



# Lezione I

Pier Luca Maffettone [pierluca.maffettone@unina.it](mailto:pierluca.maffettone@unina.it)  
<http://wpage.unina.it/p.maffettone/>

Gaetano D'Avino [gadavino@unina.it](mailto:gadavino@unina.it)

**Tipo di Esame:** Scritto

**Attività durante il Corso:** 2 PI facoltative (11/11 o 13/11, 16/12 o 18/12), Esercizi da svolgere in gruppo facoltativi

**Strumenti computazionali:** Matlab

**Criterio di valutazione:** Media due PI + 2 punti con tutti gli esercizi svolti **oppure** voto finale

## Testi

Douglas C. Montgomery e George C. Runger, *Applied Statistics and Probability for Engineers*, 2003, John Wiley & Sons, Inc.

Introduzione alla statistica: E. Kreyszig, *Introductory mathematical statistics*, Wiley 1970

Identificazione parametrica:

Y. Bard, *Nonlinear parameter estimation*, Academic Press, 1974

E. Walter and L. Pronzato, *Identification of Parametric Models*, Springer Masson, 1997

P. R. Bevington e D. K. Robinson, *Data reduction and error analysis for the physical sciences*, Mc Graw Hill, 1992.

# Sommario

- In questa lezione introdurremo i concetti di base della statistica e la relativa terminologia
- Questi concetti includono:
  - L'incertezza che caratterizza i dati sperimentali
  - Il ruolo della statistica nella sperimentazione ingegneristica e scientifica
  - La distinzione tra campione e popolazione
  - Il modello del processo
  - Il modello dell'esperimento

# Motivazione

- I dati sono l'aspetto centrale degli studi sperimentali
  - Dati: informazioni espresse in forma quantitativa (tralasciamo in questo corso dati di tipo qualitativo)
- Tutti i dati ottenuti sperimentalmente sono soggetti ad errori che inducono **incertezze** nelle misurazioni
  
- Che cos'è la statistica?
- La statistica è una disciplina scientifica che permette di ottenere deduzioni logiche valide a partire da dati sperimentali incerti.
  - Quotidianamente usiamo la statistica (anche se forse non ci facciamo troppo caso)
    - Previsioni del tempo
    - Sondaggi di opinione
    - Previsioni di crescita dell'economia italiana
    - Etc.

**La Statistica è la scienza che risolve problemi in presenza di incertezza**

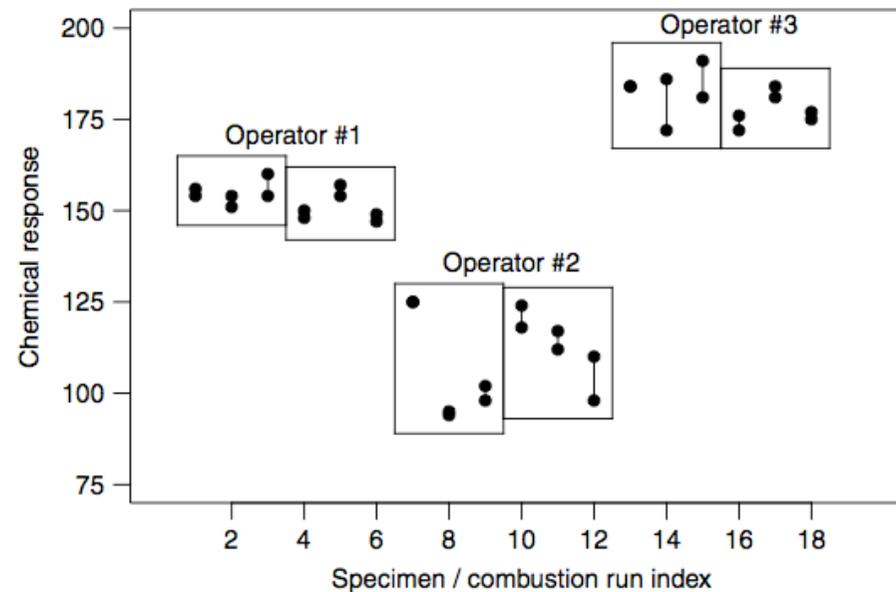
# Motivazione

- Esempio: Analisi chimica

TABLE 1.1 Results of an Experiment to Identify Sources of Variation in Chemical Analyses<sup>a</sup>

Operator	Specimen	Combustion Run	Chemical Analysis	
			1	2
1	1	1	156	154
		2	151	154
		3	154	160
	2	4	148	150
		5	154	157
		6	147	149
2	3	7	125	125
		8	94	95
		9	98	102
	4	10	118	124
		11	112	117
		12	98	110
3	5	13	184	184
		14	172	186
		15	181	191
	6	16	172	176
		17	181	184
		18	175	177

<sup>a</sup> Adapted from Snee, R. D. (1983). "Graphical Analysis of Process Variation Studies," *Journal of Quality Technology*, 15, 76-88. Copyright, American Society for Quality Control, Inc., Milwaukee, WI. Reprinted by permission.



i risultati sperimentali indicano che le fonti primarie di variazione in queste analisi chimiche sono le differenze sistematiche (bias) tra gli operatori e, in alcuni casi, la (casuale) incoerenza delle analisi chimiche eseguite da un singolo operatore.

