , ci si domanda quale è il valore  $\alpha$  di  $\chi^2$  tale che la probabilità che  $S \leq \alpha$  sia il 95%:

$$P(S_{10} \leq \alpha) = 95\%$$

Le tavole in corrispondenza della riga 10, colonna 4 danno  $\alpha = 18,3^*$ .

Il  $\chi^2$  può fornire un criterio generale ed obiettivo per decidere se una certa equazione, una certa legge descrive bene oppure no i risultati sperimentali.

Infatti supponiamo di aver misurato le quantità  $\{Y_i, X_i\}$  legate tra di loro da una equazione  $y = f(x, p)$ , la cui forma viene ipotizzata in base a considerazioni fisiche etc. e i parametri  $p$  sono determinati con uno dei metodi di fit discussi prima, per es. il metodo del minimo  $\chi^2$ .

Il  $\chi^2$  in questo caso vale

$$\chi^2 = \sum_{i=1}^N \frac{(Y_i - f(X_i, p))^2}{\sigma_{y_i}^2} = S$$

Quanto più  $S$  è piccolo, tanto meno si discosta la curva teorica dai punti sperimentali.

Supposto che gli errori siano stati valutati in modo corretto, vediamo quale è il significato da attribuire al numero  $S$ . Si cerca sulle tavole in app. in corrispondenza al numero di gradi di libertà dove si trova  $S$ . Quando si fa un fit il numero di gradi di libertà  $\nu$  è il numero  $N$  delle misure meno il numero  $n_p$  di parametri della funzione  $f(x, p)$ :  $\nu = N - n_p$ .

Per esempio supponiamo  $S = 7.2$   $\nu = 5$ . Si ottiene  $P(\chi^2 \leq S) \sim 80\%$ . Questo significa che ripetendo le misure tante volte, nell'80% dei casi si otterrà un valore minore di 7.2, cioè 7.2 è un valore non molto buono per il  $\chi^2$ , ma nemmeno assurdo, nel 20% dei casi può succedere che venga un risultato più alto.

Convenzionalmente si accetta come limite per i risultati buoni quello del 95% da un lato e dello 0.5% dall'altro, perché anche se a prima vista sembra che la cosa migliore sia ottenere dei  $\chi^2$  molto piccoli (o addirittura zero), in realtà

---

\* Se il numero dei gradi di libertà è molto elevato:  $n \geq 30$ , invece delle tavole del  $\chi^2$  si usano quelle della distribuzione normale, perché in questo limite  $\sqrt{2\chi^2 - \sqrt{2n-1}} = z$  è una variabile casuale normale (a media zero e varianza 1).

questi casi vanno guardati con sospetto. Di solito  $\chi^2$  piccoli si ottengono perché gli errori sono sovrastimati (cioè si assume per  $\sigma^2$  un valore più grande di quello ragionevole); oppure i risultati sono truccati, cioè lo sperimentatore (consciamente o no) tende ad attribuire al risultato della misura il valore che gli piacerebbe.

Vediamo ora un'altra applicazione del test del  $\chi^2$ .

Supponiamo di fare un esperimento che abbia come oggetto l'osservazione di possibili eventi  $E_1, E_2, \dots, E_K, \dots, E_N$ . Registriamo la frequenza di questi possibili eventi:  $O_1, O_2, \dots, O_K, \dots, O_N$  (oppure la frequenza relativa).

