3-Trattamento delle Misure topografiche

Premessa

Tutto il nostro **Trattamento** era in passato chiamato “Teoria delle Misure” ed era basato sulla gestione di quelli che si chiamavano “*errori accidentali*” (adesso “fluttuazioni accidentali”) e per essi effettivamente serve. Serve anche per i cosi detti *“errori grossolani*” ma meglio ricorrere per questi al *metodo delle reiterazioni dei pesi* per eliminarli o ridurli o ad altri test e processi.

Serve poco o nulla per gli “*errori sistematici*”: ma purtroppo anche questi intervengono quando si misura, così come gli altri errori e sotto certi aspetti sono anche più pericolosi (possono avere anche valori più elevati). Occorrerebbe trovare un modo per tenere conto di tutto e non solamente di una parte. La Statistica quindi sotto questo punto di vista non serve o serve a poco.

Si dovrebbe agire logicamente in questo modo:

1. Vedere se le misure possono essere affette da errori sistematici
2. Se sì, vedere come fare per eliminare la sistematicità nell’organo di misura (ma l’errore sistematico di taratura di un campione non può essere eliminato; si deve ricorrere alla ripetizione con campioni diversi, indipendenti)
3. Se no, applicare la teoria statistica che viene qui descritta per eliminare gli errori (fluttuazioni) accidentali.

Occorre fare anche attenzione a pensare che le distribuzioni di errore siano solamente quelle gaussiane e fare attenzione ai casi di variabili tra loro dipendenti. Si vedrà nel dettaglio:

- Il concetto di grandezza e della sua **misura diretta**
- Si introducono i concetti di: *variabile statistica* ad una o più dimensioni e della rappresentazione sintetica; di *eventi aleatori* o stocastici, di *variabili casuali* e di alcune *distribuzioni* importanti
- Si affronta il problema della stima dei parametri, o dell’i*nferenza* statistica, in una popolazione composta da un numero limitato di estrazioni a caso, con il *principio di massima verisimiglianza* che porta al *principio dei minimi quadrati*, per ottenere media e varianza empiriche.
- Si esamina il concetto di *media ponderata*, con un esempio
- Si passa quindi alle **misure indirette**, distinguendo i tre casi classici (con l’importante caso della sovrabbondanza, il 3 °) , nei sottocasi lineari e non .
- Si approfondisce il calcolo della varianza dell’unità di peso e si fornisce la risoluzione, con il pratico metodo reiterativo, di un piccolo sistema (3 equazioni in 3 incognite)

3 -INDICE TRATTAMENTO delle MISURE

Introduzione

Individuare le LEGGI della NATURA

Analisi del FENOMENO REALE

Schematizzazione di
alcune sue caratteristiche

Nasce il FENOMENO ASTRATTO

Fare esempio dall'OTTICA:
definizione delle ABERRAZIONI

Affinità procedurale

LE GRANDEZZE

NON esiste un’unica misura vera
di una GRANDEZZA

Cos’è una GRANDEZZA
(secondo Euclide,
secondo Russell, ecc.)

Introduzione del concetto
di QUANTITA’

Classi di grandezza di DIVISIBILITA’
Classi di grandezza di DISTANZA

Le grandezze si riuniscono
in CLASSI

Ad es. 1 km, 1 m, ecc. sono grandezze
della Classe delle Lunghezze

Possibilità di stabilire CONFRONTI
tra grandezze della stessa classe, ecc.

Corrispondenza tra rapporti di
grandezze e numeri reali positivi

Introduzione

MISURARE

Misurare una Grandezza = ottenere una CORRISPONDENZA con dei numeri

Operazione di MISURA estranea alla Grandezza. E' un mezzo umano di rappresentare la grandezza

Come si fa una MISURA DIRETTA

Definire le unità campioni u.c. Si sommano le u.c., si confrontano con la grandezza e si fa un'operazione di conteggio

Sperimentalmente le misure cambiano nel tempo

Distinguere tra ERRORI GROSSOLANI, SISTEMATICI, ACCIDENTALI

GRANDEZZE

- **Euclide** (V libro degli “Elementi”):

1. le grandezze vengono ripartite in CLASSI, formate ciascuna di ENTI;
2. si stabiliscono opportuni procedimenti di “confronto” tra due grandezze della stessa classe;
3. il *rapporto* tra due grandezze è sempre un numero reale positivo (*misura*);
4. il rapporto è 1 se le due grandezze sono uguali.

- **Eudosso** da Cnido (membro dell’Accademia platonica)

1. La grandezza non è concepita come un numero, ma come un’ entità geometrica che può variare in modo continuo
2. Definisce il concetto di rapporto tra grandezze e quello di uguaglianza o proporzione tra due rapporti di grandezze (sia razionali che irrazionali) .

- **Russell** (“I principi della matematica”):

1. le grandezze sono concetti astratti che comprendono una molteplicità di termini;
2. se una grandezza si “particolarizza”, diventa una QUANTITA’ di grandezza.

Le “*quantità*” non possono essere maggiori o minori (relazioni valide solo per grandezze della stessa classe). Hanno solamente la proprietà dell’uguaglianza.

GRANDEZZE secondo Russell

CLASSI di GRANDEZZE di

DIVISIBILITA':
un Tutto (o un Insieme finito o limitato) si correla con il numero di parti semplici che lo compongono.
Ha significato la "somma".

DISTANZA:
sono quelle grandezze che devono avere un punto o un evento di riferimento (ad es. il tempo).
Non ha significato l'operazione di "somma".

COME SI FA UNA MISURA DIRETTA:

1. Definire le unità campione (u.c.);
2. Si sommano le u.c. e si confrontano con la grandezza da misurare;
3. Si fa un operazione di conteggio delle u.c.

FLUTTUAZIONI ACCIDENTALI

Introduzione del concetto di
FLUTTUAZIONE ACCIDENTALE:
Tutte misure vere della stessa grandezza

Popolazioni di MISURE POSSIBILI

Estrarre un unico valore rappresentativo:
la **MEDIA**

Accoppiare la valutazione numerica
della variabilità dei dati:
la **VARIANZA**

Non è facile dare la definizione di scienza STATISTICA.

Ma così è anche per altre scienze.

Ogni scienza è in continuo progresso, quindi non la si può circoscrivere e quindi definire perfettamente.

Senza dubbio la STATISTICA richiede “conteggi”.

Permette di dare un valore quantitativo (=misura) ai fenomeni collettivi.

FENOMENO COLLETTIVO:

- relativo ad una collettività di casi singoli (popolazione di individui)
- relativo ad un caso singolo, su cui si effettua una collettività di osservazioni

Quasi tutte le scienze fanno ricorso alla **STATISTICA**: ad esempio

- scienze che hanno per oggetto una collettività di enti, della quale studiano un aspetto
- scienze sperimentali, ecc.

Occorre dire ancora qualcosa sulle “**corrispondenze numeriche**” o “**tabelle**”, in generale.
Esse possono essere relative a:
- distribuzioni
- non distribuzioni

Le **distribuzioni** possono essere:
- di frequenza
- di quantità

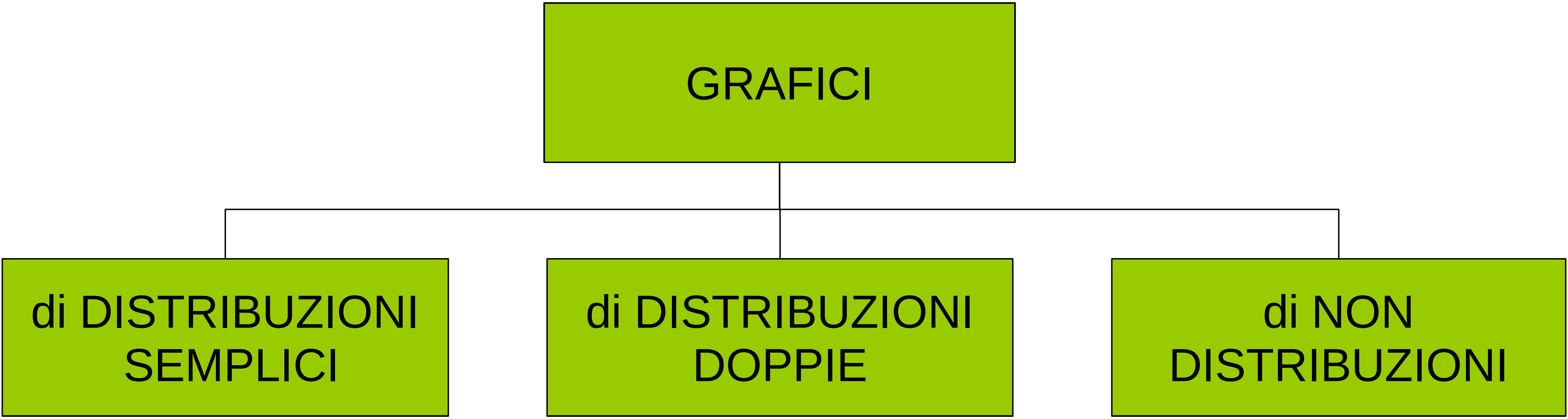
Le **non distribuzioni** si distinguono in:
- serie storiche
  - **di stato**
  - **di movimento**
- serie territoriali

Esempio di **distribuzione di quantità**:

La Banca XY ha conteggiato che nel 2005 si sono avuti:

•	cambiali ordinarie	€ 3.284.500
•	tratte	€ 5.242.825
•	assegni bancari	€ 2.854.300

GRAFICI



GRAFICI

- di DISTRIBUZIONI SEMPLICI
- di DISTRIBUZIONI DOPPIE
- di NON DISTRIBUZIONI

GRAFICI

DISTRIBUZIONI SEMPLICI

- Secondo un **carattere di qualsiasi tipo**:
    - diagrammi simbolici o pictogrammi
    - grafici:
        - ➢ a segmenti
        - ➢ a nastri
        - ➢ a colonne
    - areogrammi

- Secondo un **carattere ordinato**:
    - rettilineo → ISTOGRAMMi
    - ciclico:
        - ➢ grafico a raggi
        - ➢ grafico a settori circolari di uguale ampiezza

- Secondo un **carattere quantitativo rettilineo discreto**

GRAFICI di

DISTRIBUZIONI DOPPIE

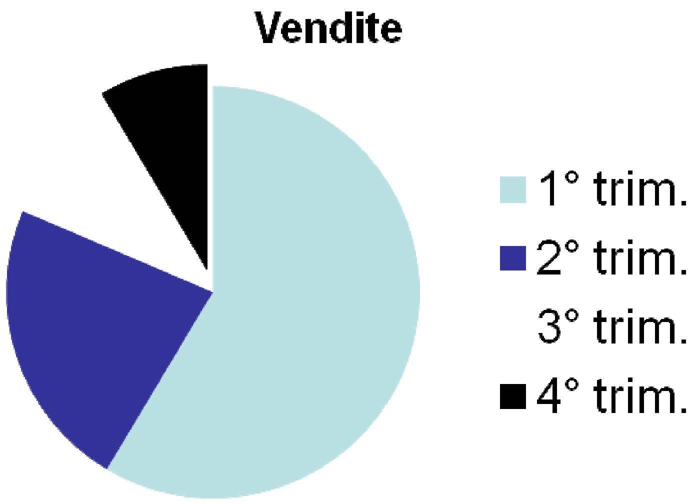
- A colonne suddivise
- A cerchi suddivisi
- Cartodiagrammi a rappresentazioni suddivise:
  - lineare
  - areale
- Stereogrammi

GRAFICI di

NON DISTRIBUZIONI

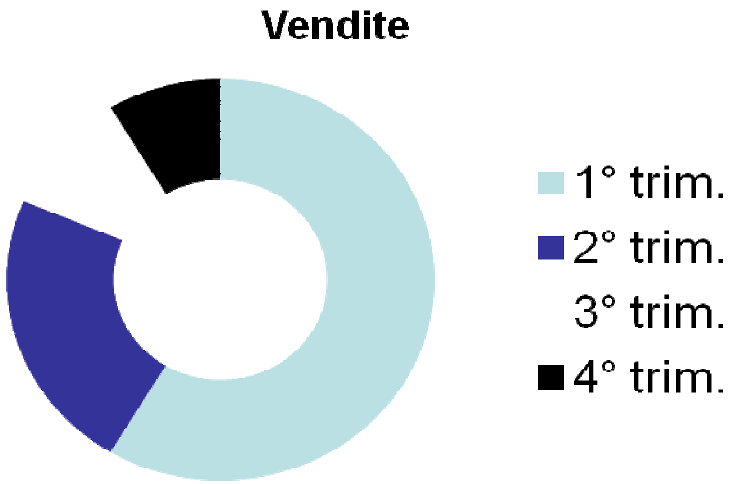
- Delle serie territoriali
- Cartogrammi:
  - a tratteggio
  - a colonne
  - a cerchi
  - misto (colonne e cerchi)
- Delle serie storiche

Vari tipi di diagrammi: alcuni esempi

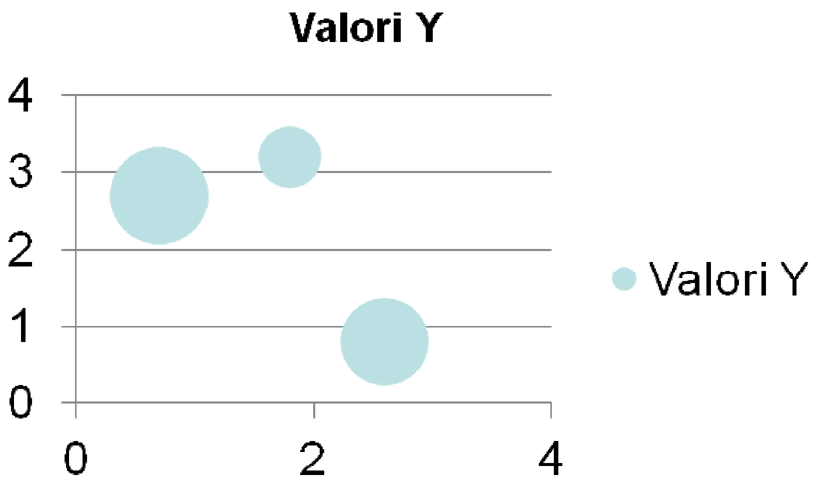


1-a torta

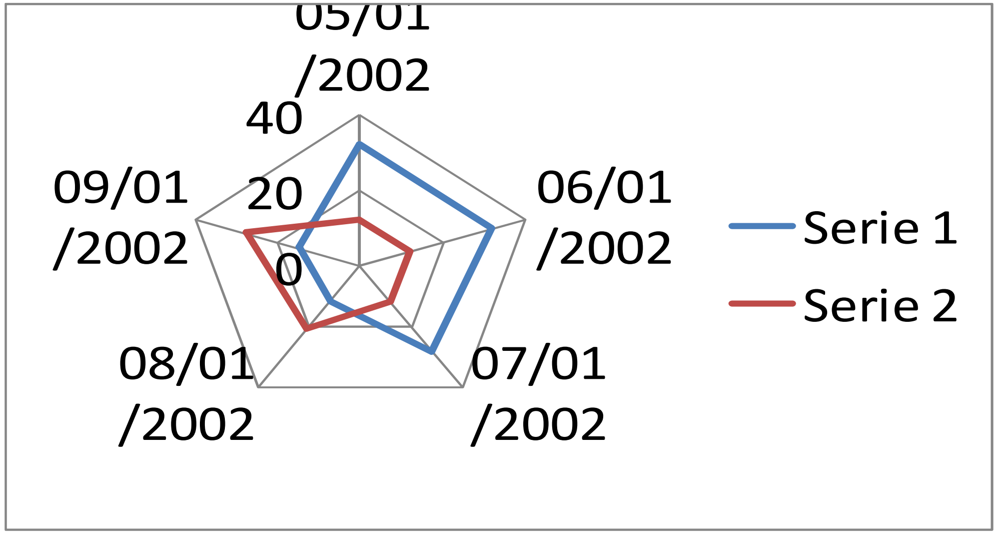
FIG.1



2-ad anello



3-a bolle

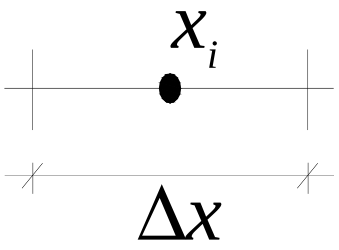


4-radar

VARIABILE STATISTICA AD UNA DIMENSIONE

Si prenda in esame un insieme ben determinato di *individui* (**popolazione**), caratterizzati da un *attributo* X che può assumere *valori argomentali* $x_1, x_2, \ldots.., x_n$ diversi.

Classificare la popolazione secondo l'attributo X vuole dire esaminare tutti gli individui della popolazione e determinare,per ognuno di essi, il valore argomentale x_i; quindi si procede ad un raggruppamento di valori argomentali (e quindi di individui), definendo un *intervallo di raggruppamento* Δx: si assume che tutti i valori argomentali compresi in questo intervallo, assumano il valore argomentale x_i intermedio.



Si potrà allora scrivere una corrispondenza numerica:

$$X\begin{cases} x_1\, x_2 \,.....\, x_n \\ F_1\, F_2 \,.....\, F_n \end{cases}$$

dove gli x_i sono ordinati in senso crescente e gli F_i rappresentano il numero di individui che hanno il valore argomentale x_i (*frequenza assoluta*). E' evidente che (essendo N= numero totale individui):

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

Si possono introdurre le frequenze relative: $f_i = \frac{F_i}{N}$

$$X\begin{cases} x_1\, x_2 \,.....\, x_n \\ f_1\, f_2 \,.....\, f_n \end{cases} \qquad \sum_{i=1}^{n} f_i = 1$$

Questa corrispondenza numerica si chiama **variabile statistica ad 1** dimensione (ma si possono definire, in maniera analoga, anche se più complessa, variabili a 2 o più dimensioni).

Rappresentazione grafica della v.s. ad 1 dimensione:

ISTOGRAMMA

Si divida l’asse x secondo l’intervallo di raggruppamento Δx_i; la mezzeria di ogni intervallo sia in corrispondenza del valore x_i della variabile. Per ogni intervallo costruiamo un rettangolo di altezza y_i tale che:

$$y_i * \Delta x = f_i$$

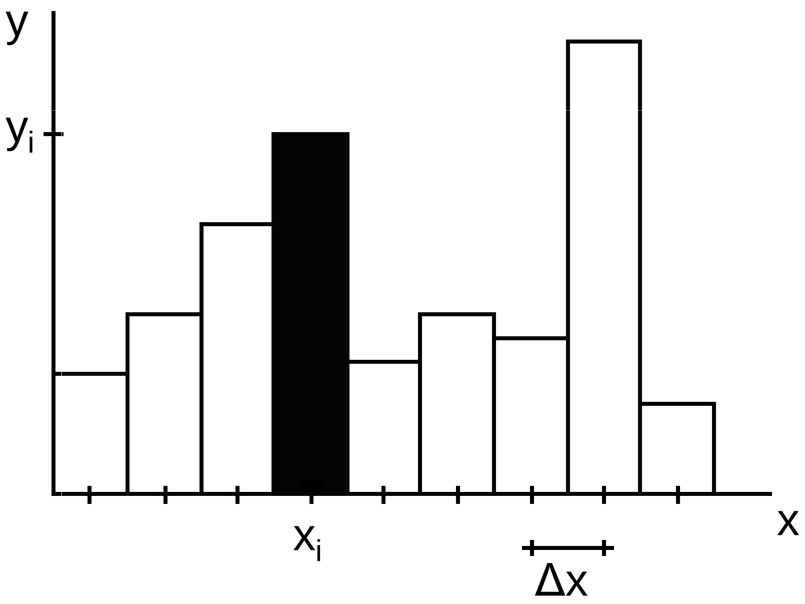
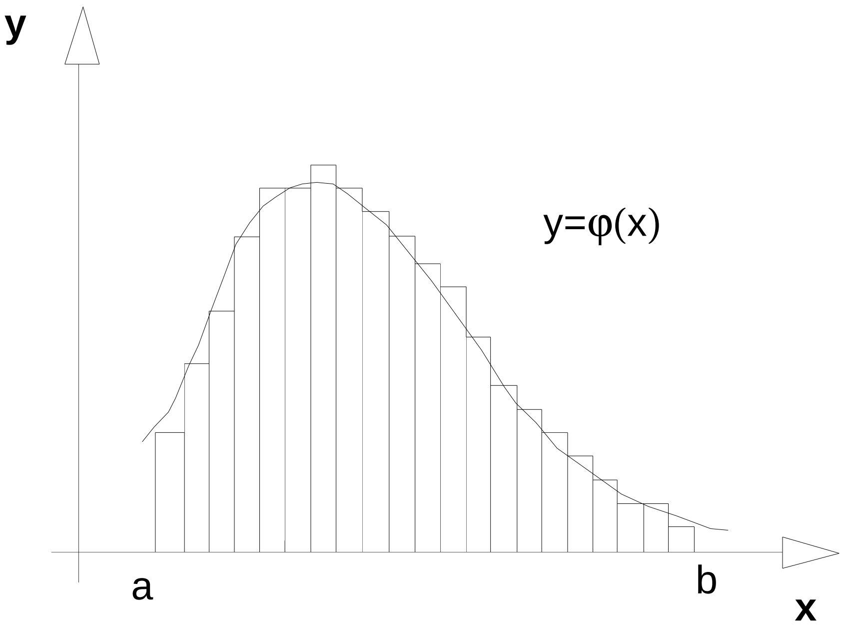


FIG.2

L’area racchiusa da ogni rettangolo rappresenta la frequenza relativa f_i , corrispondente al valore argomentale x_i.

L’istogramma dà una rappresentazione visiva di come i valori argomentali siano distribuiti tra gli individui della popolazione.

Per un numero N di individui molto elevato, con valori x_i che si susseguono ad intervalli molto ristretti, se non esistono discontinuità nelle frequenze di gruppi vicini, l'istogramma può essere rappresentato con una curva continua $y = \varphi(x)$ chiamata **curva di frequenza** (sempre che non si introducano elementi estranei o incongruenti con il tipo di variabile esaminato).



$$\int_a^b \varphi(x) \cdot dx = 1$$

FIG.3

RAPPRESENTAZIONE SINTETICA di una Variabile Statistica

Occorre introdurre il concetto di MOMENTO, che dà informazioni quantitative globali sulla distribuzione.

Si chiama *momento del* K^{mo} *grado* della variabile:

$$m_k(x) = \sum_{i=1}^{n} x_i \cdot f_i = \sum x_i^k \frac{F_i}{N} \cong \frac{1}{N} \sum_{r=1}^{N} x_r^k$$

Per variabile continua:

$$m_k(x) = \int_a^b x^k \cdot \varphi(x) dx$$

Esaminiamo alcuni momenti rappresentativi:

- Momento di 1° grado:

$$m_1(x) = \sum_{i=1}^{n} x_i \cdot f_i$$

MEDIA

Indica genericamente quale è la posizione della variabile sull’asse delle x, perché i valori argomentali x_i si distribuiscono intorno al valore della media.