# Motivazione

- Questo esempio illustra **tre caratteristiche generali** del modello statistico e di analisi degli esperimenti.
  - **Le considerazioni statistiche dovrebbero essere incluse nella fase di progettazione di ogni esperimento.**
    - la natura dei dati da raccogliere,
    - le misure da fare
    - le incertezze connesse alle misure
    - i fattori che possono influenzare le grandezze misurate
  - **un progetto statistico della sperimentazione**
    - Deve consentire la caratterizzazione delle incertezze che andranno modellate
      - Errore sperimentale è introdotto quando le condizioni di prova non sono identiche:
      - Errori di misura derivano dall'incapacità di ottenere esattamente la stessa misura su due prove successive, quando tutte le condizioni sperimentali sono immutate.
  - **un'analisi statistica dei risultati sperimentali che consente di trarre deduzioni sulle relazioni tra le condizioni sperimentali e le grandezze misurate.**
    - I risultati di questa fase potrebbero suggerire una sperimentazione supplementare. L'analisi dovrebbe guidare lo sperimentatore ad una progettazione adeguata e, se necessario, un modello più adeguato del processo di misurazione.

# Motivazione

## Fase di progettazione

- Cosa intendiamo misurare?
- Quanto ampie saranno le variazioni?
- Quali sono le grandezze che influenzano il fenomeno?

## Fase di progettazione dell'esperimento

- Controllare le sorgenti di variabilità note
- Permettere la stima delle dimensioni delle variazioni incontrollate
- Permettere lo studio di modelli adeguati

## Fase di analisi statistica

- Inferenza
- Guida a progettazione seguente
- Suggestire modelli più adeguati

# Motivazione

- I dati vengono raccolti con un certo obiettivo
- L'obiettivo determina sia il tipo di dati da raccogliere che gli strumenti statistici da utilizzare per la loro analisi.