D'altra parte possiamo fare certe ipotesi sul fenomeno che stiamo studiando le quali ci permettono di determinare le frequenze (frequenze relative) teoriche:  $e_1, e_2, \dots, e_K, \dots, e_N$ . Ci saranno differenze tra le frequenze osservate e quelle teoriche. Per determinare se le differenze sono significative (quindi eventualmente rifiutare le ipotesi che hanno portato a calcolare  $e_i$ ) si usa la quantità

$$\sum_{i=1}^n \frac{(O_i - e_i)^2}{e_i} \quad (*)$$

che viene considerata ancora una variabile  $\chi^2$ , anche se in realtà non lo è. Formalmente ricorda un  $\chi^2$  perché si suppone che sia Poissoniana la distribuzione di ognuna delle frequenze  $O_i$ :  $O_i$  misurato rappresenta il valor medio della distribuzione campione,  $e_i$  rappresenta il valor medio della parent distribution, e inoltre per la poissoniana  $\sigma_i^2 = e_i$ . I gradi di libertà sono  $n - m - 1$ , dove  $n$  è il numero delle osservazioni,  $m$  è il numero dei parametri che servono per calcolare le probabilità degli eventi  $E_1, E_2, \dots, E_K, \dots, E_N$  quindi le frequenze teoriche; inoltre c'è la seguente relazione:

$$\sum_{i=1}^n \frac{O_i}{N} = 1$$

od anche

$$\sum_{i=1}^n O_i = N$$

$N$  è il numero degli eventi osservati

e questa relazione causa la perdita di un altro grado di libertà.

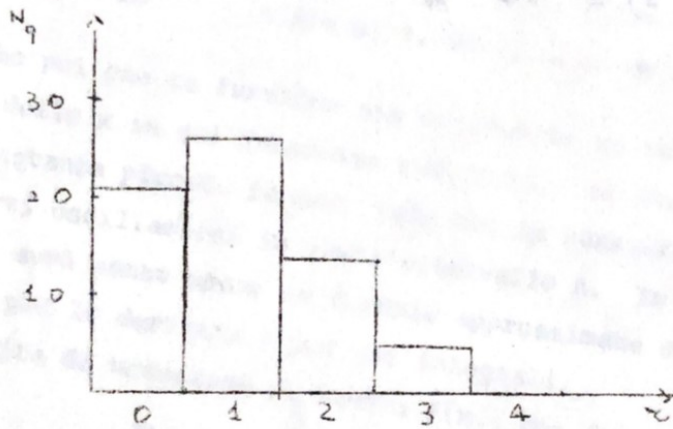
Poiché la quantità (\*) è considerata un  $\chi^2$  ad essa si attribuisce lo stesso significato probabilistico trattato precedentemente.

Esempio:

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 nel quadrato.

Quindi si fa una tabella suddividendo i quadrati in funzione del numero dei

N° dei dischi r per quadrato	N° quadrati con r dischi	Frequenza relativa	Freq.reat. teorica
0	21	21/64	0.37
1	26	26/64	0.37
2	13	13/64	0.18
3	4	4/64	0.06
4	0	0	0.02



dischi che essi contengono: 0, 1, 2, 3 ... Per calcolare le frequenze teoriche si fa l'ipotesi che la distribuzione della variabile  $r$  sia una Poissoniana:

$$P(r) = \frac{\mu^r e^{-\mu}}{r!}$$

$\mu$  è stimato in questo modo:  $\frac{n^\circ \text{totale eventi}}{n^\circ \text{degli intervalli}} = \frac{n^\circ \text{totale dischi}}{n^\circ \text{degli intervalli}} = \frac{64}{64} = 1$ .

$$1) P(0) = \frac{1^0}{0!} e^{-1} = \frac{1}{e}$$

$$2) P(1) = \frac{1}{e}$$

$$3) P(2) = \frac{1}{e} \cdot \frac{1}{2}$$

$$4) \begin{cases} P(3) = \frac{1}{e} \cdot \frac{1}{3!} \\ P(4) = \frac{1}{e} \cdot \frac{1}{4!} \end{cases} \Rightarrow P(r > 2) = 1 - \left( \frac{1}{e} + \frac{1}{e} + \frac{1}{e \cdot 2} \right) = 1 - \frac{5}{2} \cdot \frac{1}{e} = 0.08 =$$