- Momento di 2° grado:

$$m_2(x) = \sum_{i=1}^{n} x_i^2 \cdot f_i$$

VALORE
QUADRATICO MEDIO

(questo valore ha, per noi, poco significato pratico).

- Si consideri una nuova variabile statistica, così definita:

$$v_i = x_i - M$$

Variabile SCARTO

Questa variabile ha la stessa distribuzione della variabile X:

$$v \begin{cases} v_1 v_2 \ldots . \; v_n \\ f_1 \; f_2 \ldots . \; f_n \end{cases} \qquad \sum_{i=1}^{n} f_i = 1$$

- momento di 1° grado: $m_1(v) = \sum_{i=1}^{n} v_i f_i = \sum_{i=1}^{n} (x_i - M) f_i = 0$
- momento di 2° grado: $m_2(v) = \sum_{i=1}^{n} f_i (x_i - M)^2 = \sum_{i=1}^{n} f_i v_i^2$

VARIANZA
σ^2

La varianza indica come la variabile è distribuita intorno alla media: è, cioè, un indice di dispersione.

Per quanto riguarda il calcolo della media:

$$M = m_1(x) = \sum_{i=1}^{n} x_i f_i = \sum_{i=1}^{n} \frac{x_i F_i}{N} = \frac{1}{N}\sum_{r=1}^{N} x_r$$

Per quanto riguarda la varianza σ²:

$$m_2(v) = \sigma^2 \sum_{i=1}^{n} (x_i - M)^2 \cdot f_i = \sum_{i=1}^{n} \frac{(x_i - M)^2 F_i}{N}$$

$$\sum_{i=1}^{n} (x_i - M)^2 \cdot F_i = \boxed{\text{devianza della distribuzione}}$$

$$\sigma^2 = \sum_{i=1}^{n} (x_i - M)^2 \cdot f_i = \sum_{i=1}^{n} \left(x_i^2 + M^2 - 2x_i M\right) \cdot f_i =$$

$$= \sum_{i=1}^{n} x_i^2 f_i + M^2 \sum_{i=1}^{n} f_i - 2M \sum_{i=1}^{n} x_i f_i = m_2(x) - M^2$$

$$\sigma = \boxed{\text{Devianza standard o s.q.m.}}$$

VALORI CENTRALI

In statistica, oltre la media, si definiscono altri valori centrali nella distribuzione:

- MEDIANA: è quel valore dell’argomento in corrispondenza del quale l’ordinata ha la caratteristica di dividere la curva di frequenza in due parti di aree equivalenti;
- MODA o NORMA: è quel valore dell’argomento per il quale si ha il massimo di frequenza (ordinata) f_i.

Si possono avere più mode:

- unimodale
- bimodale
- trimodale
- ecc.

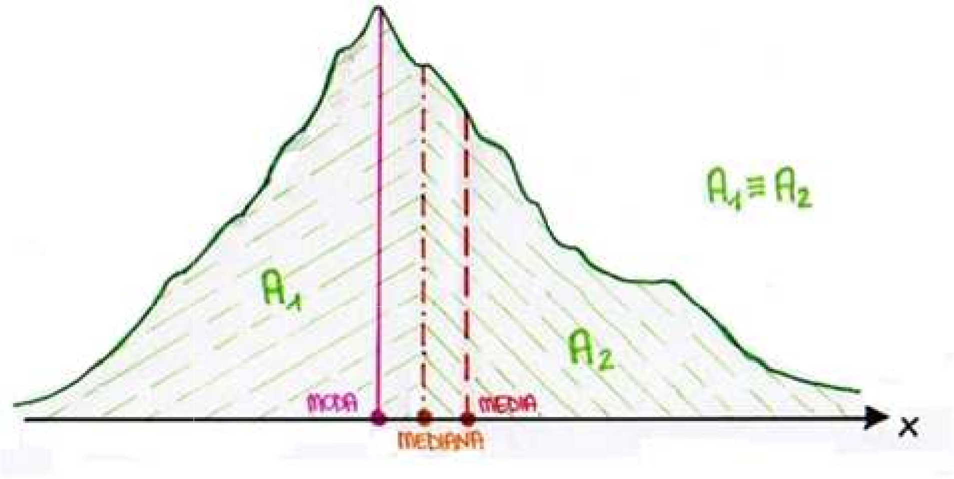


FIG.4

DISUGUAGLIANZA di TCHEBYCHEFF
(o teorema di BYENAYME'-TCHEBYCHEFF)

La varianza σ^2 è un indice di dispersione abbastanza significativo: si può infatti dimostrare che la maggior parte dei valori argomentali x_i è contenuta entro i limiti M±3σ, qualunque sia la distribuzione. Infatti, scriviamo:

$$\sigma^2 = \sum_{i=1}^{n} v_i^2 f_i = v_1^2 f_1 + v_2^2 f_2 + ... + v_m^2 f_m + ... + v_{n-1}^2 f_{n-1} + v_n^2 f_n$$

(le v_i^2 siano ordinate in ordine crescente).

Si prenda un valore intermedio v_m^2 e ad esso diamo il valore c^2. Si azzerino tutti i valori v_i^2 minori di c^2. Si pongano uguali a c^2 tutti quelli maggiori. Sarà allora:

$$\sigma^2 \geq c^2\left(f_m + f_{m+1} + ... + f_n\right)$$

Si pone $f_m + f_{m+1} + ... + f_n = f^*$ da cui $\sigma^2 \geq c^2(1 - f_*)$.

Si ricavi f_* (frequenza relativa degli scarti inferiori a c):

Imponendo c=λσ si ottiene:

$$\boxed{f_* \geq 1 - \frac{1}{\lambda^2}}$$

La disuguaglianza di Tchebycheff:

$$f_* \geq 1 - \frac{1}{\lambda^2}$$

è significativa per λ>1. In particolare:

- per λ=2 $$f_* \geq 1 - \frac{1}{4} = 0,75$$
- per λ=3 $$f_* \geq 1 - \frac{1}{9} = 0,89$$

cioè: per qualsiasi variabile statistica, almeno il 75% degli scarti ha un valore inferiore o uguale a ±2σ, e almeno l'89% ha un valore inferiore o uguale a ±3σ.

Si può dire anche che la frequenza relativa dei termini che sono non interni ad un intorno della media, non può superare un certo limite (è un caso particolare del più generale teorema di Markov).

FUNZIONE CUMULATIVA DI FREQUENZA

E’ la variabile originaria:

$$X \begin{cases} x_1\ x_2\ \dots\ x_n \\ f_1\ f_2\ \dots\ f_n \end{cases}$$

Si definisce funzione cumulativa di frequenza la seguente:

$$Y \begin{cases} x_1\ x_2\ \dots\ x_n \\ t_1\ t_2\ \dots\ t_n \end{cases}$$

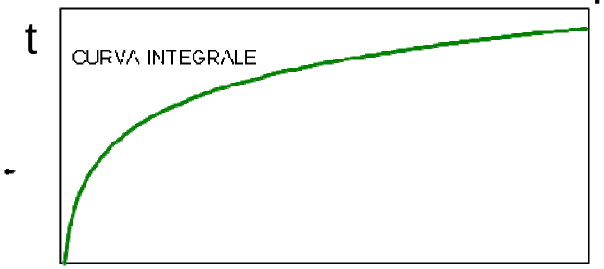


FIG.5

con: $t_1 = f_1$

$t_2 = f_1 + f_2$

$t_n = f_1 + f_2 + \dots + f_n = \Sigma f_i = 1$

Il valore t_i rappresenta la frequenza relativa di tutti gli individui che hanno un valore argomentale inferiore od uguale a x_i.

Le curve cumulative di frequenza sono assai utili per trovare rapidamente il numero di individui che hanno valori argomentali compresi in un intervallo qualsiasi.

EVENTO ALEATORIO O STOCASTICO

E’ un evento del quale non si può prevedere la modalità di uscita.
Ad esempio. pensare ad una moneta che viene lanciata: non è mai possibile prevedere se, a lancio effettuato, si verificherà l’evento “testa” o l’evento “croce”.

ESTRAZIONE a CASO

Un particolare evento aleatorio è l’estrazione a caso. Si pensi ad un’urna con un certo numero N_1 di palline, tutte uguali tra loro, ma con scritto un numero distintivo su ogni pallina. All’atto dell’estrazione tutte le palline sono identiche e chi effettua l’estrazione non deve poter scegliere e cioè non deve esserci ragione di estrazioni preferenziali.
Si estrae una pallina, si legge il numero sopra scritto e la si rigetta all’interno dell’urna.
Si continua in queste estrazioni un numero di volte N_2, grande a piacere ($N_2 >> N_1$). Si ottiene la corrispondenza:

$$X \left\{ \begin{array}{llll} x_1 & x_2 & \dots & x_n \\ f'_1 & f'_2 & \dots & f'_n \end{array} \right.$$

che si chiama VARIABILE STATISTICA CAMPIONE.

LEGGE EMPIRICA del CASO

Se incrementiamo il numero N_2 di estrazioni a caso, sino a valori molto grandi, ci si accorge che:

- i valori argomentali x_i della variabile campione assumeranno tutti i valori propri della popolazione assoggettata alle estrazioni a caso;
- le frequenze f_i corrispondenti ad ogni valore argomentale tendono a stabilizzarsi ed a convergere verso il valore proprio della frequenza di quell'argomento nella popolazione, quale si sarebbe ottenuto facendo un'indagine di tipo statistico.

Da questa "legge" si possono trarre due conclusioni:

1) **non** si può mai prevedere l'estrazione di un determinato individuo o di un valore argomentale. **E'** prevedibile, invece, la distribuzione, e cioè: dopo un certo numero N_2 di estrazioni a caso, si può prevedere la distribuzione dei valori argomentali (ovvero le frequenze) della variabile campione;
2) Esiste la possibilità di *postulare* l'esistenza di una popolazione. Cioè tutte le volte che l'indagine su un fenomeno aleatorio mostra una "stabilità statistica", si può essere autorizzati ad immaginare una popolazione "ignota", ma "determinata", che fa da sfondo al fenomeno aleatorio.

Questa popolazione si chiama POPOLAZIONE POSSIBILE.

VARIABILI CASUALI

Se si effettua una serie sempre crescente di estrazioni a caso da popolazione di composizione nota, o da popolazione possibile, si ottiene una specie di VARIABILE LIMITE, caratterizzata ancora da una serie di valori argomentali, in corrispondenza dei quali si definiscono delle “frequenze limite”: solo che, essendo il numero di estrazioni a caso infinito, la variabile che così si definisce non ha il carattere di un’operazione già eseguita e cristallizzata, ma di una VARIABILE TEORICA in divenire, definibile cioè dopo un numero di estrazioni infinito. A questa variabile si da’ il nome di VARIABILE CASUALE ed alle frequenze limite, relative ai valori argomentali, si da’ il nome di PROBABILITA’:

$$X_c \begin{cases} x_1 \;\; x_2 \ldots \;\; x_n \\ \\ p_1 \;\; p_2 \ldots \;\; p_n \end{cases} \qquad\qquad \Sigma \; p_i \; = \; 1$$

PROBABILITÀ

• definizione *a posteriori*: è il limite a cui tende la frequenza di un dato valore argomentale della variabile campione, quando il numero di prove tende all'infinito

• definizione *a priori* :   è il rapporto tra

$$= \frac{\text{numero casi favorevoli al presentarsi dell'evento}}{\text{numero di casi possibili}}$$

Noi scegliamo la "definizione a posteriori" che presuppone che siano effettivamente state fatte delle prove e che da queste si giunga alla definizione della variabile casuale.

Per variabile casuale continua si definisce:

$$dp = f(x) \cdot dx$$ essendo f(x) = densità di probabilità

$$F(x) = \int_0^x f(x) \cdot dx$$ essendo F(x) = funzione di distribuzione

E per i MOMENTI :
$$m_k(x) = \int_a^b x_i^k f(x) dx$$

Proprietà delle probabilità composte:

se un evento aleatorio A risulta dal concorso di due eventi aleatori indipendenti B e C di probabilità p_b e p_c, la probabilità dell’evento A è data **dal prodotto** delle probabilità di B e C:

$$p_a = p_b \cdot p_c$$

Se l’ordine di uscita degli eventi C e B <u>non</u> ha importanza, si ha:

$$p'_a = 2 \cdot p_b \cdot p_c$$

In generale:

$$p'_a = n! \cdot p_b \cdot p_c \cdot \ldots$$

essendo *n!* in numero delle disposizioni possibili degli n eventi attesi.

Proprietà delle probabilità totali:

se un evento aleatorio A può presentarsi con modalità diverse B e C ognuna delle quali ha probabilità p_b e p_c, la probabilità di A è data dalla **somma** delle probabilità:

$$p_a = p_b + p_c$$

Alcuni esempi

1°caso: ci sono 4 palline colorate in modo diverso: come possono essere sistemate in fila? La 1° posizione può essere occupata da una pallina qualsiasi (4 diversi modi di occupare la 1° posizione); ci sono 3 modi di occupare la 2° posizione, 2 modi di occupare la 3° posizione, 1 modo di occupare la 4 posizione. Quindi 4x3x2x1= 24

R.: Ci sono 24 modi di sistemare le palline.

2°caso: si vogliono sistemare 7 persone attorno ad un tavolo rotondo

a) ognuno può sedere in un posto qualsiasi; la 1° persona può sedere in un posto qualsiasi, le restanti 6 persone in 6! = 720 modi diversi

b) due persone non devono essere vicine: le 2 persone possono essere sistemate in 2 ! modi diversi e le altre 5 in 5! modi diversi. Quindi 2! 5! =240 modi diversi di sistemare le 7 persone, con due che siedono vicini.

R.:si dovrà quindi fare 720-240 = 480 modi di sistemare le persone, tenendo conto della restrizione imposta.

3°caso : scegliere tra un totale di 9 persone, un comitato composto da sole 5 persone.

R.: $\binom{9}{5}$ = 9!/5! =9x8x7x6x5/ 5! = 126 modi di scegliere il comitato

4° caso : in un dado che probabilità ha l'uscita del 3 o del 5 ?

R.: 1/6 + 1/6 = 1/3 (probabilità più vantaggiosa, totale) .

5° caso : qual è la probabilità di uscita del 3 e del 5 ?

R.: 1/6 x 1/6 = 1/ 36 (probabilità composta).

EVENTO ALEATORIO COMPOSTO con 2 o più eventi aleatori. OPERAZIONI con VARIABILI CASUALI

Se un fenomeno aleatorio risulta dalla concomitanza di due o più fenomeni aleatori, il problema è quello di definire la variabile casuale relativa al fenomeno aleatorio composto.

Devono essere note:

- le variabili casuali componenti
- la modalità matematica con cui si ottiene il risultato

Si abbia la variabile casuale:

$$X \begin{cases} x_1 & x_2 & \dots & x_n \\ p_1 & p_2 & \dots & p_n \end{cases} \qquad \sum_{j=1}^{n} p_j = 1$$

e la variabile casuale

$$Y \begin{cases} y_1 & y_2 & \dots & y_m \\ g_1 & g_2 & \dots & g_m \end{cases} \qquad \sum_{k=1}^{n} g_k = 1$$

Vogliamo ottenere la variabile casuale “somma” delle due: cioè, riconducendoci allo schema dell’estrazione a caso, supponiamo di estrarre a caso un individuo della popolazione possibile definito dalla variabile X e di estrarre a caso un individuo dalla variabile Y.

Nell’eseguire un numero molto grande di doppie estrazioni, tutti i valori x_i si combineranno con i valori y_i e daranno luogo ad una serie di *t* valori della “somma” (con modalità matematica definita).

$$s_i = x_j + y_k$$

e quindi si avrà la variabile composta

$$S\begin{cases} s_1 \quad s_2 \quad \dots \quad s_t \\ f_1 \quad f_2 \quad \dots \quad f_t \end{cases} \qquad \sum_{i=1}^{n} f_i = 1$$

La deduzione della distribuzione della variabile composta può alle volte essere molto complessa e spesso occorre ricorrere a soluzioni approssimate.

Cerchiamo le relazioni che permettono, noti i momenti di 1° e 2° grado della X e della Y, di ricavare i momenti di 1° e 2° grado della variabile casuale composta S:

$$m_{1(s)} = \sum_{i=1}^{t} s_i f_i = \sum_{j=1}^{n} \sum_{k=1}^{m} \left(x_j \pm y_k\right) f_{jk} = \sum_{j=1}^{n} \sum_{k=1}^{m} \left(x_j \pm y_k\right) p_j g_k =$$

$$= \sum_{j=1}^{n} \left(x_j p_j \sum_{k=1}^{m} g_k \pm p_j \sum_{k=1}^{m} y_k g_k \right) = \sum_{j=1}^{n} x_j p_j \pm \sum_{j=1}^{n} p_j \sum_{k=1}^{m} y_k g_k$$

Modalità
Si = xj + yk

REGOLA N.1: il valore medio di una variabile casuale somma (o differenza) di due variabili casuali è pari alla somma (o differenza) dei valori medi delle variabili casuali componenti:

$$m_{1(s)} = m_{(x)} \pm m_{(y)}$$

$$m_{2(s)} = \sum_{i=1}^{t} s_i^2 f_i = \sum_{j=1}^{n} \sum_{k=1}^{m} \left(x_j \pm y_k\right)^2 f_{jk} = \sum_{j=1}^{n} \sum_{k=1}^{m} \left(x_j \pm y_k\right)^2 p_j g_k =$$

$$= \sum_{j=1}^{n} \sum_{k1}^{m} x_j^2 p_j g_k + \sum_{j=1}^{n} \sum_{k=1}^{m} y_k^2 p_j g_k + 2 \sum_{j=1}^{n} \sum_{k=1}^{m} x_j y_k p_j g_k$$

REGOLA N.2: il valore quadratico medio di una variabile somma (o differenza) di due variabili casuali è la somma di valori quadratici medi delle variabili componenti aumentata (o diminuita) del doppio prodotto delle due medie.

$$m_{2(s)} = m_{2(x)} + m_{2(y)} \pm 2 m_{1(x)} \cdot m_{1(y)}$$

$$\sigma^2_{(s)} = m_{2(v)} = \sigma^2_x + \sigma^2_y$$

REGOLA N.3: la VARIANZA di una variabile casuale somma (o differenza) di due variabili casuali è la somma delle varianze casuali componenti.

In generale, data la funzione:

$$W = aX \pm bY \pm cZ$$

si ha (REGOLA N.4):

$$m_{1(W)} = am_{1(X)} \pm bm_{1(Y)} \pm cm_{1(Z)}$$

$$\sigma^2_{(W)} = a^2\sigma^2_{(X)} + b^2\sigma^2_{(Y)} + c^2\sigma^2_{(Z)}$$

Riepilogo

- Popolazioni di individui definibili con esattezza, essendo gli individui osservabili uno per uno: tipico della variabile statistica: risultato di **esperimenti** (concreta, finita, discreta).

- Popolazioni definite e nelle quali si fanno estrazioni a caso molto numerose: variabile statistica campione.

- Popolazione ben definita, ma non si possono osservare tutti gli individui per motivi economici, di tempo, ecc : si crea una variabile campione.

- Caso delle “popolazioni possibili ”, definibili facendo un numero infinito di prove. Si costruisce una variabile casuale, che possa rappresentare questa popolazione, basata su certe ipotesi: sono **modelli interpretativi** (quindi astratti, illimitati, continui).

- Postulazione di una identità formale tra VARIABILE STATISTICA e VARIABILE CASUALE, che discende dalla loro completa (formale) indistinguibilità.

PRODOTTO di DUE VARIABILI CASUALI

Il valore medio del prodotto di due variabili casuali è il prodotto dei valori medi delle variabili casuali componenti:

$$m_{1(P)} = m_{1(X)} \cdot m_{1(Y)}$$

Il valore quadratico medio del prodotto di due variabili casuali è uguale al prodotto dei valori quadratici medi delle variabili componenti:

$$m_{2(P)} = m_{2(X)} \cdot m_{2(Y)}$$

La varianza del prodotto di due variabili casuali vale:

$$\sigma^2_{(P)} = m_{2(P)} - m^2_{1(P)} = m_{2(X)} \cdot m_{2(Y)} - m^2_{1(X)} \cdot m^2_{1(Y)} = \sigma^2_{(X)} \cdot \sigma^2_{(Y)} + m^2_{1(Y)} \cdot \sigma^2_{(X)} + m^2_{1(X)} \cdot \sigma^2_{(Y)}$$

essendo:

$$m_{2(X)} = \sigma^2_{(X)} \cdot m^2_{1(X)}$$

$$m_{2(Y)} = \sigma^2_{(Y)} \cdot m^2_{1(Y)}$$

QUADRO SINOTTICO
(come variabili limiti-teoriche)

- **PROBABILITA’** = Limite cui tende la frequenza per N→∞
- Analoga definizione dei momenti
- Variabili casuali = combinazione di variabili statistiche
  - Operazioni sulle variabili casuali
    - 1° regola: M_s
    - 2° regola: q_s
    - 3° regola: σ_s^2
    - 4°regola
  - Probabilità
    - Totali: $p_t = p_a + p_b + p_c$
    - Composte: $p_t = p_a \times p_b \times p_c$

DISTRIBUZIONE BINOMIALE o di BERNOULLI

Pensiamo ad una popolazione di cui un attributo è caratterizzato da due soli valori argomentali. Si ha allora una variabile casuale del tipo

$$X = \begin{cases} x_1 \;\; x_2 \\ q \;\;\; p \end{cases} \qquad q + p = 1$$

che si può scrivere

$$X = \begin{cases} 0 & 1 \\ q & p \end{cases}$$

La si combini *n* volte, sommando i valori argomentali:

$$X = \begin{cases} 0 & 1 \\ q & p \end{cases} \qquad X = \begin{cases} 0 & 1 \\ q & p \end{cases} \qquad \Rightarrow \qquad X' = \begin{cases} 0+0 & 0+1=1 & 1+1=2 \\ & 1+0=1 & \\ q^2 & qp+qp & p^2 \end{cases}$$

si ottiene cioè

$$X' = \begin{cases} 0 & 1 & 2 \\ q^2 & 2qp & p^2 \end{cases}$$

Se si combina la nuova variabile X’ con la variabile originaria X si avrà:

$$X''=\left\{\begin{array}{cccc} 0 & 1 & 2 & 3 \\ q^3 & 3pq^2 & 3qp^2 & p^3 \end{array}\right.$$

Se si continua per n combinazioni, la variabile Y così definita prende il nome di DISTRIBUZIONE BINOMIALE:

$$Y=\left\{\begin{array}{cccccc} 0 & 1 & 2 & \dots & n-1 & n \\ q^n & npq^{n-1} & \binom{n}{2}p^2q^{n-2} & \dots & \binom{n}{n-1}p^{n-1}q & p^n \end{array}\right.$$

Note la media e la varianza della variabile X, si ottengono la media e la varianza della Y:

$$X=\begin{cases} m_{1_{(X)}}=0\cdot q+1\cdot p=p \\ m_{2_{(X)}}=0^2\cdot q+1^2\cdot p=p \\ \sigma^2_{(X)}=m_{2_{(X)}}-m^2_{1_{(X)}}=p-p^2=p(1-p)=p\cdot q \end{cases}$$

$$Y=\begin{cases} m_{1_{(Y)}}=n\cdot p \\ \sigma^2_{(Y)}=n\cdot p\cdot q \end{cases}$$

Si è visto che (distribuzione di Bernoulli) la probabilità i-esima è:

$$P(i)=\binom{n}{i}p^i \cdot q^{n-i}$$

n molto grande

Ponendo *v* = *i-np* e supponendo *v* molto piccolo, dopo parecchie manipolazioni, si ricava:

$$P(v)=\frac{1}{\sqrt{2\pi npq}} \cdot e^{-\frac{v^2}{2npq}}$$

e ponendo:

$$h=\frac{1}{\sqrt{2npq}}=\frac{1}{\sqrt{2}\cdot\sigma}$$

si ottiene:

$$P(v)=\frac{h}{\sqrt{\pi}} \cdot e^{-h^2v^2}$$

espressione approssimata della probabilità associata al valore argomentale *v*

Rappresentazione grafica della distribuzione di Bernoulli:

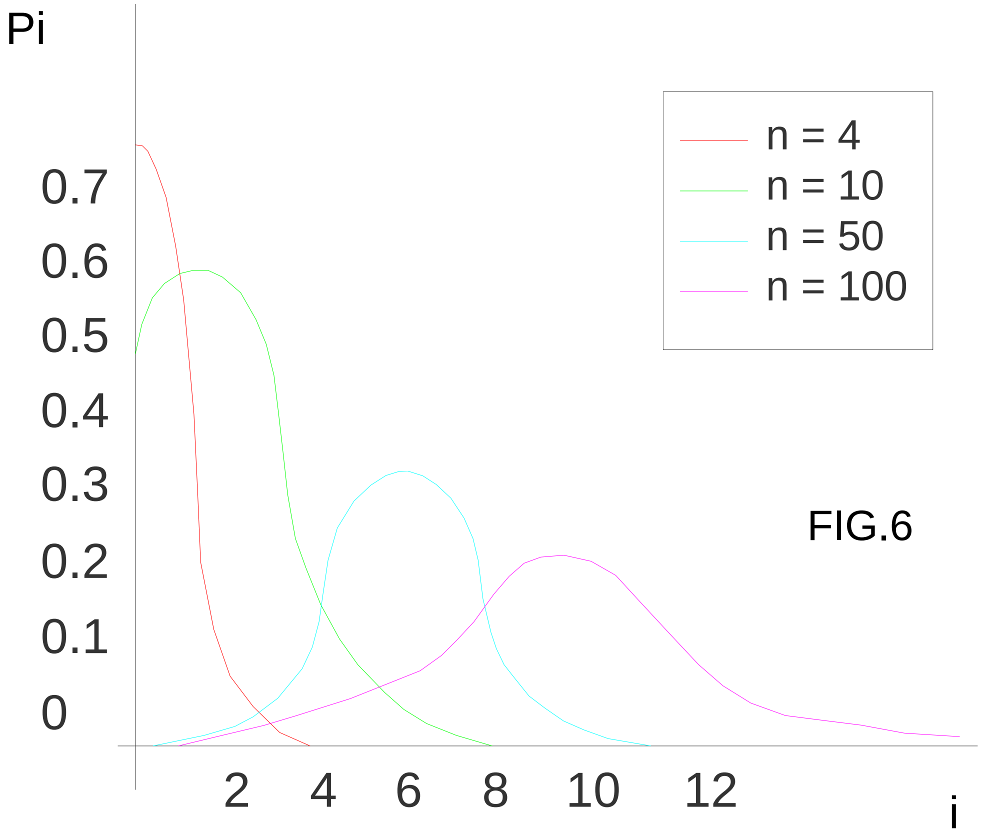


FIG.6

p = 0.1

q = 0.9

Cenni sulla DISTRIBUZIONE di POISSON

E’ detta anche “*degli eventi rari*” e viene costruita sull’ipotesi che la probabilità p dell’evento x_2 non sia costante, ma vari al variare del numero n di combinazioni che si effettuano (p=l/n).

n = numero delle combinazioni molto alto
p = probabilità dell’evento favorevole molto piccola
q = 1-p, prossima ad 1

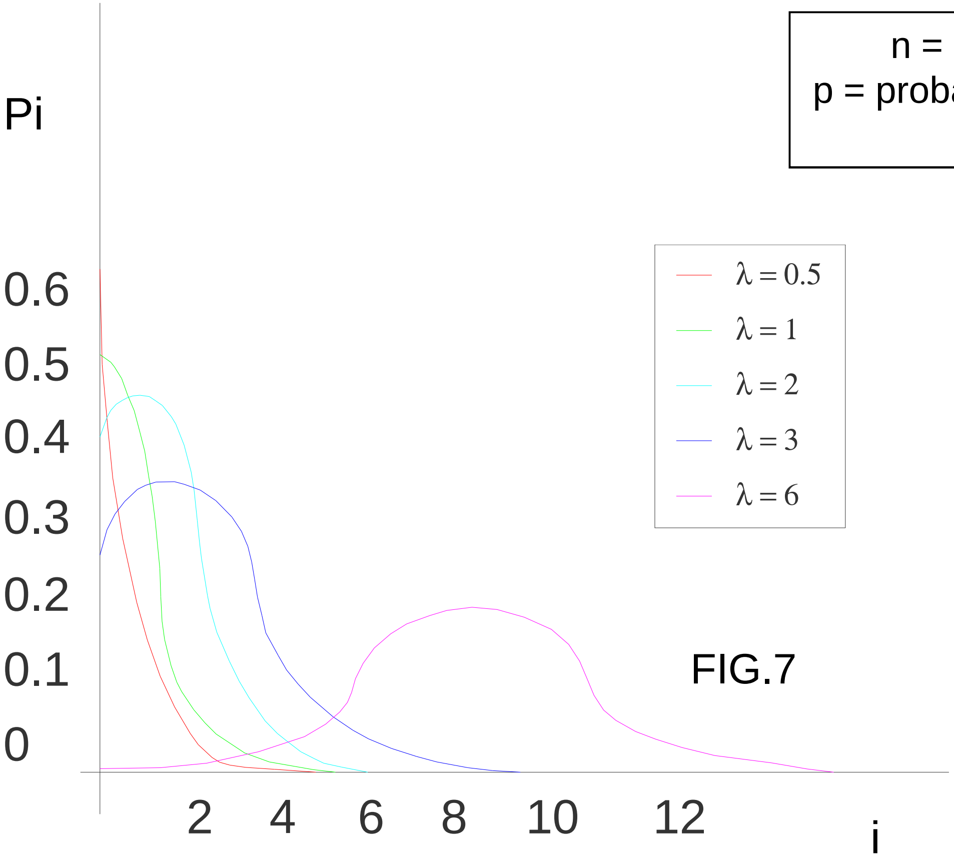


FIG.7

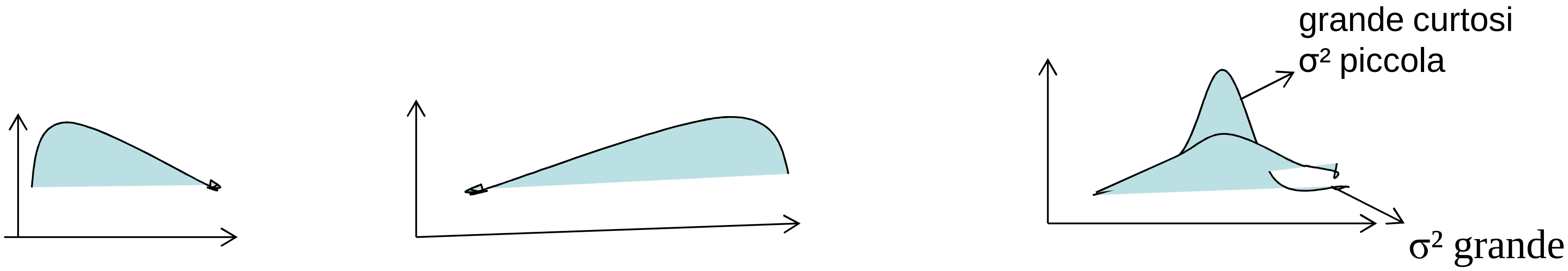
Λ= cost.positiva assegnata
I = valori argom. 0,1,2,..

PROBABILITA’ DELL’EVENTO i:

$$\frac{\lambda^i \cdot e^{-\lambda}}{i!}$$

Media = varianza = λ

In generale però le **distribuzioni** non sono simmetriche rispetto al loro massimo, ma hanno una delle loro “code” più lunga dell’altra e cioè:



Asimmetrica a sin.

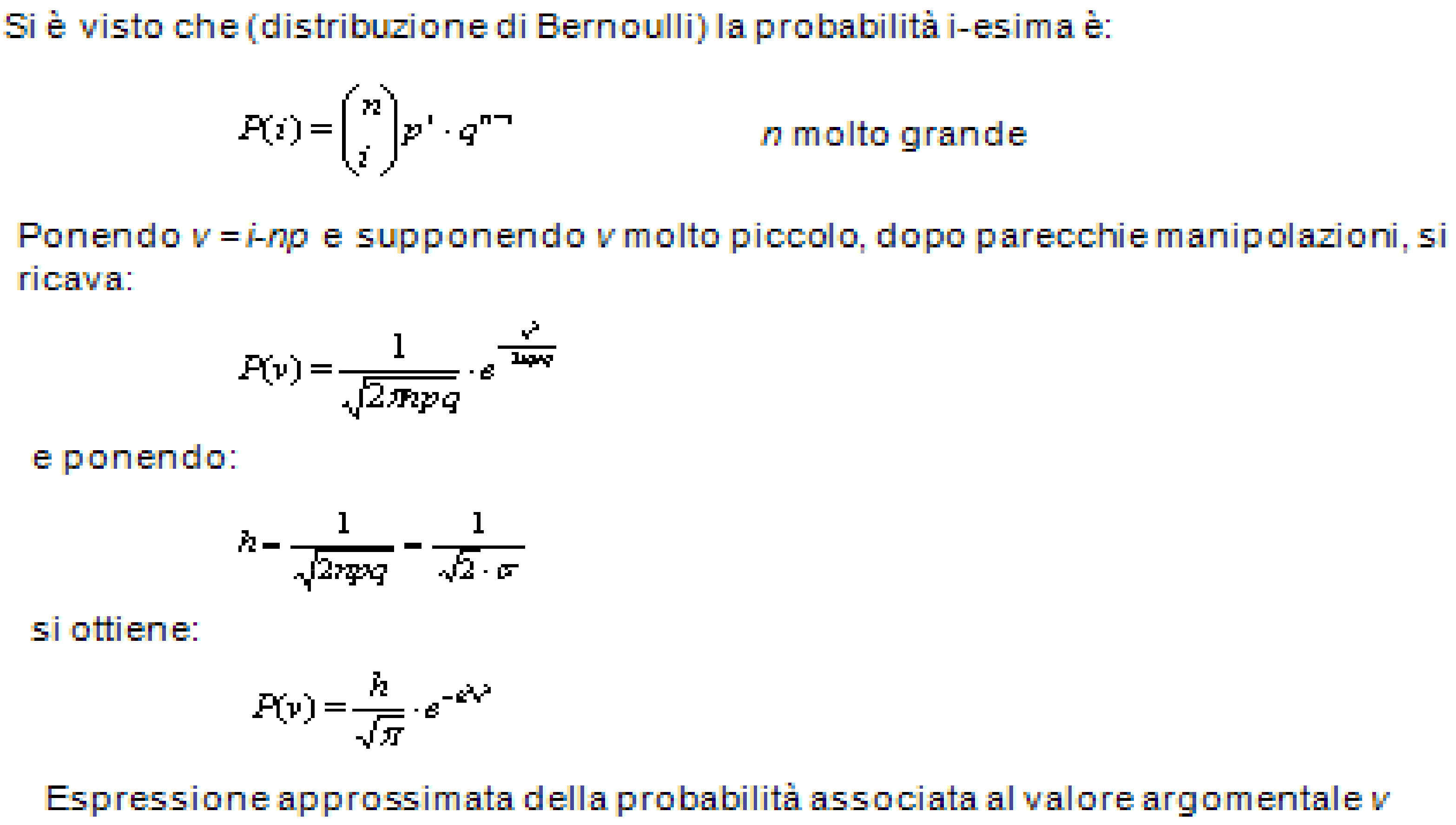
Asimmetrica a dx

FIG.8

Il *coefficiente di simmetria* $\alpha_3 = \mu_3 / \sigma^3$ permette di dire se la distribuzione è simmetrica a destra (positiva) o a sinistra (negativa). Se $\alpha_3=0$ la distribuzione è simmetrica. Il ***coefficiente di curtosi*** $\alpha_4 = \mu_4 / \sigma_4$ misura il grado di ripidezza della distribuzione.

Nella curva “normale” il coefficiente è pari a 3 ed il coefficiente di simmetria è “zero”.

Verso la distribuzionale normale



Si è visto che (distribuzione di Bernoulli) la probabilità i-esima è:

$$P(i) = \binom{n}{i} p^{i} \cdot q^{n-i}$$

n molto grande

Ponendo *v* = *i-np* e supponendo *v* molto piccolo, dopo parecchie manipolazioni, si ricava:

$$P(v) = \frac{1}{\sqrt{2\pi npq}} \cdot e^{\frac{v^2}{2npq}}$$

e ponendo:

$$h = \frac{1}{\sqrt{2npq}} = \frac{1}{\sqrt{2} \cdot \sigma}$$

si ottiene:

$$P(v) = \frac{h}{\sqrt{\pi}} \cdot e^{-h^2 v^2}$$

Espressione approssimata della probabilità associata al valore argomentale *v*

DISTRIBUZIONE NORMALE o di GAUSS

Considerata continua la variabile casuale così determinata, definiamo la probabilità infinitesima dp che lo scarto sia compreso tra v e (v+dv):

$$dp = \frac{h}{\sqrt{\pi}} \cdot e^{-h^2 v^2} dv \qquad e = \lim_{m \to \infty} \left(1 + \frac{1}{m}\right)^m = 2,718281828459045...$$

Una distribuzione di questo tipo si chiama NORMALE.

Sappiamo in generale che:

$$v_i = x_i - M$$

$$h^2 = \frac{1}{2\sigma^2}$$

E quindi:

$$\boxed{y(v) = \frac{1}{\sqrt{2\pi} \cdot \sigma} e^{\frac{-(x_i - M)^2}{2\sigma^2}}}$$

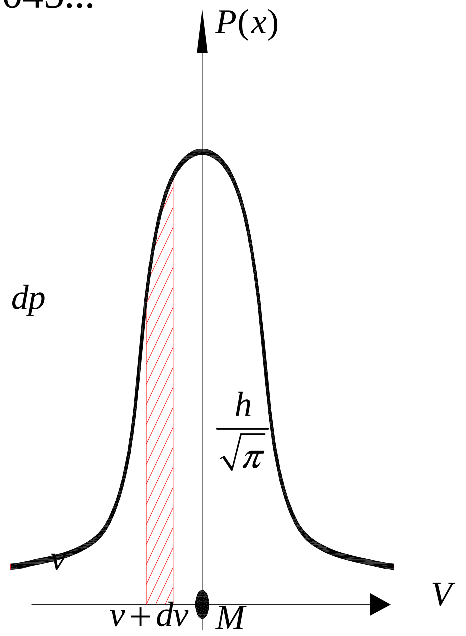


FIG.9

(definizione di funzione NORMALE definita soltanto dai parametri M e σ^2)

Alcune proprietà:

$$\int_{-\infty}^{+\infty} dp = \int_{-\infty}^{+\infty} \frac{h}{\sqrt{\pi}} \cdot e^{-h^2 v^2} dv = 1$$

$$\int_{-\infty}^{+\infty} v \cdot dp = m_{(v)} = \int_{-\infty}^{+\infty} v \frac{h}{\sqrt{\pi}} \cdot e^{-h^2 v^2} dv = 0$$

Cioè la media della variabile scarto è nulla.

La curva delle fluttuazioni accidentali deve essere simmetrica, rispetto alla retta x=a, essendo **a** l’ascissa del massimo. Si ipotizza inoltre che la curva abbia come campo di definizione tutto l’asse reale e che questo sia un asintoto della curva.

La curva normale y=f(x) deve avere equazione:

$$\frac{dy}{dx} = \frac{-y(x-a)}{b^2}$$

per y non negativa

Calcoliamo l’integrale dell’equazione differenziale; per prima cosa si separano le variabili: $\frac{dy}{y} = \frac{x-a}{b^2}dx$

E quindi $\int \frac{dy}{y} = -\int \frac{x-a}{b^2}dx$ o anche $\log y = -\frac{(x-a)^2}{2b^2} + c$

Posto $c = \log y_0$ si ha $\log y = -\frac{(x-a)^2}{2b^2} + c$

Ossia $\log y - \log y_0 = -\frac{(x-a)^2}{2b^2}$ e cioè $\frac{y}{y_0} = e^{-\frac{(x-a)^2}{2b^2}}$

O anche $y = y_0 e^{-\frac{(x-a)^2}{2b^2}}$

Si determina $y_0 = \frac{1}{\sqrt{2\pi}\cdot b}$ con $a = m \quad b^2 = \sigma^2$

E quindi

$$\boxed{y = \frac{1}{\sqrt{2\pi}\cdot\sigma} e^{\frac{-(x-m)^2}{2\sigma^2}}}$$

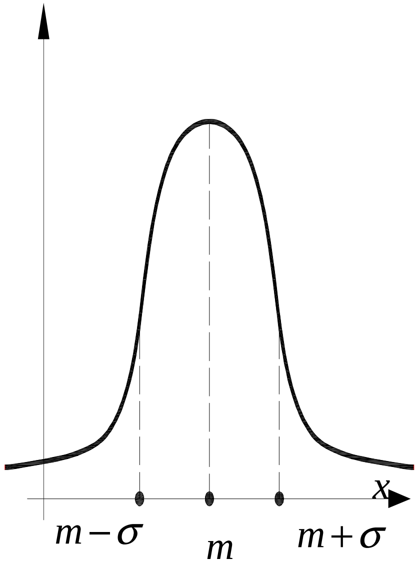


FIG.10

Si sono già viste alcune proprietà:

1) $\int_{-\infty}^{+\infty} dp = \int_{-\infty}^{+\infty} \frac{h}{\sqrt{\pi}} \cdot e^{-h^2v^2} dv = 1$ (cioè l' area compresa sotto la funzione vale 1)

2) m= $\int_{-\infty}^{+\infty} v \cdot dp = m_{(v)} = \int_{-\infty}^{+\infty} v \frac{h}{\sqrt{\pi}} \cdot e^{-h^2v^2} dv = 0$

(Cioè la media della variabile scarto è nulla).

3) **varianza**= $\int_{-\infty}^{+\infty} v^2 \frac{h}{\sqrt{\pi}} e^{-h^2v^2} dv = \frac{1}{2h^2} = \sigma^2$

h si chiama *misura di precisione*.

Si introduce la **VARIABILE SCARTO STANDARDIZZATA**:

$$z = \frac{x_i - M}{\sigma}$$

(il rapporto è un numero puro: ciò rende possibile fare confronti tra distribuzioni secondo caratteri o parametri diversi). L'espressione della nuova funzione diventa:

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

Questa nuova variabile ha **media nulla** e **varianza** uguale a 1.

La distribuzione normale è interessante perché si trova tabulata e fornisce la probabilità che un valore argomentale z generico sia compreso tra due limiti **a** e **b**. Dal valore di z si passa quindi al valore di x e di v. In particolare, per il valore dello scarto v: per

λ = 1 si trova il valore di probabilità 0,683

λ = 2 si trova il valore di probabilità 0,954

λ = 3 si trova il valore di probabilità 0,997

Ciò vuol dire che nelle distribuzioni normali, la popolazione di misure possibili ha:

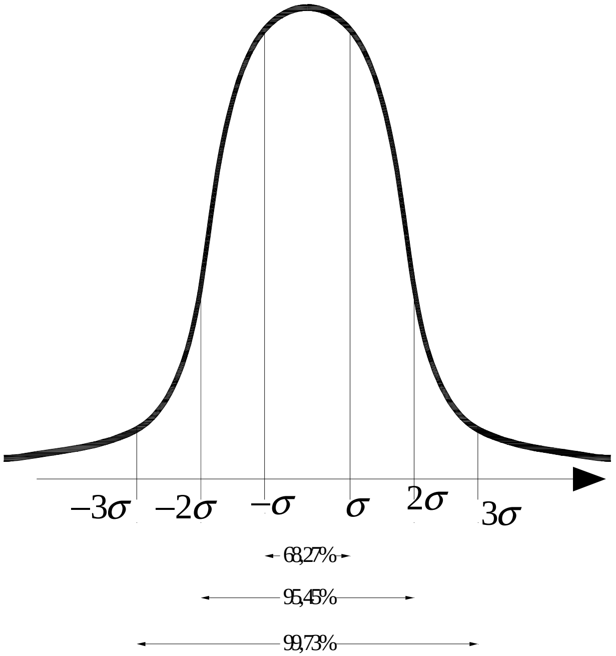
il 68,3% degli scarti compresi tra ±σ

il 95,4% degli scarti compresi tra ±2σ

il 99,7% degli scarti compresi tra ±3σ

Nel caso delle popolazioni di misure possibili, a distribuzione normale, si pone lo scarto pari a **±3σ** come limite praticamente invalicabile (concetto di **TOLLERANZA**)

FIG.11



Alcune distribuzioni teoriche continue

Si riportano alcune distribuzioni utili in casi particolari :

Distribuzione Chi Quadro χ^2

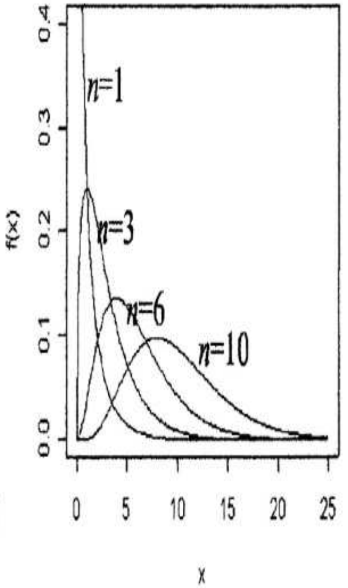
È la distribuzione della somma di n variabili casuali VC normali indipendenti

$\chi^2 = N\, s^2/ \sigma^2 = [\, (x_1-\overline{x})^2 + (x_2-\overline{x})^2 + \dots (x_n-\overline{x})^2\,] / \sigma^2$

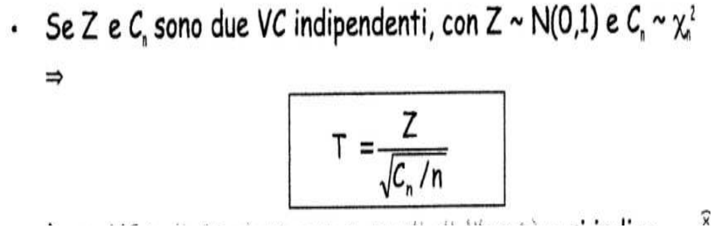
assume valori in [0 , + ∞]

Il valore medio della distribuzione vale il numero di gradi di libertà **n** e la varianza vale **2n.**

Per n → ∞ la distribuzione tende alla normale.