$$P(3) + P(4)$$

$$S = \sum_{i=1}^4 \frac{(O_i - e_i)^2}{e_i} = \frac{\left(\frac{26}{64} - 0.37\right)^2}{0.37} + \frac{\left(\frac{21}{64} - 0.37\right)^2}{0.37} + \frac{\left(\frac{13}{64} - 0.18\right)^2}{0.18} + \frac{\left(\frac{4}{64} - 0.08\right)^2}{0.08} = 0.015$$

$$\nu = 4 - 2 = 2$$
$$P(\chi^2 \leq S) \simeq 0.7\%$$





---

## CAPITOLO VI

Vogliamo trattare alcuni problemi riguardanti una funzione  $f(x)$  tabulata, cioè si conoscono i valori della funzione  $f(x)$  in punti discreti dell'asse delle  $x$ . Si può conoscere o no la forma analitica di  $f(x)$ . Si pongono i seguenti problemi: come derivare la funzione  $f(x)$  nota solo per punti, come integrarla, quale valore assume  $f(x)$  per  $x_i < x < x_{i+1}$ .

### 1 Formule approssimate per la derivata di una funzione $f(x)$ data per punti

Supponiamo che i punti della tabella siano ugualmente spazati di una quantità  $h$ :

$$x_i = x_0 \pm ih \quad i = 0, 1, 2, \dots, n$$

Supponiamo poi che la funzione sia derivabile su tutto l'intervallo delle  $x$  in cui interessa elaborarla, ed inoltre che  $h$  sia abbastanza piccolo in modo tale che la funzione non compia forti oscillazioni in quell'intervallo  $h$ . In queste condizioni avrà senso usare le formule approssimate che si otterranno per le derivate e per gli integrali.

Per semplicità di notazione si indica  $f(x_i)$  con  $f_i$  e  $\left(\frac{d^k f(x)}{dx^k}\right)\Big|_{x=x_i}$  con  $f_i^{(k)}$ .

Il valore  $f_m$  di  $f(x)$  in ogni punto  $x_m$  si può esprimere in termini di  $f(x)$  e della sua derivata in un altro punto  $x_i$  per mezzo di uno sviluppo in serie di Taylor intorno ad  $x_i$ :

$$f_m = f(x_i \pm mh) = f_i \pm mh f'_i + \frac{(mh)^2}{2!} f''_i \pm \frac{(mh)^3}{3!} f'''_i \dots$$

Quindi per  $m = 1$

$$1) f_{i+1} = f_i + h f'_i + \frac{h^2}{2} f''_i + \frac{h^3}{6} f'''_i + \frac{h^4}{24} f^{IV}_i + \dots$$

Una espressione approssimata per la derivata di  $f(x)$  in  $x_i$  è:

$$f'_i = \frac{f_{i+1} - f_i}{h} - \left( \frac{h}{2} f''_i + \frac{h^2}{6} f'''_i + \dots \right)$$

trascurando i termini tra parentesi si ottiene

$$f'_i = \frac{f_{i+1} - f_i}{h}$$

L'errore è dato dai termini entro parentesi, dei quali il termine predominante è:  $-\frac{h}{2} f''_i$ .

Per  $m = -1$  si ottiene

$$2) \ f_{i-1} = f_i - h f'_i + \frac{h^2}{2} f''_i - \frac{h^3}{6} f'''_i + \frac{h^4}{24} f^{IV}_i + \dots$$

Sottraendo  $f_{i-1}$  da  $f_{i+1}$  si ottiene un'altra espressione approssimata per la derivata prima:

$$f'_i = \frac{f_{i+1} - f_{i-1}}{2h} \quad e_i = -\frac{h^2}{6} f'''_i$$

Sommando  $f_{i+1}$  con  $f_{i-1}$  si ottiene un'espressione approssimata per  $f''_i$ :

$$f''_i = \frac{1}{h^2} (f_{i+1} - 2f_i + f_{i-1}) \quad e_i = -\frac{h^2}{12} f^{IV}_i$$

## 2 Formule approssimate per gli integrali di una funzione $f(x)$ , data per punti

Valgono le stesse ipotesi fatte al paragrafo 1 e si usano le stesse notazioni.