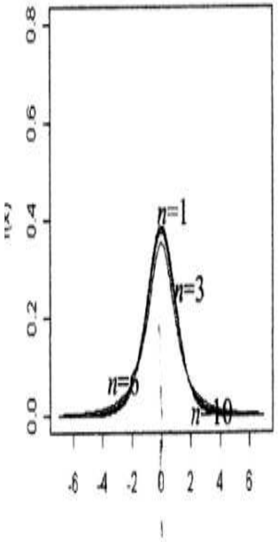
Distribuzione t di Student



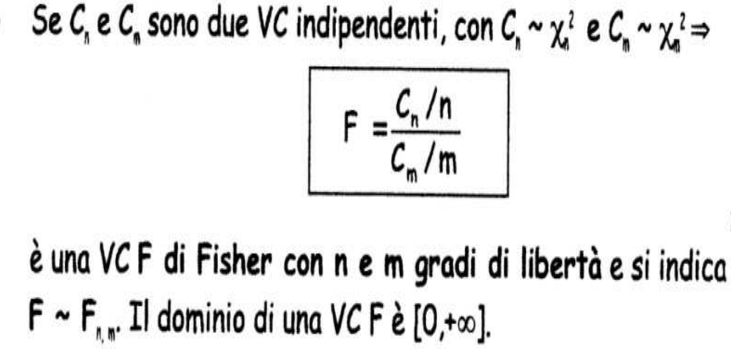
- Se Z e C_n sono due VC indipendenti, con $Z \sim N(0,1)$ e $C_n \sim \chi^2_n$ $\Rightarrow$

$$\boxed{T = \frac{Z}{\sqrt{C_n/n}}}$$

Il valore medio vale “zero”
La varianza vale n/(n-2) .Tende a una normale per n → ∞



Distribuzione F di Fisher

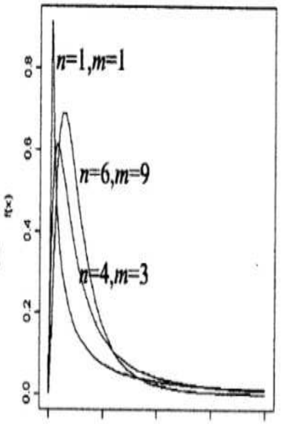


Se C_n e C_m sono due VC indipendenti, con $C_n \sim \chi^2_n$ e $C_m \sim \chi^2_m \Rightarrow$

$$\boxed{F = \frac{C_n/n}{C_m/m}}$$

è una VC F di Fisher con n e m gradi di libertà e si indica $F \sim F_{n,m}$. Il dominio di una VC F è [0,+∞].

Per n → ∞ la distribuzione tende alla normale.



VARIABILE STATISTICA DOPPIA
(o a due dimensioni)

x y	X_1	X_2	X_3	…	…	X_i	…	X_m	
Y_1	F_{11}	F_{21}	F_{31}	…	…	F_{i1}	…	F_{m1}	Q_1
Y_2	F_{12}	F_{22}	F_{32}	…	…	F_{i2}	…	F_{m2}	Q_2
Y_3	…	…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…	…	…
Y_j	F_{1j}	F_{2j}	F_{3j}	…	…	F_{jj}	…	F_{mj}	Q_j
…	…	…	…	…	…	…	…	…	…
Y_n	F_{1n}	F_{2n}	F_{3n}	…	…	F_{in}	…	F_{mn}	Q_n
	P_1	P_2	P_3	…	…	P_i	…	P_m	N

$i = 1 \div m \quad j = 1 \div n \quad F_{ij} = \text{Frequenze assolute}$ Essendo:

$$P_i = \sum_{1}^{n} F_{ij} \quad Q_j = \sum_{1}^{m} F_{ij} \quad \sum_{i=1}^{m} P_i = \sum_{j=1}^{n} Q_j = N$$

Per avere invece le frequenze relative occorre fare: $\dfrac{F_{ij}}{P_i} \quad \dfrac{F_{ij}}{Q_j}$

- Si definiscono le VARIAZIONI STATICHE MARGINALI

$$X\begin{cases} x_1 x_2 \ldots\ldots\ldots x_m \\ p_1 p_2 \ldots\ldots\ldots p_m \end{cases} \qquad Y\begin{cases} y_1 y_2 \ldots\ldots\ldots y_m \\ Q_1 Q_2 \ldots\ldots\ldots Q_m \end{cases}$$

ed i rispettivi momenti (media e varianza)

$$M_{(x)} = \sum_{i=1}^{m} x_i \frac{P_i}{N} \qquad M_{(Y)} = \sum_{j=1}^{n} y_j \frac{Q_j}{N}$$

$$\sigma_x^2 = \sum_{i=1}^{m} \left(x_i - M_{(x)}\right)^2 \cdot \frac{P_i}{N} \qquad \sigma_Y^2 = \sum_{j=1}^{n} \left(y_j - M_{(Y)}\right)^2 \cdot \frac{Q_j}{N}$$

Si possono ottenere altre VARIABILI STATISTICHE SEMPLICI, del tipo:

$$Y_i\begin{cases} y_1 y_2 \ldots y_j \ldots y_n \\ f_{i1} f_{i2} \ldots f_{ij} \ldots f_{in} \end{cases}$$

Per $i = 1 \div m$

La scrittura precedente equivale a scrivere m variabili del tipo :

$$Y_1 \begin{cases} y_1 \, y_2 \cdots y_i \\ f_{11} \; f_{12} \cdots f_{1n} \end{cases}$$ di P$_1$ individui

$$Y_2 \begin{cases} y_1 \, y_2 \cdots y_i \\ f_{21} \; f_{22} \cdots f_{2n} \end{cases}$$ di P$_2$ individui …

fino a Y m $$Y_1\text{m} \begin{cases} y_1 \, y_2 \cdots y_i \\ f_{m1} \; f_{m2} \cdots f_{mn} \end{cases}$$ di P$_m$ individui

In base alle relazioni scritte prima, per ogni colonna sarà:

- media parziale $$M_{[y_{(i)}]} = M_{y_i} = \sum_{j=1}^{n} y_i \cdot \frac{f_{ij}}{P_i}$$

- varianza parziale $$\sigma^2[Y_{(i)}] = \sigma^2{}_{Y_i} = \sum \left(y_i - M_{y_i}\right)^2 \cdot \frac{f_{ij}}{P_i}$$

Tabella a 4 colonne ed m righe

X	M_{yi}	σ^2_{yi}	P_i	
X_1	M_{y1}	σ^2_{y1}	P_1	$M_{y1} \pm 2\sigma_{y1}$
X_2	M_{y2}	σ^2_{y2}	P_2	$M_{y2} \pm 2\sigma_{y2}$
…	…	…	…	…
X_i	M_{yi}	σ^2_{yi}	P_i	$M_{yi} \pm 2\sigma_{yi}$
…	…	…	…	…
X_m	M_{ym}	σ^2_{ym}	P_m	$M_{ym} \pm 2\sigma_{ym}$

Curva di regressione

Se si riportano i dati relativi alla media su un grafico, si può costruire la curva di regressione:

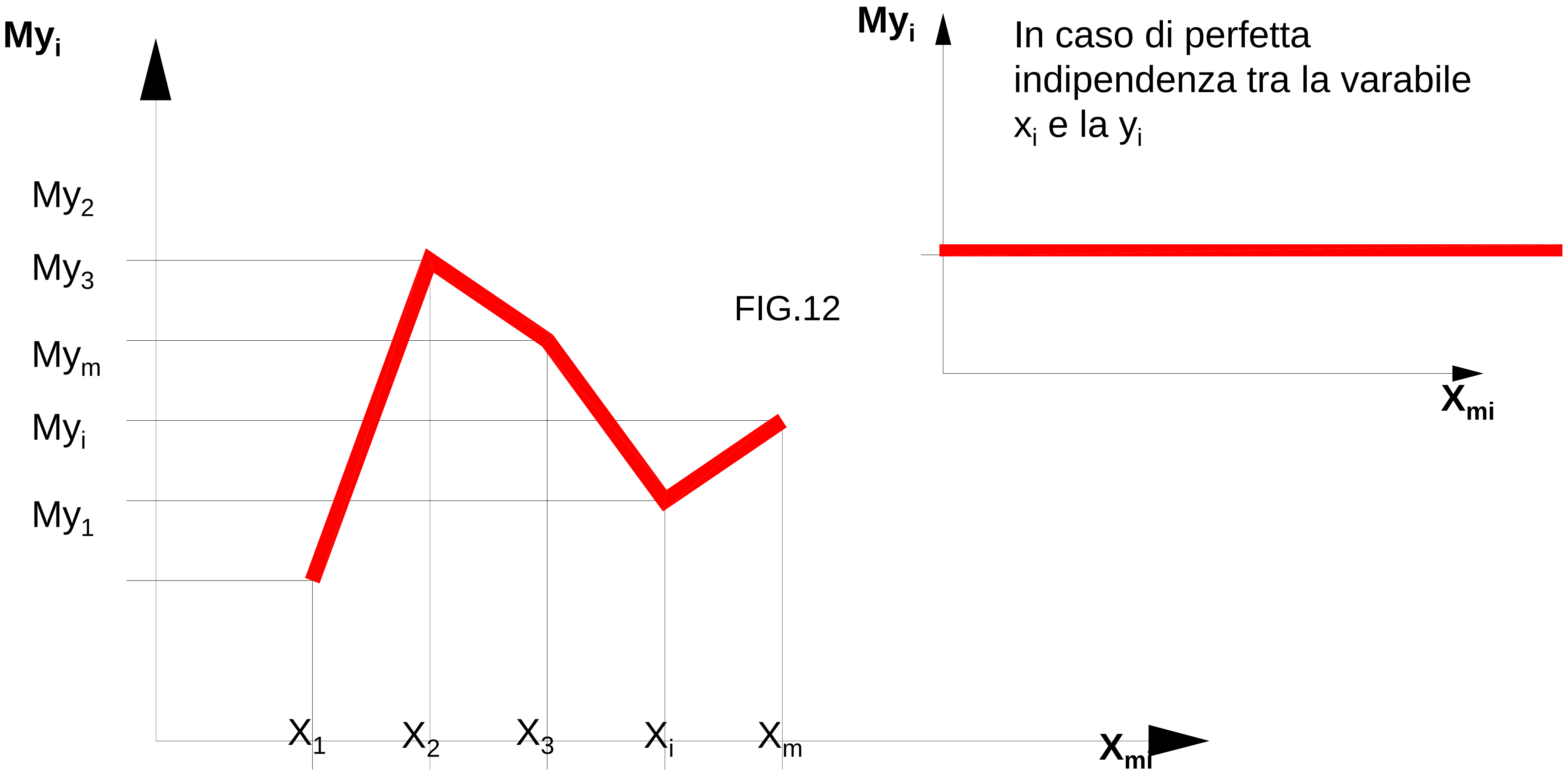
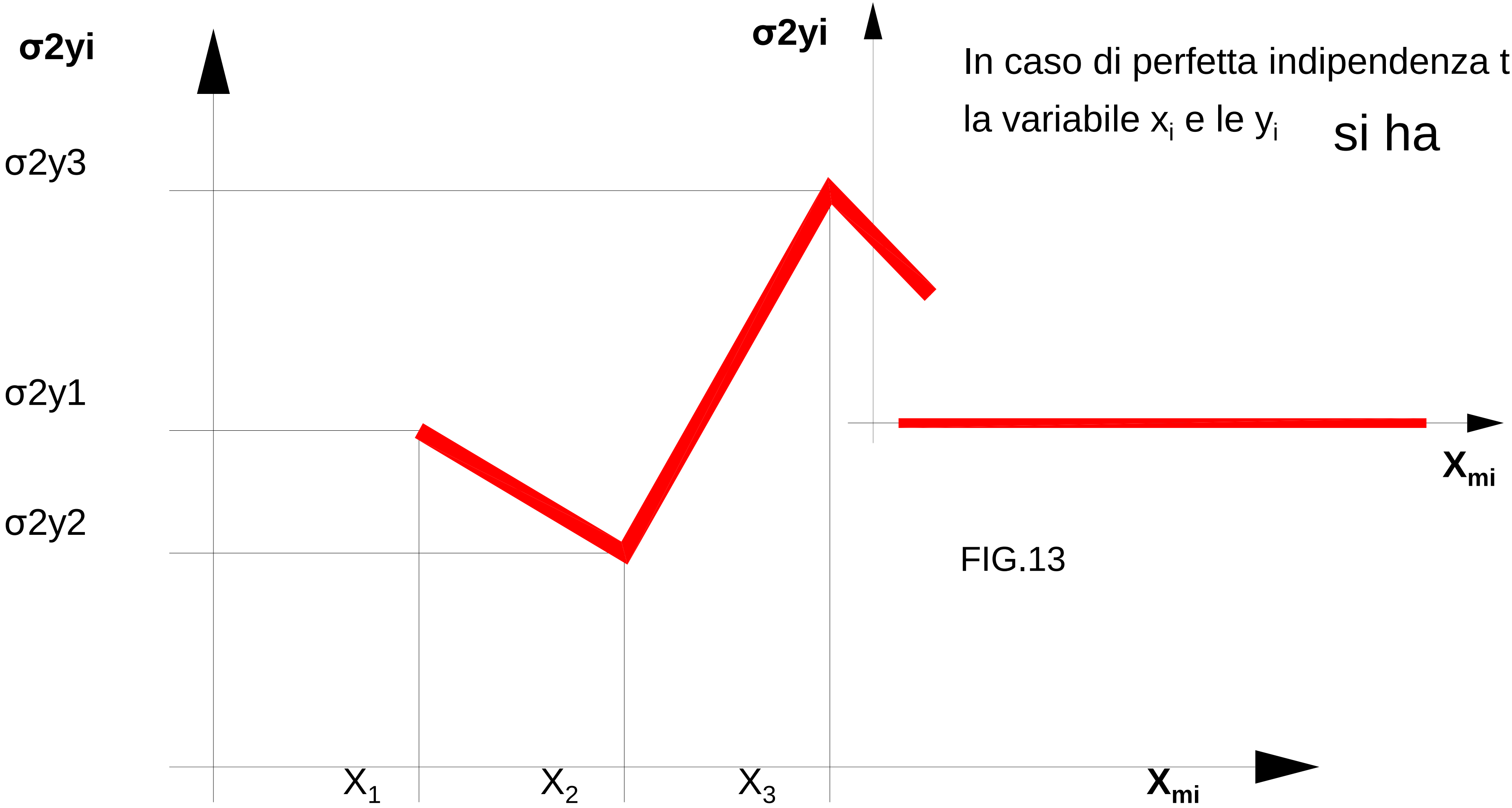


FIG.12

In caso di perfetta indipendenza tra la varabile x_i e la y_i

Curva di variabilità

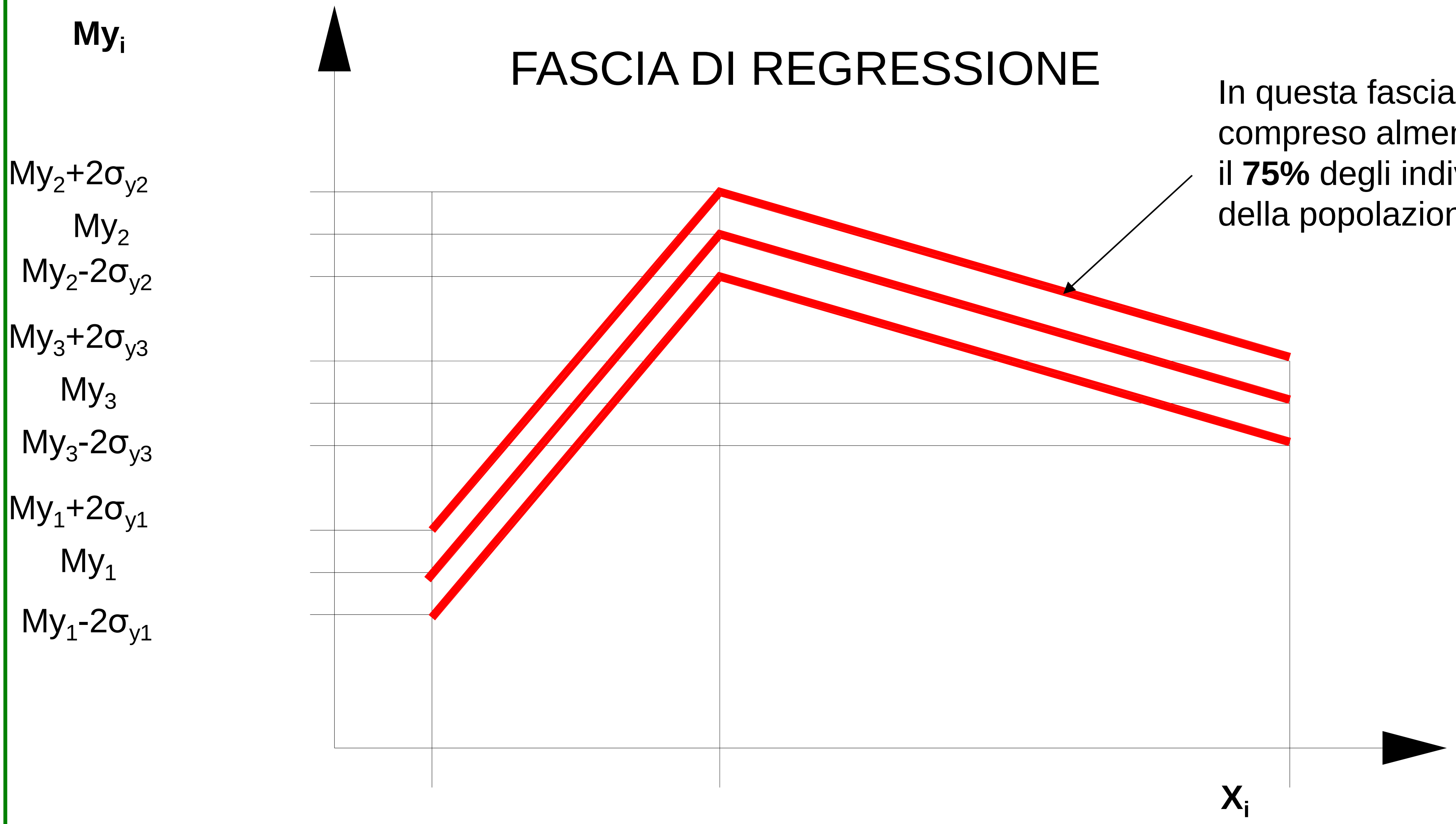
Se si riportano i dati relativi alla varianza su un grafico, si può costruire la curva di variabilità:



In caso di perfetta indipendenza tra la variabile x_i e le y_i si ha

FIG.13

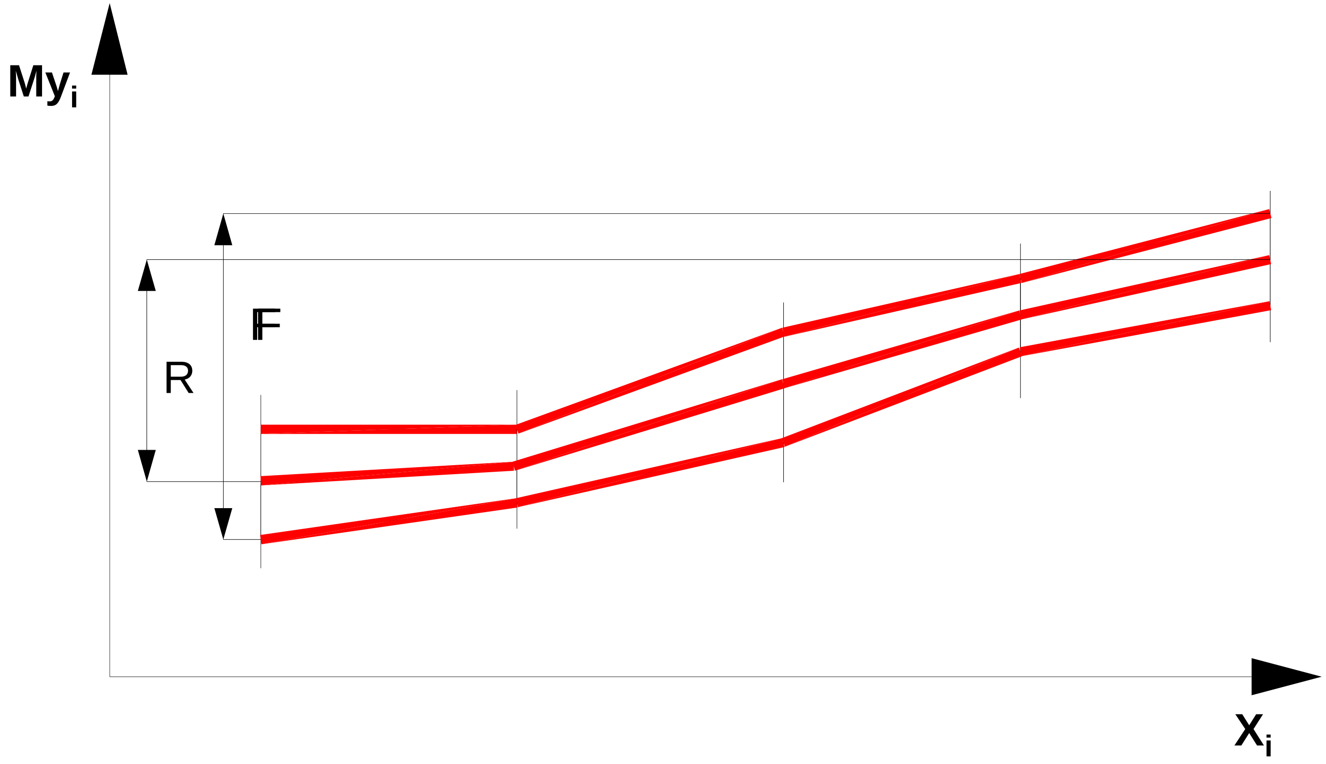
FASCIA DI REGRESSIONE



In questa fascia è
compreso almeno
il **75%** degli individui
della popolazione

FIG.14

Si prenda un caso qualsiasi di **fascia di regressione**. Si possono presentare 3 casi:



a) R / F = 1

Perfetta dipendenza

b) Rapporto intermedio tra 0 e 1

c) R / F = 0

perfetta indipendenza

FIG.15

VARIABILE STATISTICA DOPPIA
o a due dimensioni

Volendo tradurre in termini analitici le considerazioni svolte in precedenza, si consideri una nuova variabile statistica doppia così costruita:

$y*$ \ X	X1	X2	…	Xi	…	Xm	
My1	P1	0	…	0	…	0	P1
My2	0	P2	…	0	…	0	P2
...	**...**	**...**	…	**...**	…	**...**	**...**
Myi	0	0	…	Pi	…	0	Pi
...	**...**	**...**	…	**...**	…	**...**	**...**
Mym	0	0	…	0	…	Pm	Pm
	P1	P2	…	Pi	…	Pm	**N**

Si definisce **varianza della tabella derivata della perfetta dipendenza:**

$$\sigma^2_{y^*} = \sum_{i=1}^{m} \left(M_{y_i} - M_y\right)\cdot \frac{P_i}{N} = \sigma^2_{y_{PD}}$$

Il rapporto $\dfrac{\sigma^2_{y_{PD}}}{\sigma^2_{y_m}} = \eta^2_x$ si chiama INDICE di **PEARSON**, e sostituisce l'analogo rapporto R/F

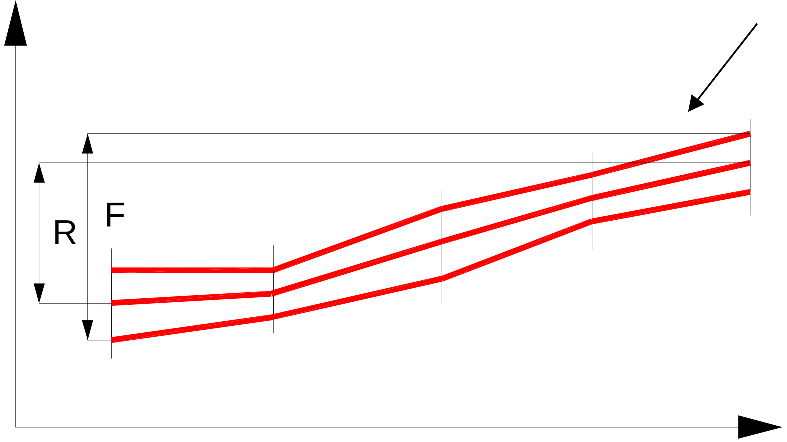


FIG.16

Ricordando il significato di varianza si può ancora scrivere:

$$\eta^2_x = \frac{\sigma^2_{y_{PD}}}{\sigma^2_{y_m}} = \frac{\sum_{i=1}^{m} \left(M_{y_i} - M_y\right)^2 \cdot P_i}{\sum_{j=1}^{n} \left(y_j - M_y\right)^2 \cdot Q_j}$$

$\eta^2_x = 0$ Si ha perfetta **INDIPENDENZA**

$\eta^2_x = 1$ Si ha perfetta **DIPENDENZA**

L'indice di PEARSON non dà però informazioni circa la FORMA della CORRELAZIONE tra X e Y.

Lo studio della forma della correlazione presenta diverse difficoltà. Se si fa l'ipotesi che questa forma sia lineare, si può prendere in esame il *coefficiente di correlazione lineare* r_{xy}**,** così definito:

$$r_{xy}=\frac{\sum_{i=1}^{m}\sum_{j=1}^{n}v_{x_i}v_{x_j}\cdot\frac{f_{ij}}{N}}{\sigma_x\,\sigma_y}=\frac{\sum_{i=1}^{m}\sum_{j=1}^{n}\left(x_i-M_x\right)\left(y_j-M_y\right)\cdot\frac{f_{ij}}{N}}{\sigma_x\sigma_y}=$$

$$=\frac{\sum_{i=1}^{m}\sum_{j=1}^{n}x_i y_i\frac{f_{ij}}{N}-M_xM_y}{\sigma_x\sigma_y}=\frac{\sum_{i=1}^{m}\sum_{j=1}^{n}x_i M_x\frac{P_i}{N}-M_xM_y}{\sigma_x\sigma_y}=\frac{\sigma_{xy}}{\sigma_x\sigma_y}$$

$r_{xy}=\pm1$ Caso di **PERFETTA** DIPENDENZA **LINEARE**

$r_{xy}\;=0$ Caso di **PERFETTA** INDIPENDENZA (la correlazione è nulla o non facilmente interpretabile. In questo caso si studia il coefficiente η_x^2)

TEOREMA di BONFERRONI : $0\le r_{xy}^2\le1$

È stato introdotto, implicitamente, il concetto di **MOMENTO MISTO:**

$$m_{xy} = \sum_{i=1}^{m} \sum_{j=1}^{n} x_i y_j \, \frac{f_{ij}}{N}$$

Il momento misto calcolato rispetto alle **medie** prende il nome di **COVARIANZA:**

$$\sigma_{xy} = \sum_{i=1}^{m} \sum_{j=1}^{n} \left(x_i - M_y \right) \cdot \left(y_i - M_y \right) \cdot \frac{f_{ij}}{N} = m_{xy} - M_x M_y$$

Quindi la covarianza è uguale al momento misto diminuito del prodotto tra le medie delle variabili statistiche marginali (abbr: v.s.m.)

L'espressione della funzione densità di probabilità, nel caso in cui le due variabili X e Y siano **indipendenti** e distribuite con la **legge normale** è:

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

Se invece le due variabili X e Y sono correlate, l'espressione diventa:

$$f\left(x,y\right)=\frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}e^{\frac{-1}{2\cdot\left(1-\rho^2\right)}\left[\left(\frac{x-M_x}{\sigma_x}\right)^2-2\rho\cdot\frac{x-M_x}{\sigma_x}\cdot\frac{y-M_y}{\sigma_y}+\left(\frac{y-M_y}{\sigma_y}\right)^2\right]}$$

ρ è il COEFFICIENTE di CORRELAZIONE e può assumere valori compresi tra -1 e 1.

Funzione DENSITA'di PROBABILITA'

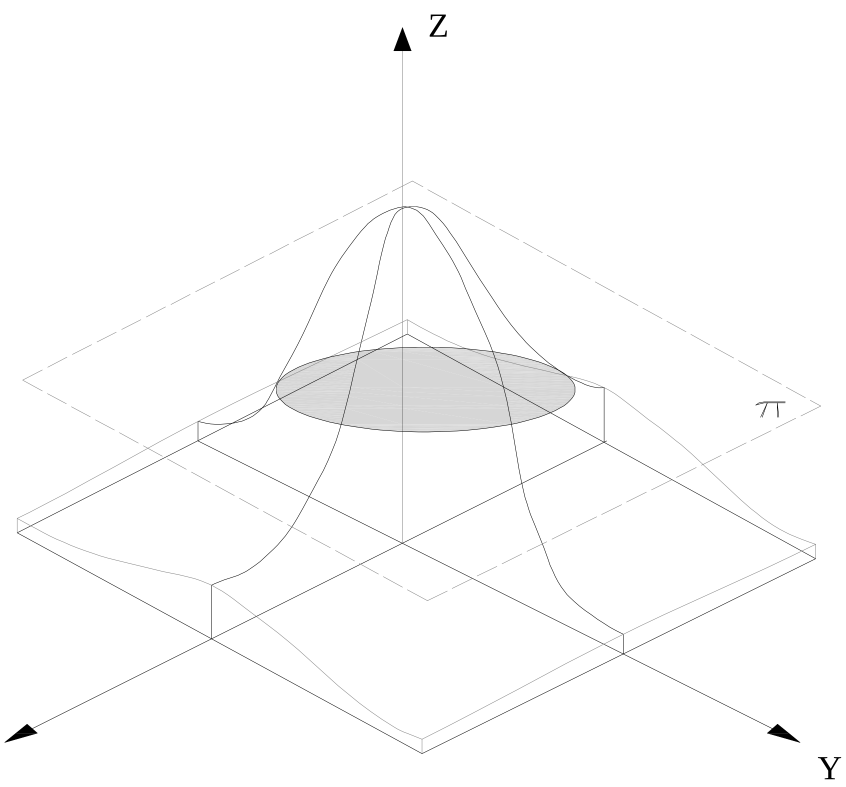


FIG. 17

Nel piano z=cost, la funzione **densità di probabilità** assume la seguente forma

$$\frac{1}{2\left(1-\rho^2\right)}\left[\left(\frac{x-M_x}{\sigma_x}\right)^2-2\rho\cdot\frac{x-M_x}{\sigma_x}\cdot\frac{y-M_y}{\sigma_y}+\left(\frac{y-M_y}{\sigma_y}\right)^2\right]=K$$

Questa curva è sempre un **ellisse**.
Ponendo K=1 si ottiene **l'ellisse STANDARD**.
Ponendo K=4 si ottiene **l'ellisse di UGUALE CONCENTRAZIONE**

ESEMPIO di VARIABILE STATISTICA DOPPIA

	X1=1.70	**X2=1.75**	**X3=1.80**	**X4=1.85**	**M(x)=1.78**
Y1=65	2	0	1	0	**Q1=3**
Y2=70	1	1	1	0	**Q2=3**
Y3=75	2	2	2	1	**Q3=7**
Y4=80	0	3	3	2	**Q4=8**
Y5=85	0	0	1	3	**Q5=4**
M(y)=76.4	**P1=5**	**P2=6**	**P3=8**	**P4=6**	**N=25**

VARIABILI
STATISTICHE
MARGINALI

$$X\left\{\begin{array}{cccc}1.70 & 1.75 & 1.80 & 1.85\\ 5 & 6 & 8 & 6\\ 0.2 & 0.24 & 0.32 & 0.24\end{array}\right.$$

$$Y\left\{\begin{array}{ccccc}65 & 70 & 75 & 80 & 85\\ 3 & 3 & 7 & 8 & 4\\ 0.12 & 0.12 & 0.28 & 0.32 & 0.16\end{array}\right.$$

Sviluppando i calcoli per ricavare media e varianza rispetto a X si ottiene:

$$M_{(x)} = 1.70 \cdot 0.2 + 1.75 \cdot 0.24 + 1.80 \cdot 0.32 + 1.85 \cdot 0.24 = 1.78$$

$$\sigma^2_{(x)} = (1.70 - 1.78)^2 \cdot 0.2 + (1.75 - 1.78)^2 \cdot 0.24 + (1.80 - 1.78)^2 \cdot 0.32 + (1.85 - 1.78)^2 \cdot 0.24$$

$$\sigma^2_{(x)} = 0.0236$$

$$\sigma^2_{(x)} = \pm 0.15$$

E rispetto a Y sarà:

$$M_{(y)} = 76.4$$

$$\sigma^2_{(x)} = 37.05$$

$$\sigma^2_{(x)} = \pm 6.08$$

Si hanno le seguenti variabili statistiche semplici:

$$Y_i = \begin{cases} 65 & 70 & 75 & 80 & 85 \\ f_{i1} & f_{i2} & f_{i3} & f_{i4} & f_{i5} \end{cases} \qquad i = 1 + 4$$

$$\begin{cases} M_{(y1)} = 65 \cdot 0,4 + 70 \cdot 0,2 + 75 \cdot 0,4 = 70 \\ \sigma^2_{(y1)} = (65 - 70)^2 \cdot 0,4 + (75 - 70)^2 \cdot 0,4 = 20 \end{cases}$$

$$\begin{cases} M_{(y2)} = 70 \cdot 0,16 + 75 \cdot 0,333 + 80 \cdot 0,5 = 76,2 \\ \sigma^2_{(y2)} = (70 - 76,2)^2 \cdot 0,16 + (75 - 76,2)^2 \cdot 0,33 + (80 - 76,2)^2 \cdot 0,5 = 13,86 \end{cases}$$

$$\begin{cases} M_{(y3)} = 65 \cdot 0,125 + 70 \cdot 0,125 + 75 \cdot 0,25 + 80 \cdot 0,375 + 85 \cdot 0,125 = 76,25 \\ \sigma^2_{(y3)} = (65 - 76,25)^2 \cdot 0,125 + (70 - 76,25)^2 \cdot 0,125 + (75 - 76,25)^2 \cdot 0,25 + (80 - 76,25)^2 \cdot 0,375 + (85 - 76,25)^2 \cdot 0.125 = 35,93 \end{cases}$$

$$\begin{cases} M_{(y4)} = 75 \cdot 0,16 + 80 \cdot 0,33 + 85 \cdot ,5 = 80,9 \\ \sigma^2_{(y4)} = (75 - 80,9)^2 \cdot 0,16 + (80 - 80,9)^2 \cdot 0,33 + (85 - 80,9)^2 \cdot 0,5 = 11,36 \end{cases}$$

Si possono raccogliere i risultati in una tabella a 4 colonne e m righe:

X	$M_{(yi)}$	$s^2_{(yi)}$	P_i	$s_{(yi)}$	$2s_{(yi)}$	$3s_{(yi)}$
1,7	70	20	5	4,47	8,9	13,4
1,75	76,2	13,86	6	3,72	7,4	11,2
1,8	76,25	35,93	8	5,99	11,9	17,9
1,85	80,9	11,36	6	3,37	6,7	10,1

Vediamo ora la curva di regressione, la curva di variabilità e la fascia di regressione

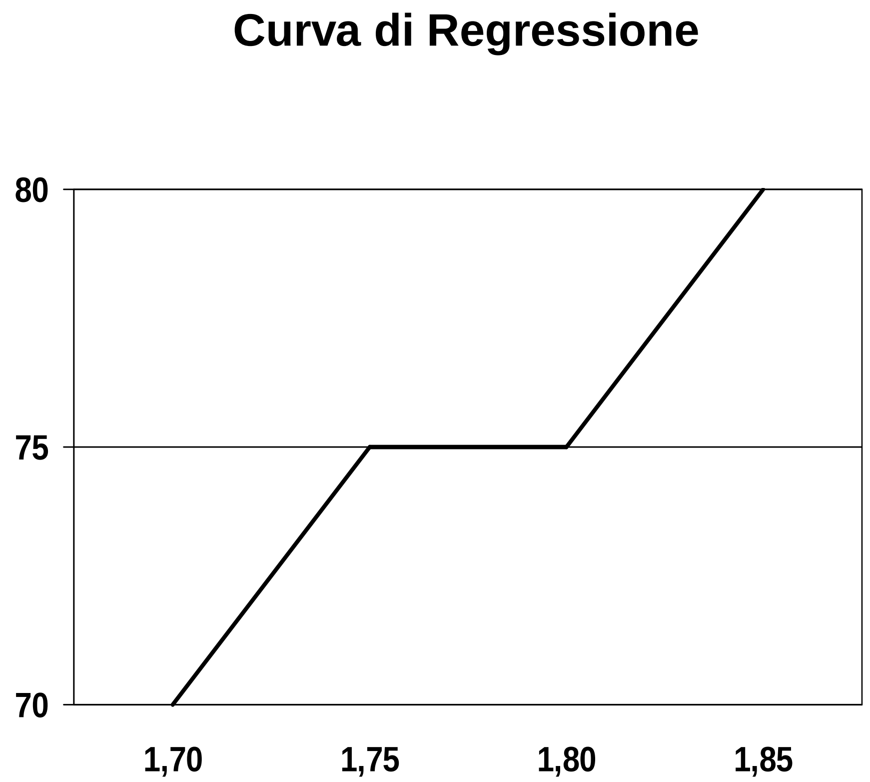

Curva di Regressione

FIG.18

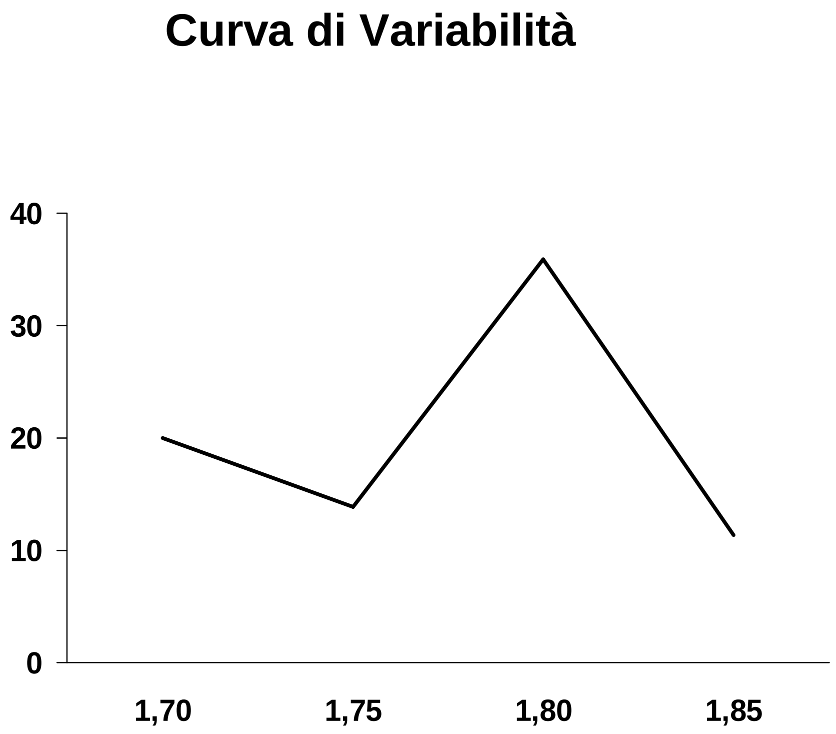

Curva di Variabilità

FIG.19

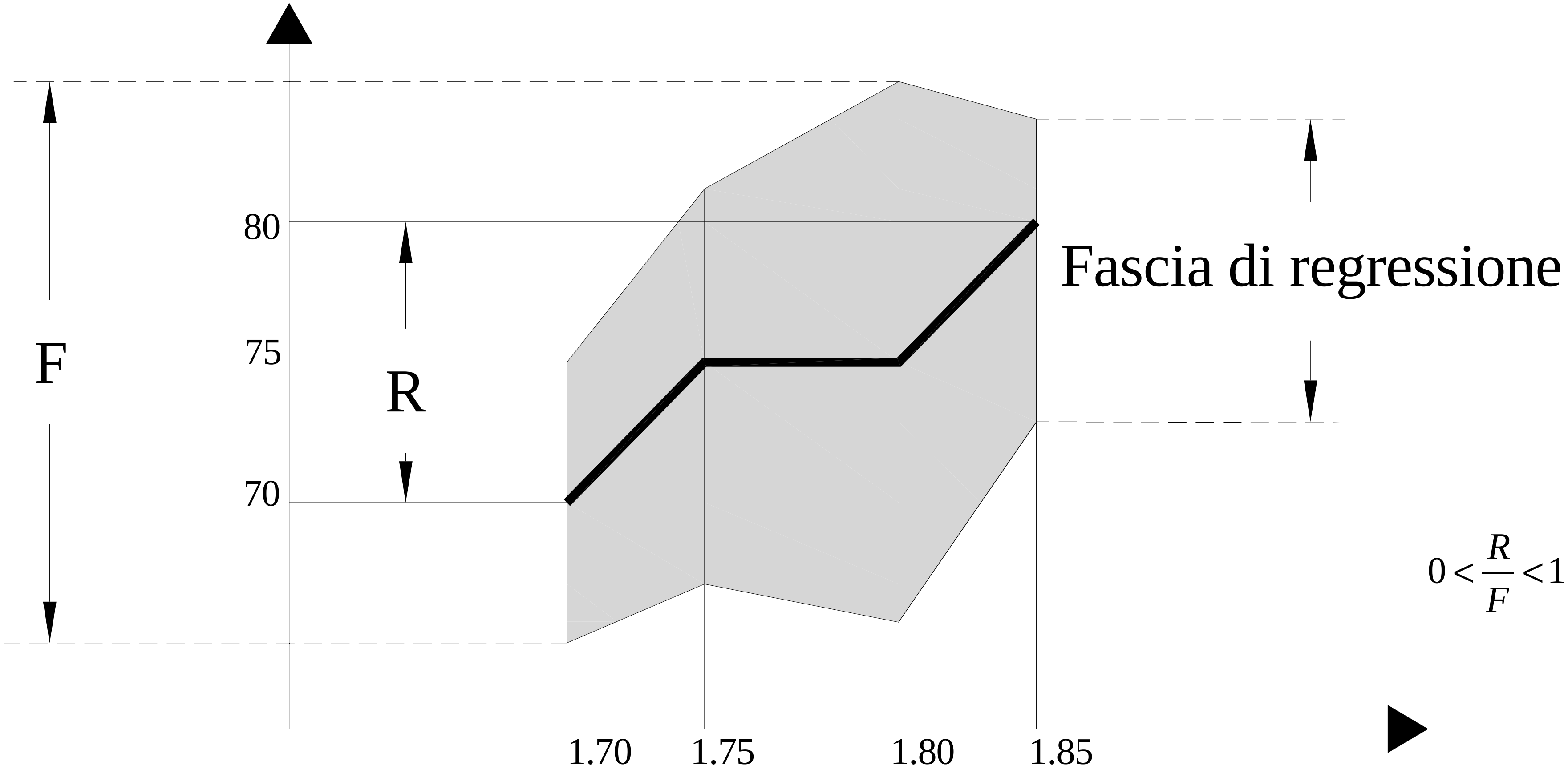


FIG.20

Si costruisce la tabella derivata della **perfetta dipendenza**:

X My_i	1,7	1,75	1,8	1,85	
70	5	0	0	0	5
76,2	0	6	0	0	6
76,25	0	0	8	0	8
80,9	0	0	0	6	6
Myi=76,1	5	6	8	6	25

$$M(y^*) = My_i = 76{,}1$$

$$\sigma^2_{y^*} = (70-76{,}1)^2 \cdot 0{,}2 + (76{,}2-76{,}1)^2 \cdot 0{,}24 + (76{,}25-76{,}1)^2 \cdot 0{,}32 + (80{,}9-76{,}1)^2 \cdot 0{,}24 = 13{,}06 = \sigma^2_{y_{pd}}$$

Da cui si può ricavare l'indice di Pearson:

$$\eta_x^2 = \frac{\sigma_{y_{pd}}^2}{\sigma_y^2} = \frac{13{,}06}{37{,}05} = 0{,}35$$