L'integrale di una funzione  $f(x)$  si fa integrando il suo sviluppo in serie di Taylor termine per termine.

Formule approssimate e gli errori corrispondenti possono essere determinati dalla serie integrata.

Indichiamo con  $I_1(i)$  l'integrale di una funzione  $f(x)$  da  $x = x_i$  a  $x = x_i + h$ , cioè l'integrale su 1 passo di larghezza  $h$  a partire da  $x_i$ :

$$I_1(i) = \int_{x_i}^{x_i+h} f(x) dx = \int_0^h f(z + x_i) dz \quad (z = x - x_i)$$

$$f(x_i + z) = f_i + z f'_i + \frac{z^2}{2} f''_i + \frac{z^3}{6} f'''_i + \dots$$

integrando termine per termine si ottiene:

$$\int_0^h f(z + x_i) dz = f_i \int_0^h dz + f'_i \int_0^h z dz + f''_i \frac{1}{2} \int_0^h z^2 dz + f'''_i \frac{1}{6} \int_0^h z^3 dz + \dots$$

$$I_1(i) = hf_i + \frac{h^2}{2} f'_i + \frac{h^3}{6} f''_i + \dots$$

Sostituiamo ad  $f'_i$  la formula a due punti della derivata:

$$f'_i = \frac{f_{i+1} - f_i}{h} - \left( \frac{h}{2} f''_i + \frac{h^2}{6} f'''_i + \dots \right)$$

Si ottiene

$$I_1(i) = hf_i + \frac{h^2}{2} \frac{f_{i+1} - f_i}{h} - \frac{h^2}{2} \left( \frac{h}{2} f''_i + \frac{h^2}{6} f'''_i \right) + \frac{h^3}{6} f''_i + \dots$$

Se si trascurano i termini in  $h^3 f''_i$  si ottiene una formula approssimata (a 2 punti) per l'integrale su 1 solo intervallo

$$I_{12}(i) = \frac{h}{2} (f_i + f_{i+1}) \quad e_i = -\frac{h^3}{12} f''_i$$

Questa è la nota regola del trapezio.

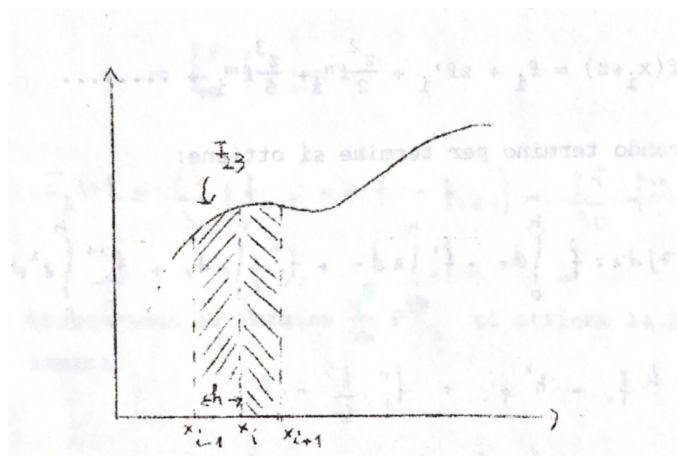
Per avere l'integrale su  $n$  passi si sommano gli integrali fatti su ciascun passo.

L'integrale fatto in questo modo non converge molto rapidamente.

Molto usata è la formula di Simpson:

$$I_{23}(i) = \frac{h}{3} (f_{i-1} + 4f_i + f_{i+1}) \quad e_i = -\frac{h^5}{90} f^{IV}_i$$

( $I_{23}$  significa integrale fatto a 3 punti su due intervalli).



converge rapidamente l'integrale su tutto l'intervallo dell'asse delle  $x$  che interessa calcolato iterando questa formula e sommando su tutti i passi.