	X		1,70	1,75	1,80	1,85	1,78
y	V_y	V_x	-0,08	-0,03	0,02	0,07	
65	-11,4		2\25 **0,08**	**0**	1\25 **0,04**	**0**	3\25 **0,12**
70	-6,4		1\25 **0,04**	1\25 **0,04**	1\25 **0,04**	**0**	3\25 **0,12**
75	-1,4		2\25 **0,08**	2\25 **0,08**	2\25 **0,08**	1\25 **0,04**	7\25 **0,28**
80	3,6		**0**	3\25 **0,12**	3\25 **0,12**	2\25 **0,08**	8\25 **0,32**
85	8,6		**0**	**0**	1\25 **0,04**	3\25 **0,12**	4\25 **0,16**
76,4			5\25 **0,20**	5\25 **0,24**	8\25 **0,32**	6\25 **0,25**	

$$\sigma_{xy} = 0,283$$ Covarianza

$$r_{xy} = \frac{\sigma_{xy}}{\sigma_x \cdot \sigma_y}$$ Coefficiente di correlazione lineare

DIAGRAMMA a DISPERSIONE

Riportando i risultati dell’esempio precedente su un grafico, si avrà:

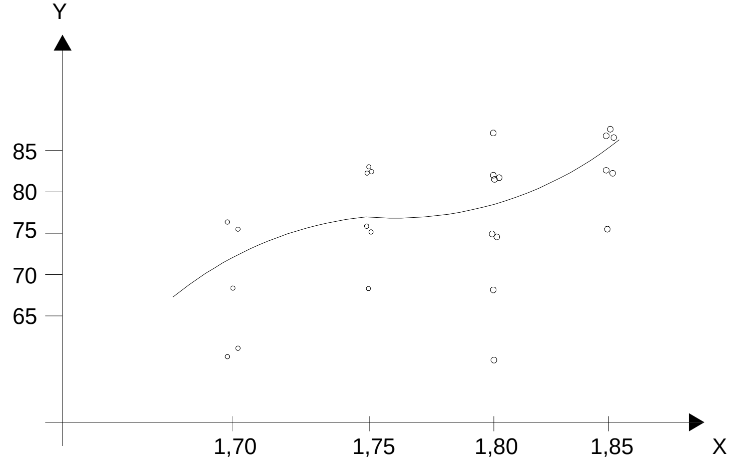
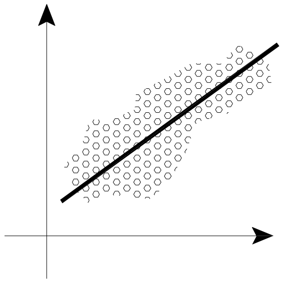


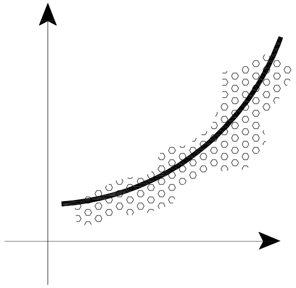
FIG.21

Si hanno diversi tipi di curva interpolatrice:



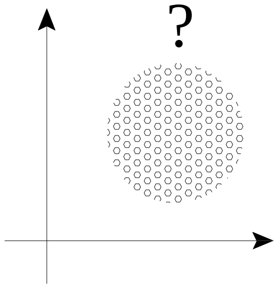
Relazione Lineare

$$y = a + bx$$



Relazione non Lineare

$$y = a + bx + cx^2$$



Esiste più di una sola curva di un certo tipo che interpola l'insieme dei dati. Per evitare valutazioni personali, è necessario definire quale può essere la "migliore curva interpolatrice": se, ad esempio, si utilizza *il metodo dei minimi quadrati*, la migliore curva interpolatrice è quella che ha la proprietà di rendere minima la quantità:

$$\boxed{d_1^2 + d_2^2 + ... + d_n^2 = \min}$$

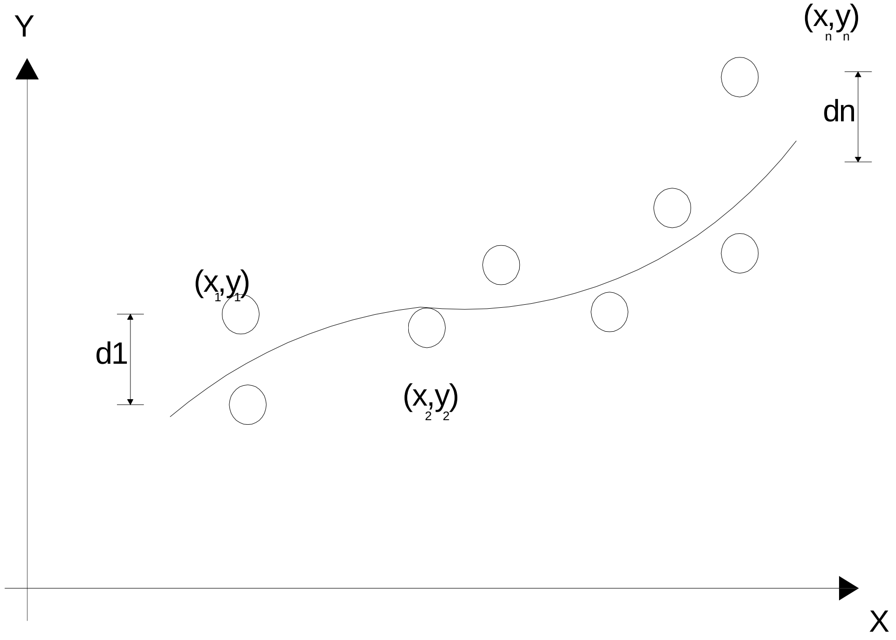


FIG.22

PROBLEMA delle STIME dei PARAMETRI di UNA DISTRIBUZIONE

Un problema rilevante della statistica matematica è quello dell'"**inferenza statistica**" cioè il problema di come si possono trovare informazioni sulla costituzione di una popolazione basandosi su *un numero limitato* di estrazioni, costituenti dei campioni, estratti a caso dalla popolazione stessa.

Il problema si può scindere in due:

- Stabilire come i parametri della popolazione si possono ricavare dal campione.
- Stabilire che precisione hanno questi parametri (qui si tralascia questa trattazione)

Si abbia un fenomeno aleatorio X definito da una distribuzione avente densità di probabilità f(x), con valori argomentali definiti in un intervallo a-b. Da questa popolazione possibile, che ha media **m** e varianza **σ^2** (valori teorici), sono stati estratti a caso **n** valori

$$x_1 \qquad x_2 \qquad \ldots \qquad x_n$$

dai quali si vogliono dedurre, nella migliore delle maniere, le **stime dei parametri teorici m** e **σ^2-**

Queste stime (o valori empirici) saranno una funzione degli n valori estratti ed in generale si potrà **porre :**

$$\overline{m} = h(x_1, x_2, \ldots, x_n)$$

$$\overline{\sigma^2} = g(x_1, x_2, \ldots, x_n)$$

dove **h** e **g** sono chiamate funzioni “estimatori”.

Vediamo quali criteri si possono usare per scelta degli estimatori:

1. Una stima è **consistente** se al tendere all’infinito del numero n di campioni estratti, essa tende al valore teorico

2. Una stima è affetta da **errore sistematico** se la media della popolazione delle stime non coincide con la media della distribuzione da cui è estratto il campione (**BIASED)**

3. Un stima è **efficiente** se fra tutte le stime confrontabili è quella che ha varianza minima

4. Una stima è più **plausibile** quando la si ottiene utilizzando il principio di **massima verisimiglianza**

Principio della Massima Verisimiglianza

Ricordiamo che la funzione **f(x)** che rappresenta una distribuzione di probabilità, dipende dai momenti della distribuzione.

Con riferimento al campione $\chi_1\ \chi_2\ \ldots\ \chi_n$ estratto a caso, si può osservare che per ogni valore **dx** è definita la probabilità infinitesima

$$dp = f(x)dx$$

Si osserva inoltre che la probabilità composta dagli **n** valori indipendenti $\chi_1\ \chi_2\ \ldots\chi_n$

vale dP= dp_1. dp_2…….dp_n e cioè:

$$dP = f(x_1) \cdot f(x_2) \cdot \ldots f(x_n) \cdot d(x)^n$$

Se in questa espressione potessimo mettere i valori teorici (incogniti) della media e della varianza, otterremmo un certo valore per **dP**; se invece introduciamo le stime (che si considerano note) di questi valori, avremo ovviamente dei valori diversi di **dP** !

Fra tutti i valori delle stime, sarà “plausibile” assumere quelli che rendono massima **dP**.

Il massimo di **dP** si ottiene quando è massima la funzione:

$$V = f(x_1) \cdot f(x_2) \cdot \ldots f(x_n)$$

detta “**funzione di VERISIMIGLIANZA**”

E cioè basta fare

$$\frac{\partial V}{\partial m} = 0 \qquad \frac{\partial V}{\partial(\sigma^2)} = 0$$

Queste equazioni consentono di ricavare $\overline{m}$ e $\overline{\sigma^2}$

Principio dei Minimi Quadrati

Data una grandezza **X** da misurare, giacchè non si può effettuare un numero troppo grande di osservazioni, si estrae un campione limitato

$$x_1 \quad x_2 \quad \ldots \quad x_n$$

Si vuole vedere come da questo campione si possano dedurre le stime $\overline{m}$ e $\overline{\sigma}^2$: si ricorre al principio di MASSIMA VERISIMIGLIANZA.

Per ogni x_i si ha la seguente funzione ***densità di probabilità*** *(ipotesi di distribuzione normale) :*

$$f(x_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{\frac{-(x_i - m)^2}{2\sigma^2}}$$

mentre la ***funzione di VERISIMIGLIANZA*** è

$$V(x_1, x_2 \ldots x_n, \sigma^2) = \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \cdot e^{\frac{-1}{2\sigma^2}\sum_{i=1}^{n}(x_i - m)^2}$$

Questa funzione diventa massima quando è minimo l’esponente dell’esponenziale e cioè

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

ovvero:

$$\boxed{\sum_{i=1}^{n} v_i^2 = \min}$$

Questo esprime **<u>il principio dei minimi quadrati</u>**: cioè nel caso di una distribuzione gaussiana, hanno massima verisimiglianza quelle stime che soddisfano alla condizione che la sommatoria dei quadrati degli scarti sia minima.

MEDIA e VARIANZA EMPIRICHE

1) Dal principio dei minimi quadrati derivando rispetto ad **m** si ha

$$\boxed{\sum_{i=1}^{n} v_i^2 = \min}$$

$$\frac{\delta}{\delta m}\left(\sum_{i=1}^{n}(x_i - m)^2\right) = -2\sum_{i=1}^{n}(x_i - \overline{m}) = 0$$

ed uguagliando a zero si ottiene

$$\sum_{i=1}^{n} x_i - n\overline{m} = 0 \qquad \overline{m} = \frac{\sum x_i}{n} = \frac{x_1 + x_2 + \ldots + x_n}{n}$$

questa stima della media è consistente e non è affetta da errore sistematico, infatti ha una media:

$$M\left[\overline{m}\right] = \frac{1}{n}\left(M\left[x_1\right] + M\left[x_2\right] + \ldots + M\left[x_n\right]\right) = n \cdot \frac{m}{n} = m$$

ed ha come varianza:

$$\sigma^2\left(\overline{m}\right) = \frac{\sigma^2}{n} = \frac{\sigma^2 + \sigma^2 + \ldots + \sigma^2}{n^2} = \frac{n\sigma^2}{n^2}$$

2) Per la stima della σ^2 consideriamo invece la **V**

$$V = f\left(x_1\right)\cdot f\left(x_2\right)\cdot\ldots\ldots\cdot f\left(x_n\right) = \frac{1}{\left(2\Pi\sigma^2\right)^{\frac{n}{2}}}\cdot e^{\frac{-1}{2\sigma^2}\sum_{i=1}^{n}\left(x_i - m\right)^2}$$

Operiamo una trasformazione con il seguente algoritmo per semplificare il calcolo delle derivate

$$\ln V = -\frac{n}{2}\ln 2\Pi - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\cdot\sum_{i=1}^{n}\left(x_i - m\right)^2$$

Anche adesso avrà massima verisimiglianza per quel valore della stima σ^2 che rende max **lnV**:

$$\frac{\partial\ln V}{\partial\sigma^2} = 0$$

$$\frac{\partial\ln V}{\partial\sigma^2} = 0$$

$$\ln V = -\frac{n\,2\hat{\sigma}}{2\hat{\sigma}^2} + \frac{1\cdot 2\hat{\sigma}}{2\hat{\sigma}^4}\cdot\sum_{i=1}^{n}\left(x_i - \overline{m}\right)^2 = 0$$

$$-n\hat{\sigma}^2 + \sum_{i=1}^{n}\left(x_i - \overline{m}\right)^2 = 0$$

Da cui

$$\boxed{\hat{\sigma}^2 = \sum_{i=1}^{n}\left(x_i - \overline{m}\right)^2\cdot\frac{1}{n}}$$

questa stima della varianza è però affetta da errore sistematico

Infatti

$$\frac{1}{n}\sum_{i=1}^{n}\left(\left(x_i-m\right)+\left(m-\overline{m}\right)\right)^2=\frac{1}{n}\sum_{i=1}^{n}\left(x_i-m\right)^2+\frac{1}{n}\sum_{i=1}^{n}\left(m-\overline{m}\right)^2+\frac{2}{n}\sum_{i=1}^{n}\left(x_i-m\right)\cdot\left(m-\overline{m}\right)$$

Prendiamo in esame il termine:

$$\frac{2}{n}\sum_{i=1}^{n}\left(x_i-m\right)\cdot\left(m-\overline{m}\right)=\frac{2}{n}\left(m-\overline{m}\right)\sum_{i=1}^{n}\left(x_i-m\right)=\frac{2}{n}\left(m-\overline{m}\right)\left(\sum_{i=1}^{n}x_1-nm\right)=-2\left(\overline{m}-m\right)^2$$

E quindi

$$\frac{1}{n}\sum_{i=1}^{n}\left(x_i-m\right)^2+\frac{1}{n}\sum_{i=1}^{n}\left(m-\overline{m}\right)^2-2\left(\overline{m}-m\right)^2=$$

$$\frac{1}{n}\sum_{i=1}^{n}\left(x_i-m\right)^2+\frac{1}{n}n\left(m-\overline{m}\right)^2-2\left(\overline{m}-m\right)^2=$$

$$\hat{\sigma}^2=\frac{1}{n}\sum_{i=1}^{n}\left(x_i-m\right)^2-\left(\overline{m}-m\right)^2$$

Calcoliamo la media della popolazione $\hat{\sigma}^2$

$$M\left[\hat{\sigma}^2\right] = \frac{1}{n} M \sum_{i=1}^{n} \left[\left(x_i - m\right)^2\right] - M\left[\left(\overline{m} - m\right)^2\right] = \sigma^2 - \frac{\sigma^2}{n} = \sigma^2 \cdot \frac{n-1}{n}$$

Cioè la stima della $\hat{\sigma}^2$ è affetta da errore sistematico (la media di $\hat{\sigma}^2$ **non** coincide con σ^2)

Invece **non è** affetta da errore sistematico la seguente stima:

$$\boxed{\overline{\sigma}^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(x_i - \overline{m}\right)^2}$$

Concludendo, per determinare la misura di una grandezza **X** si eseguono **n** misure X1 X2 ... Xn

Si assume come stima della media della popolazione delle misure possibili la

$$\overline{m} = \frac{x_1 + x_2 + ... + x_n}{n}$$

Si calcola la varianza stimata:

$$\overline{\sigma}^2 = \frac{\sum_{i=1}^{n} v_i^2}{n-1}$$

E lo scarto quadratico medio (s.q.m.):

$$\overline{\sigma} = \sqrt{\frac{\sum_{i=1}^{n} v_i^2}{n-1}}$$

Si calcola lo s.q.m. della popolazione delle medie aritmetiche:

$$\overline{\sigma}_m = \frac{\overline{\sigma}}{\sqrt{n}} = \sqrt{\frac{\sum_{i=1}^{n} v_i^2}{n \cdot (n-1)}}$$

E quindi si scrive

$$\boxed{X = \overline{m} \pm \overline{\sigma}_m}$$

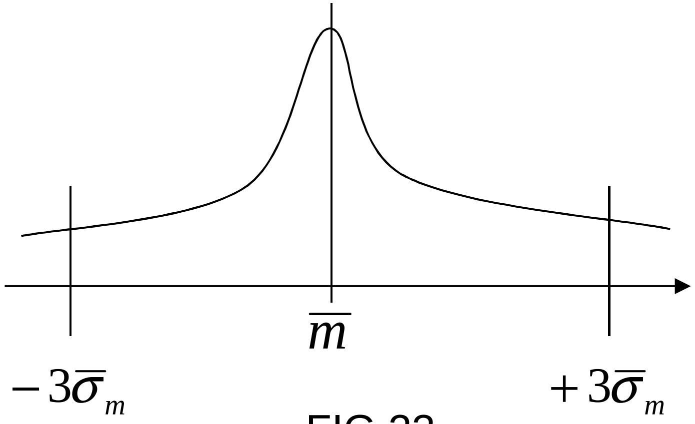


FIG.23

Alle volte, si conosce la varianza della distribuzione delle misure possibili caratteristica dello "strumento - metodo - ambiente" σ_s^2 (fornita dal *costruttore*)

Si può scrivere allora che:

$$X = \overline{m} \pm \frac{\sigma_s}{\sqrt{n}}$$

Si dovrebbe verificare (all'aumentare del numero di osservazioni) che σ_s somigli a $\overline{\sigma}$: notare che l'uso di σ_s non consente nessun controllo circa la precisione effettivamente ottenuta nelle misure fatte.

ESEMPIO di CALCOLO della MEDIA e VARIANZA, EMPIRICHE o STIMATE

La grandezza X, un angolo, è stato misurato 18 volte, con uguale precisione: calcolare la media, l’errore medio di una osservazione e l’errore medio delle media

Valori osservati x_i		Scarti		V_i^2
		-	+	
48°21’	18”		+0”.5	0.25
48°21’	15”	-2”.5		6.25
48°21’	20”		+2”.5	6.25
48°21’	17”	-0”.5		0.25
48°21’	18”		+0”.5	0.25
48°21’	21”		+3”.5	12.25
48°21’	17”	-0”.5		0.25
48°21’	19”		+1”.5	2.25
48°21’	14”	-3”.5		12.25
48°21’	16”	1”.5		2.25

Valori osservati x_i		Scarti		V_i^2
		-	+	
48°21’	15”	-2”.5		6.25
48°21’	19”		+1”.5	2.25
48°21’	16”	-1”.5		2.25
48°21’	20”		+2”.5	6.25
48°21’	18”		+0”.5	0.25
48°21’	17”	-0”.5		0.25
48°21’	19”		+1”.5	2.25
48°21’	16”	-1”.5		2.25
$\bar{x}_m = \sum \frac{x_i}{n} = 48°21'17''.5$		$\sum v_i \cong 0$		$\sum v_i^2 = 64,50$

S.q.m. di una singola osservazione:

$$\overline{\sigma} = \pm\sqrt{\frac{\sum v_i^2}{n-1}} = \pm\sqrt{\frac{64.50}{17}} = \pm\sqrt{3.794} = 1''.95$$

S.q.m. della media

$$\overline{\sigma_m} = \pm\frac{\sigma}{\sqrt{n}} = \pm\frac{1''.95}{\sqrt{18}} = 0''.46$$

Il valore attendibile dell’angolo X è:

$$X = \overline{x_m} \pm \overline{\sigma_m} = 48°21'7''.5 \pm 0''.46$$

$$0''.5$$

Grafico di $\overline{\sigma}$ e $\overline{\sigma}_m$.

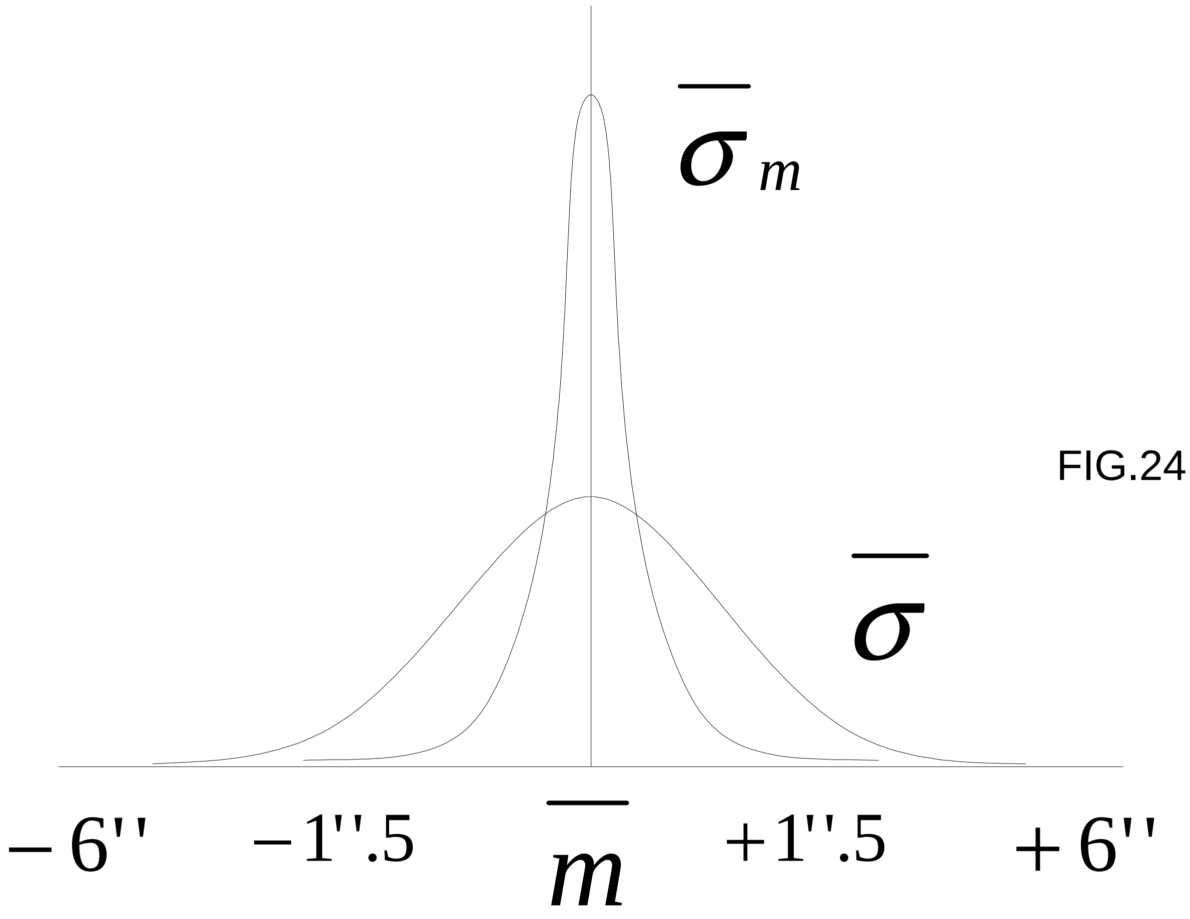


FIG.24

Media Ponderata

Una grandezza X viene misurata con strumenti, metodi, condizioni ambientali diverse:

$\overline{m_1} \pm \overline{\sigma_1}$ $\qquad$ $\overline{m_2} \pm \overline{\sigma_2}$ $\qquad$ $...\overline{m_n} \pm \overline{\sigma_n}$

Si vuole determinare la media comune alle varie distribuzioni, utilizzando il principio di massima verisimiglianza:

$$f\left(\overline{m_i}\right) = \frac{1}{\sqrt{2\pi\overline{\sigma_i^2}}} \cdot e^{\frac{-\left(\overline{m_i}-m\right)^2}{2\overline{\sigma_i^2}}}$$

(funzione densità di probabilità dell'evento $\overline{m_i}$)

che si può scrivere introducendo il concetto di peso

$$\boxed{p_i = \frac{\sigma_0^2}{\overline{\sigma_i^2}}}$$

$$f\left(\overline{m_i}\right) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \cdot \sqrt{p_i} \cdot e^{\frac{-p_i}{2\sigma_0^2}\left(\overline{m_i}-m\right)^2}$$

σ_0^2 = varianza dell’unità di peso.

Quindi:

$$V = \left(\overline{m_1}, \overline{m_2}, \dots, \overline{m_n}, m, \sigma_0^2\right) = \frac{\left(p_1, p_2, \dots, p_n\right)^{\frac{1}{2}}}{\left(2\pi\sigma_0^2\right)^{\frac{n}{2}}} e^{\frac{-1}{2\sigma_0^2}\sum_{i=1}^{n} p_i\left(\overline{m_i}-m\right)^2}$$

Il massimo della V si ottiene quando $\sum_{i=1}^{n} p_i \left(\overline{m_i} - m\right)^2 = \min$

Da cui $\sum_{i=1}^{n} p_i \overline{m_i} - m_p \sum_{i=1}^{n} p_i = 0$

La Media ponderata vale

$$\boxed{m_p = \frac{p_i \overline{m_i} + p_2 \overline{m_2} + \ldots + p_n \overline{m_n}}{p_1 + p_2 + \ldots + p_n}}$$

Si dimostra che la media ponderata è una buona stima (si tralascia la dimostrazione); Invece per la **varianz**a, così come visto in precedenza per $\hat{\sigma}^2$, che la stima è affetta da errore sistematico ed anche qui conviene assumere :

$$\overline{\sigma_o}^2 = \frac{\sum p_i \left(\overline{m_i} - m_p\right)^2}{n-1}$$

Osservazioni

La varianza dell’unità di peso σ_o^2

A priori (arbitraria)

A posteriori calcolata con:

$$\overline{\sigma_o^2} = \frac{\sum_{i=1}^{n} p_i \left(\overline{m_i} - m_p \right)^2}{n-1}$$

Il “peso” è adimensionale. E’ un fattore che permette in qualche modo di comparare tra loro le misure fatte, assegnando a ciascuna un valore arbitrario, che ne indica l’affidabilità o il peso: valori di peso bassi fanno contare meno, mentre valori di peso alti,permettono di fare contare di più le misure stimate migliori.

La varianza della media ponderata $\overline{\sigma_{mp}^{2}}$

Ricavata dal $\overline{\sigma_0^{2}}$ a priori = teoricamente dipendente dalle varianze

$\sigma_1^2, \sigma_2^2, ..., \sigma_n^2$

Ricavata dal $\overline{\sigma_o^{2}}$ a posteriori = dà l'idea della congruenza nelle varie determinazioni

$\overline{m_1}, \overline{m_2}, ... \overline{m_n}$

Procedura di calcolo

- Si fissa a piacere σ_o^2
- Si calcolano i pesi $p_i = \dfrac{\sigma_o^2}{\sigma_i^2}$
- Si calcola la media ponderata $m_p = \dfrac{p_1\overline{m_1} + p_2\overline{m_2} + \dots + p_n\overline{m_n}}{p_1 + p_2 + \dots + p_n}$
- Si calcola $\overline{\sigma_o^2} = \dfrac{\sum_{i=1}^{n} p_i\left(\overline{m}_i - m_p\right)^2}{n-1}$
- Si calcola $\overline{\sigma_{mp}^2} = \dfrac{\overline{\sigma_0^2}}{\sum p_i}$

$$\boxed{X = m_p \pm \overline{\sigma}_{mp}}$$

ESEMPIO DI MEDIA PONDERATA

Sia la grandezza X (ad esempio un angolo): si sono effettuate delle osservazioni con strumentazioni diverse, ottenendo le tre seguenti serie o popolazioni:

$\overline{m_1} \pm \overline{\sigma_1}$ 48°24’10’’ ± 5’’	$\overline{m_2} \pm \overline{\sigma_2}$ 48°24’25’’ ± 10’’	$\overline{m_3} \pm \overline{\sigma_3}$ 48°24’20’’ ± 15’’
Calcolo i pesi:		
$p_1 = \frac{\sigma_0{}^2}{\overline{\sigma_1^2}} = \frac{k}{25} = 9$	$p_2 = \frac{\sigma_0{}^2}{\overline{\sigma_2^2}} = \frac{k}{100} = 2{,}25$	$p_3 = \frac{\sigma_0{}^2}{\overline{\sigma_3^2}} = \frac{k}{225} = 1$
Ponendo k= σ_0^2 =225 (a priori)		

$\overline{m_i}$	p_i	$\overline{m_i} \cdot p_i$	$v_i = \overline{m_i} - p_i$		v_i^2	$p_i \cdot v_i^2$
10’’	9	90		-3,6	12,96	117
25’’	2,25	56,2	+11,4		112,9	276
20’’	1	20	+6,4		40,9	41

$$\sum p_i = 12{,}2 \qquad \sum \overline{m_i} \cdot p_i = 166{,}2 \qquad \sum p_i \cdot v_i^2 = 434$$

1. $$m_p = \frac{\sum_{i=1}^{n} p_i \cdot \overline{m_i}}{\sum p_i} = \frac{116{,}2}{12{,}2} = 13''.6$$

$$m_p = 48°24'13''.6$$

2.A posteriori

$$\overline{\sigma_0^2} = \sum_{i=1}^{n} \frac{p_i(\overline{m_i} - m_p)^2}{n-1} = \frac{432}{2} = 217$$

3. $$\overline{\sigma}^2_{m_p} = \frac{\overline{\sigma}_0^2}{\sum p_i} = \frac{217}{12,2} = 17,8$$

4. $$\overline{\sigma}_{m_p} = \pm 4''.2$$

In definitiva:

$$X = m_p + \overline{\sigma}_{m_p} = 48°24'13''.6 \pm 4''.2$$

Alcune brevi considerazioni sul trattamento delle misure.

Le misure di grandezze a rigore non sono proprio delle *estrazioni a caso* da *popolazioni possibili*.

Occorre precisare infatti che le estrazioni intanto non sono *a caso (aleatorie)* ma sono operazioni fatte di solito con attenzione ed inoltre che la popolazione possibile da cui si estrae deve essere pensata come una *popolazione coesa*, con individui molto simili. La si potrebbe però pensare (è una popolazione virtuale !) come si vuole, anche con individui molto diversi tra loro: ma non è il caso che interessa le misure topografiche . Quindi la si deve pensare come una popolazione possibile ma coesa. Ma coesa quanto? Dipende dalla strumentazione di misura, dalle condizioni ambientali in cui si opera, dall'operatore e dalla sua abilità e capacità nell'usare la strumentazione. Tutte informazioni che ci vengono fornite dalle stesse misure fatte, con i valori significativi ricavati (utilizzando la statistica) della *media* e della *varianza*. Alcuni studiosi rifiutano o sono molto critici su questo uso della Statistica: rifiutano infatti il concetto di popolazione possibile accoppiato al concetto di una grandezza: una grandezza ha per loro un unico valore di misura (non è pensabile che sia paragonata ad un qualcosa che pulsa e/o sia variabile). La fluttuazione accidentale è una manifestazione dell'incapacità dell'uomo e delle sue strumentazioni, di determinare quest'**unico valore**. Quindi non c'è una popolazione di individui della grandezza ma una popolazione di individui che nascono dalla misura della grandezza, fornita da strumentazioni incapaci di determinarne il vero valore. In definitiva, la popolazione si sposta dalla grandezza alla misura della grandezza, ma si ha sempre a che fare con una popolazione di individui (valori o singole osservazioni) che bisogna sapere trattare. Né si può pensare che questa *popolazione strumentale* sia sempre la stessa e quindi individuabile una volta per tutte perché cambierà per effetto degli altri parametri (ambiente ed operatore). Si parla in effetti, e ne abbiamo accennato, di un **σ_s** fornito dai costruttori degli strumenti, però non significativo per le operazioni di misura praticamente effettuate.

MISURE INDIRETTE

Abbiamo finora trattato le MISURE DIRETTE di grandezze, anche con peso diverso.Esiste però un ampio campo che è quello delle MISURE INDIRETTE di grandezze.

Si possono verificare i seguenti tre casi:

1. la grandezza incognita X da misurare indirettamente è funzione di n grandezze $X_1, X_2, ..., X_n$ misurate direttamente:

$$X = f(X_1, X_2, ..., X_n)$$

2. le grandezze incognite $X_1, X_2, ..., X_n$ sono tante quante le grandezze $L_i, M_i, ..., N_i$ misurate direttamente:(caso dell'equilibrio)

$$f_i(X_1, X_2, ..., X_m / L_i, M_i, ..., N_i) = 0 \quad \text{con i =1,2,…,m}$$

3. le grandezze misurate direttamente $L_i, M_i, ..., N_i$ sono in numero maggiore delle grandezze incognite $X_1, X_2, ..., X_m$:(caso della sovrabbondanza)

$$f_i(X_1, X_2, ..., X_m / L_i, M_i, ..., N_i) = 0 \quad \text{con i =1,2,…,n} \quad n > m$$

1°CASO: sottocaso lineare

$$X = aX_1 + bX_2 + cX_3$$

Le grandezze X_1, X_2,X_3 sono state misurate direttamente, cioè conosciamo la loro *media stimata* e la loro *varianza stimata*.

Vogliamo determinare la popolazione di misure possibili che definisce la X, cioè determinare media stimata e la varianza (ipotesi di distribuzione normale):
il sottocaso si risolve facilmente, con le regole viste nel caso di combinazioni di variabili casuali:

$$\overline{X_m} = a\overline{X_{1m}} + b\overline{X_{2m}} + c\overline{X_{3m}}$$

$$\overline{\sigma^2}_{\mathrm{x}} = a^2\overline{\sigma}^2_{x_1} + b^2\overline{\sigma}^2_{x_2} + c^2\overline{\sigma}^2_{x_3}$$

(Grandezze tra loro indipendenti)

1°CASO: sottocaso non lineare

$$X = f(X_1, X_2 ... X_n)$$

Siano $O_1, O_2 \ldots O_n$ delle stime delle medie teoriche X_{1m} , $X_{2m} \ldots$ X_{nm}.
Si può allora scrivere

$O_1 = X_{1m} + \nu_1$
$O_2 = X_{2m} + \nu_2$
……………..
$O_n = X_{nm} + \nu_n$

essendo gli scarti ν_i sufficientemente piccoli.

$$f(O_1, O_2 ... O_n) = f(X_{1m} + \nu_1, X_{2m} + \nu_2, ... X_{nm} + \nu_n) = f(X_{1m}, X_{2m} ... X_{nm}) + \left(\frac{\partial f}{\partial x_1}\right)_m \cdot \nu_1 + \left(\frac{\partial f}{\partial x_2}\right)_m \cdot \nu_2 + ... + \left(\frac{\partial f}{\partial x_n}\right)_m \cdot \nu_n$$

(sviluppo in serie di Taylor).

Si può, quindi, scrivere che:

$$f(O_1, O_2 ... O_n) - f(X_{1m}, X_{2m} ... X_{nm}) = \left(\frac{\partial f}{\partial x_1}\right)_m \cdot \nu_1 + \left(\frac{\partial f}{\partial x_2}\right)_m \cdot \nu_2 + ... + \left(\frac{\partial f}{\partial x_n}\right)_m \cdot \nu_n$$

cioè la differenza delle due funzioni è una combinazione lineare degli scarti ν_i .

Possiamo calcolarci la media della differenza delle due funzioni :

$$M[f(O_1, O_2 ... O_n) - f(X_{1m}, X_{2m} ... X_{nm})] = \left(\frac{\partial f}{\partial x_1}\right)_m \cdot M[\nu_1] + \left(\frac{\partial f}{\partial x_2}\right)_m \cdot M[\nu_2] + ... = 0$$

e cioè

$$f(O_1, O_2 ... O_n) \cong f(X_{1m}, X_{2m} ... X_{nm}) = X_m$$

dal che si vede che la media teorica di X si otterrebbe introducendo le medie teoriche delle X_i e che questa equivale anche ad introdurre le stime di O_i.

Quindi:

$$\boxed{\overline{X}_m = f(O_1, O_2 ... O_n) \cong f(\overline{X}_1, \overline{X}_2 ... \overline{X}_n)}$$

Per la varianza si ha:

$$\sigma_x^2 = \left(\frac{\partial f}{\partial x_1}\right)_m^2 \cdot \sigma^2_{x_1} + \left(\frac{\partial f}{\partial x_2}\right)_m^2 \cdot \sigma^2_{x_2} + \ldots + \left(\frac{\partial f}{\partial x_n}\right)_m^2 \cdot \sigma^2_{x_n}$$

e se si utilizzano le varianze stimate:

$$\boxed{\overline{\sigma}_x^2 = \left(\frac{\partial f}{\partial x_1}\right)_m^2 \cdot \overline{\sigma}^2_{x_1} + \left(\frac{\partial f}{\partial x_2}\right)_m^2 \cdot \overline{\sigma}^2_{x_2} + \ldots + \left(\frac{\partial f}{\partial x_n}\right)_m^2 \cdot \overline{\sigma}^2_{x_n}}$$

(<u>legge della propagazione degli s.q.m.</u>) nel caso in cui le grandezze X_1, X_2,…X_n siano indipendenti.

2°CASO: sottocaso lineare

m

$$m\left\{\begin{array}{c} a_1 X_1 + b_1 X_2 + \ldots u_1 X_m = T_1 \\ a_2 X_1 + b_2 X_2 + \ldots u_2 X_m = T_2 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_m X_1 + b_m X_2 + \ldots u_m X_m = T_m \end{array}\right\}$$

Si è supposto che nel sistema ogni equazione contenga una sola misura diretta (T_i): ciò non lede la generalità della trattazione.
Si consideri la matrice dei coefficienti e la sua inversa

$$C=\begin{pmatrix} a_1 & b_1 & \ldots & u_1 \\ a_2 & b_2 & \ldots & u_2 \\ \ldots.. & \ldots.. & \ldots.. & \ldots. \\ a_m & b_m & \ldots & u_m \end{pmatrix} \quad C^{-1}=\begin{pmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1m} \\ a_{21} & \alpha_{22} & \ldots & \alpha_{2m} \\ \ldots.. & \ldots.. & \ldots.. & \ldots. \\ \alpha_{m1} & \alpha_{m2} & \ldots & \alpha_{mm} \end{pmatrix}$$

Essendo $\alpha_{ij} = (-1)^{i+j} \cdot \frac{A_{ji}}{D}$

A_{ji} = **complemento algebrico del termine di posto ji**

D=determinante della matrice C

La matrice inversa è chiamata "<u>soluzione indeterminata</u>" del sistema.

Le soluzioni del sistema sono, per la regola di Cramer:

$$\begin{cases} x_1 = \alpha_{11} \cdot t_1 + \alpha_{12} \cdot t_2 + ... + \alpha_{1m} \cdot t_m \\ x_2 = \alpha_{21} \cdot t_1 + \alpha_{22} \cdot t_2 + ... + \alpha_{2m} \cdot t_m \\ .... \\ x_m = \alpha_{m1} \cdot t_1 + \alpha_{m2} \cdot t_2 + ... + \alpha_{mm} \cdot t_m \end{cases}$$

E quindi per le medie stimate:

$$\begin{cases} \overline{x_{1m}} = \alpha_{11} \cdot \overline{T_{1m}} + \alpha_{12} \cdot \overline{T_{2m}} + ... + \alpha_{1m} \cdot \overline{T_{mm}} \\ \overline{x_{2m}} = \alpha_{21} \cdot \overline{T_{1m}} + \alpha_{22} \cdot \overline{T_{2m}} + ... + \alpha_{2m} \cdot \overline{T_{mm}} \\ .... \\ \overline{x_{mm}} = \alpha_{m1} \cdot \overline{T_{1m}} + \alpha_{m2} \cdot \overline{T_{2m}} + ... + \alpha_{mm} \cdot \overline{T_{mm}} \end{cases}$$

E per le varianze stimate

$$\begin{cases} \overline{\sigma_{x1}^2} = \alpha_{11}^2 \cdot \overline{\sigma_{T1}^2} + \alpha_{12}^2 \cdot \overline{\sigma_{T2}^2} + ... + \alpha_{1m}^2 \cdot \overline{\sigma_{Tm}^2} \\ \overline{\sigma_{x2}^2} = \alpha_{21}^2 \cdot \overline{\sigma_{T1}^2} + \alpha_{22}^2 \cdot \overline{\sigma_{T2}^2} + ... + \alpha_{2m}^2 \cdot \overline{\sigma_{Tm}^2} \\ .... \\ \overline{\sigma_{xm}^2} = \alpha_{m1}^2 \cdot \overline{\sigma_{T1}^2} + \alpha_{m2}^2 \cdot \overline{\sigma_{T2}^2} + ... + \alpha_{mm}^2 \cdot \overline{\sigma_{Tm}^2} \end{cases}$$

Ed il primo sottocaso è così risolto.

2°CASO: sottocaso non lineare

$$f_i(X_1, X_2 ... X_m \setminus L_i, M_i, ..., N_i) = 0 \qquad \text{i = 1,2,..m}$$

Occorre effettuare la *linearizzazione* del sistema. Supponiamo noti dei valori approssimati delle incognite

$$X^0_{1m}, X^0_{2m}, ...$$

Si può porre:

$$X_{1m} = X^0_{1m} + x_1 \qquad X_{2m} = X^0_{2m} + x_2 \quad .... \qquad X_{mm} = X^0_{mm} + x_m$$

essendo le correzioni x_i incognite,ma piccole, tali che i quadrati e le potenze superiori siano trascurabili. Si ha pertanto:

$$f_i\left(X^0_{1m} + x_1, X^0_{2m} + x_2, ... X^0_{mm} + x_m / L_i, M_i, ..., N_i\right) = 0 \qquad i = 1,2,...m$$

E sviluppando in serie con sviluppo arrestato ai termini lineari:

$$f_i\left(X^0_{1m}, X^0_{2m} \dots X^0_{mm} \,/\, \overline{L}_i, \overline{M}_i, \dots, \overline{N}_i\right) + \left(\frac{\partial f_i}{\partial x_1}\right)_0 \cdot x_1 + \left(\frac{\partial f_i}{\partial x_2}\right)_0 \cdot x_2 + \dots + \left(\frac{\partial f_i}{\partial x_m}\right)_0 \cdot x_m = 0$$

e ponendo $f_i\left(X^0_{1m}, X^0_{2m} \dots X^0_{mm} \,/\, \overline{L}_i, \overline{M}_i, \dots, \overline{N}_i\right) = T_i$

$$\left(\frac{\partial f_i}{\partial x_1}\right) = a_i \qquad \left(\frac{\partial f_i}{\partial x_2}\right)_0 = b_i \dots\dots \left(\frac{\partial f_i}{\partial x_m}\right)_0 = u_i$$

Si ottiene:

$$\begin{cases} a_1 x_1 + b_1 x_2 + \dots + u_1 x_m = T_1 \\ a_2 x_1 + b_2 x_2 + \dots + u_2 x_m = T_2 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_m x_1 + b_m x_2 + \dots + u_m x_m = T_m \end{cases}$$

Cioè si ha il sistema visto nel 1° sottocaso e già trattato.

In questo caso però le T_i non sono delle misure dirette, ma sono funzioni di misure dirette e quindi vale:

$$\overline{\sigma}_{Ti}^2 = \left(\frac{\partial f_i}{\partial L_i}\right)_m^2 \cdot \overline{\sigma}_{Li}^2 + \left(\frac{\partial f_i}{\partial M_i}\right)^2 \cdot \overline{\sigma}_{Mi}^2 + \ldots + \left(\frac{\partial f_i}{\partial N_i}\right)^2 \cdot \overline{\sigma}_{Ni}^2$$

Anche qui si ipotizza che L_i, M_i, ….,N_i siano tra loro indipendenti e che tra loro indipendenti siano anche i vari T_i.

3° caso : Osservazione indiretta di m grandezze mediante un numero esuberante di misure dirette.

1° sottocaso lineare

$$
n \left| \begin{array}{l}
\overset{m}{\left\{ \begin{array}{l}
a_1 X_1 + b_1 X_2 + \dots + u_1 X_m = T_1 \\
a_2 X_1 + b_2 X_2 + \dots + u_2 X_m = T_2 \\
\dots\dots\dots\dots \quad \dots\dots\dots\dots \quad \dots\dots\dots\dots \quad \dots\dots\dots\dots \\
a_m X_1 + b_m X_2 + \dots + u_m X_m = T_m
\end{array} \right.} \\
\dots\dots\dots\dots \quad \dots\dots\dots\dots \quad \dots\dots\dots\dots \quad \dots\dots\dots \\
a_n X_1 + b_n X_2 + \dots + u_n X_m = T_n
\end{array} \right.
$$

Se introduciamo i valori stimati delle T_i , si ottiene un sistema incompatibile ed occorre introdurre degli scarti v_i incogniti:

$$a_i X_{1m} + b_i X_{2m} + \dots u_i X_{mm} - \overline{T}_{im} = v_i$$

Ma il sistema da incompatibile diventa indeterminato (m + n incognite in n equazioni).
Osserviamo però che le medie X_{im} possono essere trovate applicando il criterio di stima del “principio di massima verosimiglianza” (infatti il campione disponibile dei T_i è esuberante, essendo essi funzioni dei valori delle X_{im}).