Come si ottiene la formula di Simpson:

L'integrale va fatto su due passi simmetricamente localizzati intorno ad  $x_i$ :

$$\begin{aligned} I_2(i) &= \int_{x_{i-1}}^{x_{i+1}} f(x) dx = \int_{-h}^h f(x_i + z) dz \quad z = x - x_i \\ \int_{-h}^h f(x_i + z) dz &= \int_{-h}^h f_i dz + f'_i \int_{-h}^h z dz + \frac{1}{2} f''_i \int_{-h}^h z^2 dz + \\ &+ f'''_i \frac{1}{6} \int_{-h}^h z^3 dz + f^{IV}_i \frac{1}{24} \int_{-h}^h z^4 dz \dots \end{aligned}$$

Tenuto presente che l'integrale delle potenze dispari si annulla tra  $-h$  ed  $h$  si ottiene:

$$\begin{aligned} I_2(i) &= 2hf_i + 2\frac{h^3}{3} \frac{1}{2} f''_i + 2\frac{h^5}{5} \frac{1}{24} f^{IV}_i + \dots \\ I_2(i) &= 2hf_i + \frac{h}{3} (f_{i+1} - 2f_i + f_{i-1}) - \left( \frac{h^2}{12} f^{IV}_i \frac{h^3}{3} - \frac{h^5}{5 \cdot 12} f^{IV}_i \right) + \dots \\ I_2(i) &= \frac{h}{3} (f_{i+1} + 4f_i + f_{i-1}) - \frac{h^5}{90} f^{IV}_i \end{aligned}$$

Trascurando il termine  $\frac{h^5}{90} f^{IV}_i$  si ottiene la formula approssimata:

$$\begin{aligned} I_{23}(i) &= \frac{h}{3} (f_{i+1} + 4f_i + f_{i-1}) \\ e_i &= -\frac{h^5}{90} f^{IV}_i \end{aligned}$$

### 3 Interpolazione ed estrapolazione

Interpolazione significa calcolare  $f(x)$  in un punto  $x$  compreso tra due punti successivi della tabella. La valutazione di  $f(x)$  fuori dell'intervallo dei punti della tavola è chiamata estrapolazione.

La prima idea è approssimare la curva ad una retta, scrivere l'equazione della retta che passa per  $x_i$  ed  $x_{i+1}$  e quindi calcolare  $f(x)$ ; o in altre parole \*

$$f(x) \sim f_i + f'_i(x - x_i)$$

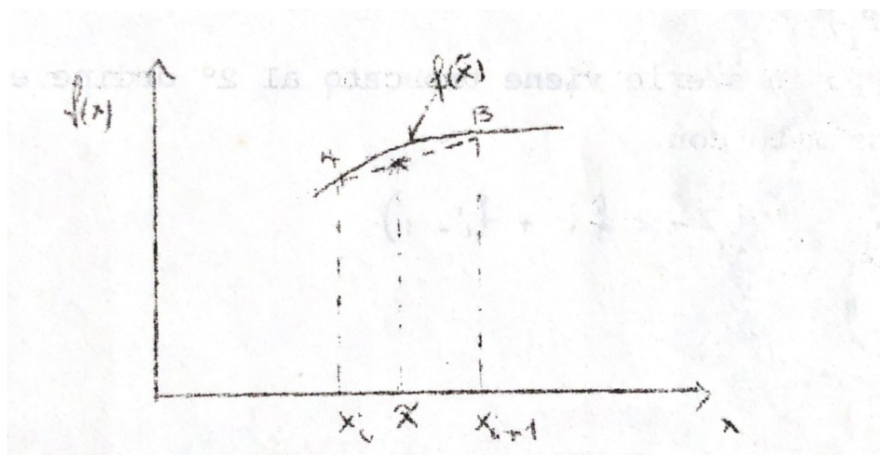
dove per  $f'_i$  si prende l'approssimazione

$$f'_i = \frac{f_{i+1} - f_i}{h}$$

Per ottenere le formule di interpolazione ed estrapolazione approssimiamo come al solito  $f(x)$  mediante una serie di Taylor troncata per es. al 2° ordine.