Ricordiamo che la funzione densità di probabilità del dato $\overline{T}_{im}$ è:

$$\varphi\left(\overline{T}_{im}, X_{1m}, X_{2m}, \ldots X_{mm}, T_{im}, \sigma_{Ti}^2\right) = \frac{1}{\sqrt{2\pi\sigma_{Ti}^2}} \cdot e^{\frac{-(\overline{T}_{im} - T_{im})^2}{2\sigma_{Ti}^2}}$$

Ed introducendo il concetto di peso

$$p_i = \frac{\sigma_0^2}{\sigma_{Ti}^2}$$

$$\varphi = \left(\frac{p_i}{2\pi\sigma_0^2}\right)^{\frac{1}{2}} \cdot e^{\frac{-p_i(\overline{T}_{im} - T_{im})^2}{2\sigma_0^2}}$$

La funzione di verosimiglianza V vale:

$$V = \left(2\pi\sigma_0^2\right)^{-\frac{n}{2}} \cdot \left(p_1 p_2 \ldots.. p_n\right)^{\frac{1}{2}} \cdot e^{-\frac{1}{2\sigma_0^2}\sum_{i=1}^{n} p_i \left(\overline{T}_{im} - T_{im}\right)^2}$$

Sarà verosimile assumere quei valori dei parametri X_{1m}, X_{2m} … che rendano massima la V: questo massimo si ha quando

$$\sum_{i=1}^{n} p_i \left(\overline{T}_{im} - T_{im}\right)^2 = \min \quad \text{o anche} \quad \sum_{i=1}^{n} p_i \cdot v_i^2 = \min$$

Il minimo si ha per quei particolari valori T_{im}^*

(essendo $T^*_{im}=a_i X^*_{1m}+b_i X^*_{2m}+\dots u_i X^*_{mm}$)

Che soddisfano il sistema

$$m\begin{cases} \dfrac{\partial}{\partial X^*_{1m}}\left\{\sum_{i=1}^{n} p_i\left(\overline{T}_{im}-T^*_{im}\right)^2\right\}=0 \\ \dfrac{\partial}{\partial X^*_{2m}}\left\{\sum_{i=1}^{n} p_i\left(\overline{T}_{im}-T^*_{im}\right)^2\right\}=0 \\ \dots\dots\dots\ \dots\dots\dots\ \dots\dots\dots\ \dots\dots\dots\ \dots \\ \dfrac{\partial}{\partial X^*_{mm}}\left\{\sum_{i=1}^{n} p_i\left(\overline{T}_{im}-T^*_{im}\right)^2\right\}=0 \end{cases}$$

Queste m equazioni aggiunte alle n equazioni del sistema generato risolvono completamente il problema.

Sviluppando il sistema prima scritto:

$$m \begin{cases} -\ 2\sum_{i=1}^{n} p_i \left(\overline{T}_{im} - T^*_{im}\right)^2 \dfrac{\partial T^*_{im}}{\partial X^*_{1m}} = 0 \\ -\ 2\sum_{i=1}^{n} p_i \left(\overline{T}_{im} - T^*_{im}\right)^2 \dfrac{\partial T^*_{im}}{\partial X^*_{2m}} = 0 \\ \quad \ldots\ldots\ldots\ldots \quad \ldots\ldots\ldots\ldots \quad \ldots\ldots\ldots\ldots \quad \ldots\ldots\ldots \\ -\ 2\sum_{i=1}^{n} p_i \left(\overline{T}_{im} - T^*_{im}\right)^2 \dfrac{\partial T^*_{im}}{\partial X^*_{mm}} = 0 \end{cases}$$

ed essendo $\dfrac{\partial T^*_{im}}{\partial X^*_{1m}} = a_i \qquad \dfrac{\partial T^*_{im}}{\partial X^*_{2m}} = b_i \ldots\ldots\ldots \qquad \dfrac{\partial T^*_{im}}{\partial X^*_{mm}} = u_i$

Si ottiene:

$$m\begin{cases} p_1\left(\overline{T}_{1m}-T^*_{1m}\right)\cdot a_1+p_2\left(\overline{T}_{2m}-T^*_{2m}\right)\cdot a_2+\ldots+p_n\,(\overline{T}_{nm}-T^*_{nm})\cdot a_n=0 \\ p_1\left(\overline{T}_{1m}-T^*_{1m}\right)\cdot b_1+p_2\left(\overline{T}_{2m}-T^*_{2m}\right)\cdot b_2+\ldots+p_n\,(\overline{T}_{nm}-T^*_{nm})\cdot b_n=0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ p_1\left(\overline{T}_{1m}-T^*_{1m}\right)\cdot u_1+p_2\left(\overline{T}_{2m}-T^*_{2m}\right)\cdot u_2+\ldots+p_n\,(\overline{T}_{nm}-T^*_{nm})\cdot u_n=0 \end{cases}$$

Prendiamo in esame ad esempio la prima equazione e sviluppiamola:

$$p_1\overline{T}_{1m}a_1-p_1T^*_{1m}a_1+p_2\overline{T}_{2m}a_2-p_2T^*_{2m}a_2+\ldots+p_n\overline{T}_{nm}a_n-p_nT^*_{nm}a_n=0$$

$$-p_1a_1\left(a_1X^*_{1m}+b_1X^*_{2m}+\ldots u_1X^*_{mm}\right)-p_2a_2\left(a_2X^*_{1m}+b_2X^*_{2m}+\ldots u_2X^*_{mm}\right)+\ldots$$
$$\ldots-p_na_n\left(a_nX^*_{1m}+b_nX^*_{2m}+\ldots u_nX^*_{mm}\right)+p_1\overline{T}_{1m}a_1+p_2\overline{T}_{2m}a_2+\ldots p_n\overline{T}_{nm}a_n=0$$

Che può essere così scritta in forma sintetica:

$$[p_i a_i a_i]X^*_{1m} + [p_i a_i b_i]X^*_{2m} + \ldots + [p_i a_i u_i]X^*_{mm} = [p_i a_i \overline{T}_{im}]$$

E così per le altre equazioni: si ottiene in definitiva il SISTEMA NORMALE

$$m\begin{cases} [p_i a_i a_i]X^*_{1m} + [p_i a_i b_i]X^*_{2m} + \ldots + [p_i a_i u_i]X^*_{mm} = [p_i a_i \overline{T}_{im}] \\ [p_i b_i a_i]X^*_{1m} + [p_i b_i b_i]X^*_{2m} + \ldots + [p_i b_i u_i]X^*_{mm} = [p_i b_i \overline{T}_{im}] \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ [p_i u_i a_i]X^*_{1m} + [p_i u_i b_i]X^*_{2m} + \ldots + [p_i u_i u_i]X^*_{mm} = [p_i u_i \overline{T}_{im}] \end{cases}$$

Questo sistema fornisce valori più plausibili delle X^*_{im}.

Introducendo i valori sopra ottenuti nel sistema delle equazioni generate, si possono calcolare gli scarti v_i^*

$$a_i X^*_{1m} + b_i X^*_{2m} + \ldots u_i X^*_{mm} - \overline{T}_{im} = v_i^*$$

Che risultano quindi anche come: $v_i^* = \overline{T}_{im}^* - \overline{T}_{im}$

È facile constatare le seguenti tendenze

$[a_i v_i^*] = 0$ $[b_i v_i^*] = 0$ $[u_i v_i^*] = 0$ (utili per il controllo dei calcoli)

In seguito verrà trattato il calcolo delle varianze.

3° Caso: sottocaso non lineare

$$\boxed{f_i\left(X_1, X_2, \ldots X_m / L_i, M_i, \ldots N_i\right) = 0} \qquad i = 1,2,\ldots n \qquad n > m$$

Si trovano i valori approssimati $X_{1m}^0, X_{2m}^0, \ldots X_{mm}^0$ delle incognite e si pone:

$$X_{1m} = X_{1m}^0 + x_i \qquad X_{2m} = X_{2m}^0 + x_2 \ldots \qquad X_{mm} = X_{mm}^0 + x_m$$

$$f_i\left(X_{1m}^0 + x_i, X_{2m}^0 + x_2, \ldots X_{mm}^0 + x_m / L_i, M_i, \ldots N_i\right) = 0$$

si sviluppa in serie:

$$f_i\left(X_{1m}^0, X_{2m}^0, \ldots X_{mm}^0 / \overline{L}_i, \overline{M}_i, \ldots \overline{N}_i\right) + \left(\frac{\partial f_i}{\partial x_i}\right)_0 \cdot x_i + \left(\frac{\partial f_i}{\partial x_2}\right)_0 \cdot x_2 + \ldots + \left(\frac{\partial f_i}{\partial x_m}\right)_0 \cdot x_m = 0$$

da cui

$$a_i x_i + b_i x_2 + \ldots + u_i x_m + T_i = 0$$

ovvero

$$a_i x_i + b_i x_2 + \ldots + u_i x_m + \overline{T}_{im} = v_i$$

essendo

$$\overline{T}_{im} = f_i\left(X_{im}^0, X_{2m}^0, \ldots X_{mm}^0 / \overline{L}_i, \overline{M}_i, \ldots, \overline{N}_i\right)$$

cioè ci siamo così ridotti al caso lineare già esaminato.

Si deve calcolare la varianza del termine noto di ogni equazione

$$\overline{\sigma}^2_{\overline{T}i} = \left(\frac{\partial f_i}{\partial \overline{\overline{L}}_i} \right)^2_m \cdot \overline{\sigma}^2_{\overline{L}i} + \left(\frac{\partial f_i}{\partial \overline{\overline{M}}_i} \right)^2_m \cdot \overline{\sigma}^2_{\overline{M}i} + \ldots + \left(\frac{\partial f_i}{\partial \overline{\overline{N}}_i} \right)^2_m \cdot \overline{\sigma}^2_{\overline{N}i}$$

-Vediamo la trattazione in NOTAZIONE MATRICIALE:

Sia A la matrice dei coefficienti e T il vettore termini noti

$$A = \begin{pmatrix} a_i & b_i & \ldots & u_i \\ a_2 & b_2 & \ldots & u_2 \\ \ldots & \ldots & \ldots & \ldots \\ a_n & b_n & \ldots & u_n \end{pmatrix} \qquad T = \begin{vmatrix} \overline{T}_{1m} \\ \overline{T}_{2m} \\ \vdots \\ \overline{T}_{nm} \end{vmatrix}$$

Si calcola il peso del termine noto generico

$$p_i = \frac{\sigma_0^2}{\sigma_{T_i}^2}$$

Si moltiplica ogni riga della matrice A per il peso che compete al termine noto e si esegue il prodotto di matrici.

$$\begin{pmatrix} a_i & a_2 & ... & a_n \\ b_1 & b_2 & ... & b_n \\ ... & ... & ... & ... \\ u_i & u_2 & ... & u_n \end{pmatrix} \cdot \begin{pmatrix} p_i a_i & p_i b_i & ... & p_1 u_i \\ p_2 a_2 & p_2 b_2 & ... & p_2 u_2 \\ ... & ... & ... & ... \\ p_n a_n & p_n b_n & ... & p_n u_n \end{pmatrix}$$

Questo prodotto permette il calcolo dei coefficienti del sistema normale.
I termini noti invece si ottengono effettuando il prodotto

$$\begin{pmatrix} a_i & a_2 & ... & a_n \\ b_1 & b_2 & ... & b_n \\ ... & ... & ... & ... \\ u_i & u_2 & ... & u_n \end{pmatrix} \cdot \begin{pmatrix} p_1 \overline{T}_{1m} \\ p_2 \overline{T}_{2m} \\ .... \\ p_n \overline{T}_{nm} \end{pmatrix} = T \cdot N$$

Il sistema normale può sintetizzarsi in

$$\boxed{A^T \cdot A \cdot X^* - T \cdot N = 0}$$

da cui

$$X^* = \left(A^T \cdot A\right)^{-1} \cdot TN$$

Calcolo delle varianze delle X_{im}^{*}

Si è visto che:

$$\begin{cases} x_1^* = \alpha_{11}\left[p_i a_i \overline{T}_{im}\right] + \alpha_{12}\left[p_i b_i \overline{T}_{im}\right] + \ldots + \alpha_{1m}\left[p_i u_i \overline{T}_{im}\right] \\ x_2^* = \alpha_{21}\left[p_i a_i \overline{T}_{im}\right] + \alpha_{22}\left[p_i b_i \overline{T}_{im}\right] + \ldots + \alpha_{2m}\left[p_i u_i \overline{T}_{im}\right] \\ \boxed{x_k^* = \alpha_{k1}\left[p_i a_i \overline{T}_{im}\right] + \alpha_{k2}\left[p_i b_i \overline{T}_{im}\right] + \ldots + \alpha_{km}\left[p_i u_i \overline{T}_{im}\right]} \\ \\ x_m^* = \alpha_{m1}\left[p_i a_i \overline{T}_{im}\right] + \alpha_{m2}\left[p_i b_i \overline{T}_{im}\right] + \ldots + \alpha_{mm}\left[p_i u_i \overline{T}_{im}\right] \end{cases}$$

sviluppiamo la X_k generica:

$$x_k^* = \alpha_{k1}\left(p_1 a_i \overline{T}_{1m} + p_2 a_2 \overline{T}_{2m} + \ldots + p_n a_n \overline{T}_{nm}\right) + \alpha_{k2}\left(p_1 b_1 \overline{T}_{1m} + p_2 b_2 \overline{T}_{2m} + \ldots + p_n b_n \overline{T}_{nm}\right) + \ldots$$

$$\ldots + \alpha_{km}\left(p_1 b_1 \overline{T}_{1m} + p_2 b_2 \overline{T}_{2m} + \ldots + p_n b_n \overline{T}_{nm}\right) =$$

$$= \overline{T}_{1m}\, p_1\left(\alpha_{k1} a_1 + \alpha_{k2} b_1 + \ldots + \alpha_{km} u_1\right) + \overline{T}_{2m}\, p_2\left(\alpha_{k1} a_2 + \alpha_{k2} b_2 + \ldots + \alpha_{km} u_2\right) + \ldots$$

$$\ldots + \overline{T}_{nm}\, p_1\left(\alpha_{k1} a_n + \alpha_{k2} b_n + \ldots + \alpha_{km} u_n\right)$$

e quindi con le solite leggi:

$$\sigma^2_{X_k} = \sigma^2_{T_{1m}} p_1^2 \left(\alpha_{k1} a_1 + \alpha_{k2} b_1 + \ldots + \alpha_{km} u_1\right)^2 + \ldots$$

$$\ldots + \sigma^2_{T_{2m}} p_2^2 \left(\alpha_{k1} a_2 + \alpha_{k2} b_2 + \ldots + \alpha_{km} u_2\right)^2 + \ldots$$

$$\ldots + \sigma^2_{T_{nm}} p_n^2 \left(\alpha_{k1} a_n + \alpha_{k2} b_n + \ldots + \alpha_{km} u_n\right)^2 =$$

ed essendo $\sigma_0^2 = \sigma^2_{T_{im}} p_i^2$

$$= \sigma_0^2 \; p_1 \left(\alpha_{k1} a_1 + \alpha_{k2} b_1 + \ldots + \alpha_{km} u_1\right)^2 + \ldots$$

$$\ldots + \sigma_0^2 \; p_1 \left(\alpha_{k1} a_2 + \alpha_{k2} b_2 + \ldots + \alpha_{km} u_2\right)^2 + \ldots$$

$$\ldots + \sigma_0^2 \; p_n \left(\alpha_{k1} a_n + \alpha_{k2} b_n + \ldots + \alpha_{km} u_n\right)^2 + \ldots$$

$$\sigma^2_{X_k} = \sigma_0^2 \left\lfloor p_1 \left(\alpha_{k1} a_i + \alpha_{k2} b_i + \ldots + \alpha_{km} u_i\right)^2 \right\rfloor$$

Ma si dimostra che $\left\lfloor p_1 \left(\alpha_{k1} a_i + \alpha_{k2} b_i + \ldots + \alpha_{km} u_i\right)^2 \right\rfloor = \alpha_{kk}$

e quindi

$$\sigma^2_{X_k} = \alpha_{kk} \cdot \overline{\sigma}_0^2$$

essendo

$$\overline{\sigma}_0^2 = \frac{\sum_{i=1}^{n} p_i \cdot \overline{v}_i^2}{n - m}$$

(calcolo a posteriori della varianza dell’unità di peso)

OSSERVAZIONI INDIRETTE

Scrivere SISTEMA EQUAZIONI GENERATE LINEARIZZATO

Scrivere SOLUZIONI SISTEMA NORMALE

Esplicitare la 1°mettendo in evidenza i termini noti

Considerare la formazione di una nuova variabile casuale e calcolarne la σ^2

farne lo sviluppo nel caso di tre sole incognite

Si ottiene $\sigma_{X_k}^2 = \alpha_{kk} \cdot \overline{\sigma}_0^2$

VALUTAZIONE di σ_0

SISTEMA delle EQUAZIONI GENERATE

Considerare le medie teoriche sia delle incognite che delle grandezze misurate

scrivere sistema equazioni generate linearizzate contenente le correzioni per avere medie teoriche

Sottrarre questo sistema dal sistema equazione generato di tipo tradizionale

Considerare il caso di
n=3 equazioni
m=2 incognite

Elevare al quadrato e sommare

Considerare la (li=lf) una nuova variabile casuale

Calcolarne il valore MEDIO

$$n\sigma_0^2 = M\{v_i^2\} + m\sigma_0^2$$

in generale

e quindi

$$\sigma_0^2 = \frac{M\{v_i^2\}}{n-m}$$

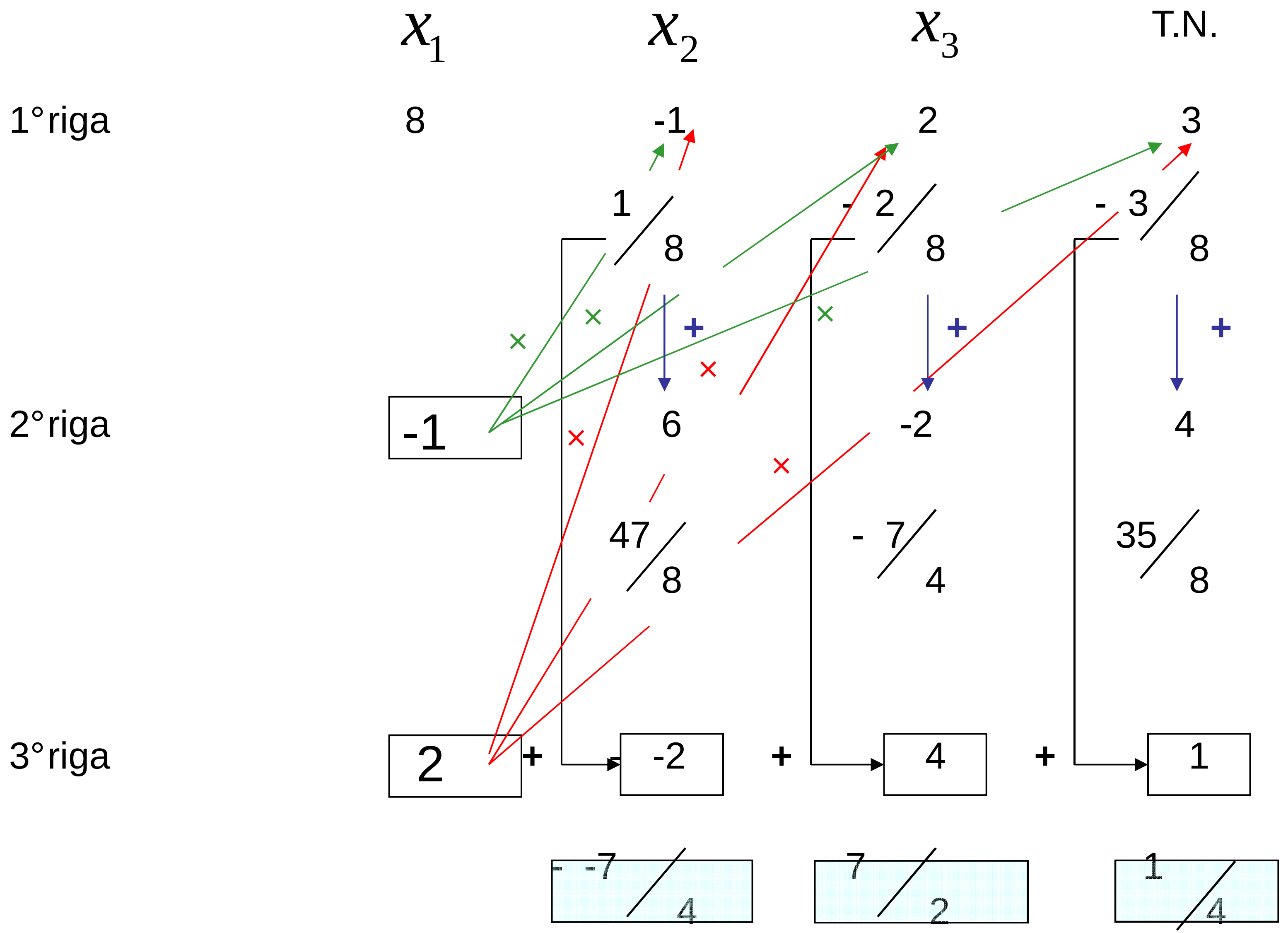
e per un valore stimato si pone

$$\overline{\sigma_0}$$

RISOLUZIONE di UN PICCOLO SISTEMA CON IL METODO REITERATIVO DELLE ELIMINAZIONI SUCCESSIVE

Si abbia ad esempio il seguente sistema di 3 equazioni in 3 incognite:

$$\begin{cases} 8x_1 - x_2 + 2x_3 + 3 = 0 \\ -x_1 + 6x_2 - 2x_3 + 4 = 0 \\ 2x_1 - 2x_2 + 4x_3 + 1 = 0 \end{cases}$$



	x_1	x_2	x_3	T.N.
1°riga	8	-1	2	3
		$1/8$	$-2/8$	$-3/8$
2°riga	-1	6	-2	4
		$47/8$	$-7/4$	$35/8$
3°riga	2	-2	4	1
		- $-7/4$	$7/2$	$1/4$

1. Si dividono i coefficienti di x_2 , x_3 e T.N. per il coefficiente di x_1 e si cambia di segno (nella 1° equazione)

2. Si moltiplica il coefficiente di x_1 della 2° equazione per il coefficiente di x_2 modificato (1° equaz.) e si somma al coeff. di x_2 della 2° equaz.
   Si moltiplica ancora il coeff. di x_1 della 2° equazione per il coefficiente di x_3 modificato (1° equaz.) e si somma al coeff. di x_3 della 2° equaz e così di seguito.

Quindi si moltiplica il coefficiente di x_1 della 3° equazione per il coefficiente di x_2 modificato (1° equazione) e si somma al coefficiente di x_2 della 3° equazione e così di seguito.

3. Si ottiene il sistema ridotto:

$$\begin{cases} \frac{47}{8} x_2 - \frac{7}{4} x_3 + \frac{35}{8} = 0 \\ -\frac{7}{4} x_2 + \frac{7}{2} x_3 + \frac{1}{4} = 0 \end{cases}$$

che può essere facilmente risolto.

SISTEMA NORMALE

RISOLUZIONE

- LINEARE
  - Algoritmi esatti:
    - A. di Gauss
    - A. di CHOLESKY
    - A. di HOUSEHOLDER,ecc
  - Algoritmi iterativi
    - Metodo del GRADIENTE CONIUGATO (CG)
    - Metodo di CHOLESKY incompleto gradiente coniugato (IGCG)
- NON LINEARE
  procedimenti operativi per la ricerca di valori appross. delle incognite