---

\* non è altro che lo sviluppo in serie di  $f(x)$  intorno ad  $x_i$  troncato al 1° ordine.



$$f(x) = f_i + f'_i(x - x_i) + \frac{1}{2}f''_i(x - x_i)^2 + \dots$$

Se tronciamo la serie di Taylor al 1° ordine e al posto di  $f'_i$  mettiamo l'espressione approssimata

$$f'_i = \frac{f_{i+1} - f_i}{h}$$

si ottiene la formula di interpolazione lineare a due punti.

$$f^{(1)}(x) = f_i + \frac{f_{i+1} - f_i}{h}(x - x_i)$$

(Significa che si è fatta passare una retta per i due punti noti  $f_{i+1}$ ,  $f_i$ ).

Se lo sviluppo in serie viene troncato al 2° ordine e  $f''_i$  viene approssimato con

$$f''_i = \frac{1}{h^2}(f_{i+1} - 2f_i + f_{i-1})$$

$f'_i$  con  $\frac{f_{i+1} - f_{i-1}}{2h}$

si ottiene la formula di interpolazione quadratica a tre punti:

$$f^{(2)}(x) = f_i + \frac{f_{i+1} - f_{i-1}}{2h}(x - x_i) + \frac{1}{2h^2}(f_{i+1} - 2f_i + f_{i-1})(x - x_i)^2$$

oppure

$$f^{(2)}(x) = f_i + \frac{f_{i+1} - f_i}{h}(x - x_i) + \frac{1}{2h^2}(f_{i+1} - 2f_i + f_{i-1})(x - x_i)^2$$

Estrapolazione:  $x > x_i$

$$f'_i = \frac{f_i - f_{i-1}}{h}$$

$$f^{(1)}(x) = f_i + \frac{f_i - f_{i-1}}{h}(x - x_i)$$

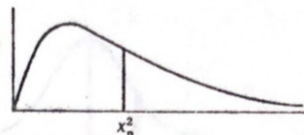


## APPENDIX

345

## Appendix IV

PERCENTILE VALUES ( $\chi_p^2$ )  
for  
THE CHI-SQUARE DISTRIBUTION  
with  $\nu$  degrees of freedom  
(shaded area =  $p$ )



$\nu$	$\chi_{.995}^2$	$\chi_{.99}^2$	$\chi_{.975}^2$	$\chi_{.95}^2$	$\chi_{.9}^2$	$\chi_{.75}^2$	$\chi_{.50}^2$	$\chi_{.25}^2$	$\chi_{.10}^2$	$\chi_{.05}^2$	$\chi_{.025}^2$	$\chi_{.01}^2$	$\chi_{.005}^2$
1	7.88	6.63	5.02	3.84	2.71	1.32	.455	.102	.0158	.0039	.0010	.0002	.0000
2	10.6	9.21	7.38	5.99	4.61	2.77	1.39	.575	.211	.103	.0506	.0201	.0100
3	12.8	11.3	9.35	7.81	6.25	4.11	2.37	1.21	.584	.352	.216	.115	.072
4	14.9	13.3	11.1	9.49	7.78	5.39	3.36	1.92	1.06	.711	.484	.297	.207
5	16.7	15.1	12.8	11.1	9.24	6.63	4.35	2.67	1.61	1.15	.831	.554	.412
6	18.5	16.8	14.4	12.6	10.6	7.84	5.35	3.45	2.20	1.64	1.24	.872	.676
7	20.3	18.5	16.0	14.1	12.0	9.04	6.35	4.25	2.83	2.17	1.69	1.24	.989
8	22.0	20.1	17.5	15.5	13.4	10.2	7.34	5.07	3.49	2.73	2.18	1.65	1.34
9	23.6	21.7	19.0	16.9	14.7	11.4	8.34	5.90	4.17	3.33	2.70	2.09	1.73
10	25.2	23.2	20.5	18.3	16.0	12.5	9.34	6.74	4.87	3.94	3.25	2.56	2.16
11	26.8	24.7	21.9	19.7	17.3	13.7	10.3	7.58	5.58	4.57	3.82	3.05	2.60
12	28.3	26.2	23.3	21.0	18.5	14.8	11.3	8.44	6.30	5.23	4.40	3.57	3.07
13	29.8	27.7	24.7	22.4	19.8	16.0	12.3	9.30	7.04	5.89	5.01	4.11	3.57
14	31.3	29.1	26.1	23.7	21.1	17.1	13.3	10.2	7.79	6.57	5.63	4.66	4.07
15	32.8	30.6	27.5	25.0	22.3	18.2	14.3	11.0	8.55	7.26	6.26	5.23	4.60
16	34.3	32.0	28.8	26.3	23.5	19.4	15.3	11.9	9.31	7.96	6.91	5.81	5.14
17	35.7	33.4	30.2	27.6	24.8	20.5	16.3	12.8	10.1	8.67	7.56	6.41	5.70
18	37.2	34.8	31.5	28.9	26.0	21.6	17.3	13.7	10.9	9.39	8.23	7.01	6.26
19	38.6	36.2	32.9	30.1	27.2	22.7	18.3	14.6	11.7	10.1	8.91	7.63	6.84
20	40.0	37.6	34.2	31.4	28.4	23.8	19.3	15.5	12.4	10.9	9.59	8.26	7.43
21	41.4	38.9	35.5	32.7	29.6	24.9	20.3	16.3	13.2	11.6	10.3	8.90	8.03
22	42.8	40.3	36.8	33.9	30.8	26.0	21.3	17.2	14.0	12.3	11.0	9.54	8.64
23	44.2	41.6	38.1	35.2	32.0	27.1	22.3	18.1	14.8	13.1	11.7	10.2	9.26
24	45.6	43.0	39.4	36.4	33.2	28.2	23.3	19.0	15.7	13.8	12.4	10.9	9.89
25	46.9	44.3	40.6	37.7	34.4	29.3	24.3	19.9	16.5	14.6	13.1	11.5	10.5
26	48.3	45.6	41.9	38.9	35.6	30.4	25.3	20.8	17.3	15.4	13.8	12.2	11.2
27	49.6	47.0	43.2	40.1	36.7	31.5	26.3	21.7	18.1	16.2	14.6	12.9	11.8
28	51.0	48.3	44.5	41.3	37.9	32.6	27.3	22.7	18.9	16.9	15.3	13.6	12.5
29	52.3	49.6	45.7	42.6	39.1	33.7	28.3	23.6	19.8	17.7	16.0	14.3	13.1
30	53.7	50.9	47.0	43.8	40.3	34.8	29.3	24.5	20.6	18.5	16.8	15.0	13.8
40	66.8	63.7	59.3	55.8	51.8	45.6	39.3	33.7	29.1	26.5	24.4	22.2	20.7
50	79.5	76.2	71.4	67.5	63.2	56.3	49.3	42.9	37.7	34.8	32.4	29.7	28.0
60	92.0	88.4	83.3	79.1	74.4	67.0	59.3	52.3	46.5	43.2	40.5	37.5	35.5
70	104.2	100.4	95.0	90.5	85.5	77.6	69.3	61.7	55.3	51.7	48.8	45.4	43.3
80	116.3	112.3	106.6	101.9	96.6	88.1	79.3	71.1	64.3	60.4	57.2	53.5	51.2
90	128.3	124.1	118.1	113.1	107.6	98.6	89.3	80.6	73.3	69.1	65.6	61.8	59.2
100	140.2	135.8	129.6	124.3	118.5	109.1	99.3	90.1	82.4	77.9	74.2	70.1	67.3

Source: Catherine M. Thompson, *Table of percentage points of the  $\chi^2$  distribution*,  
Biometrika, Vol. 32 (1941), by permission of the author and publisher.