## INDAGINE STATISTICA

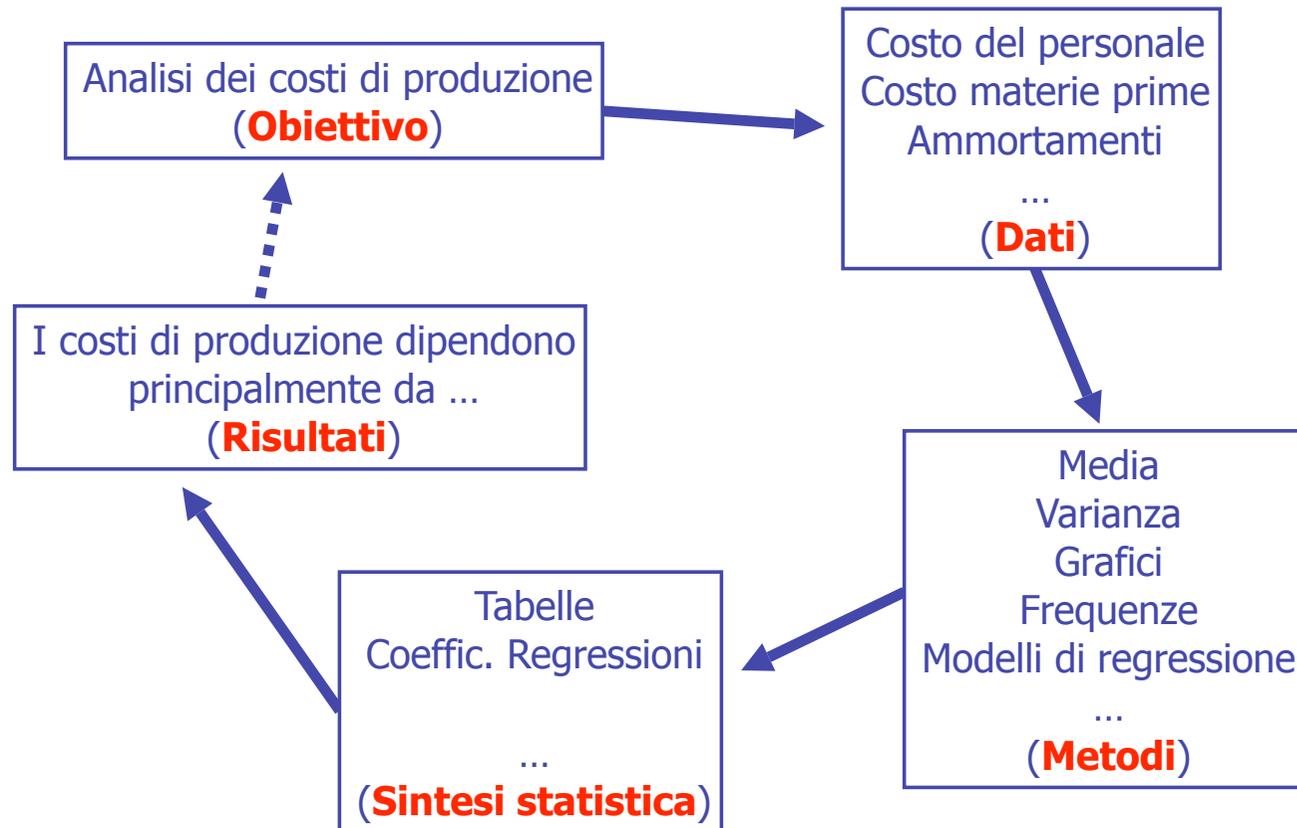


# Motivazione

- Si effettuano esperimenti
  - per scoprire qualche informazione su un particolare processo
  - per determinare l'effetto di diverse condizioni su di un fenomeno.
  - per stimare parametri di modelli matematici
- Quali sono le grandezze che influenzano il processo?
- Obiettivo: **Stima parametri da campagne sperimentali**
- Problemi connesso all'obiettivo:
  - Progettazione di campagne sperimentali adeguate
  - Quante prove sperimentali sono necessarie?
  - In quali condizioni devono essere effettuate?

# Motivazione

## ESEMPIO I



# Motivazione

## ESEMPIO II

- Costruiamo una campagna sperimentale sulla produttività di un reattore per:
  - Decidere se la produttività media in fissate condizioni operative raggiunge le condizioni di ottimo economico;
  - Determinare se una certa scelta di condizioni operative determina una produttività più alta di altre;
  - Trovare un modello matematico che permetta di prevedere la produttività al variare di condizioni operative.

# Motivazione

## ESEMPIO III

- Vogliamo analizzare come varia la tensione di vapore della CO con la temperatura.
- Crediamo che tale dipendenza si possa descrivere con la legge di Antoine

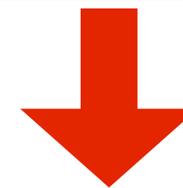
$$P = 10^{A - \frac{B}{C+T}}$$

- Viene condotta una campagna sperimentale

Original Data	
p (torr)	T (C)
1	-222,0
5	-217,2
10	-215,0
20	-212,8
40	-210,0
60	-208,1
100	-205,7
200	-201,3
400	-196,3
760	-191,3
T (C)	p (atm)
-191,3	1
-183,5	2
-170,7	5
-161,0	10
-149,7	20
-141,9	30

Modified units	
p (Pa)	T (K)
133	51,2
667	56,0
1333	58,2
2666	60,4
5333	63,2
7999	65,1
13332	67,5
26664	71,9
53329	76,9
101325	81,9
202650	89,7
506625	102,5
1013250	112,2
2026500	123,5
3039750	131,3

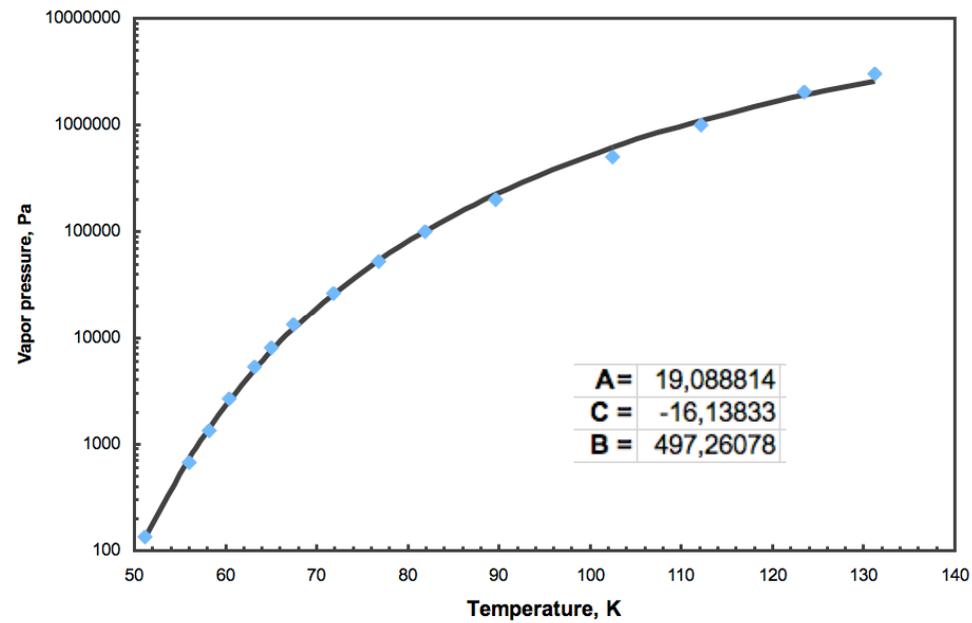
possiamo ricavare i valori dei  
parametri del modello  
operando sui dati  
sperimentali



REGRESSIONE

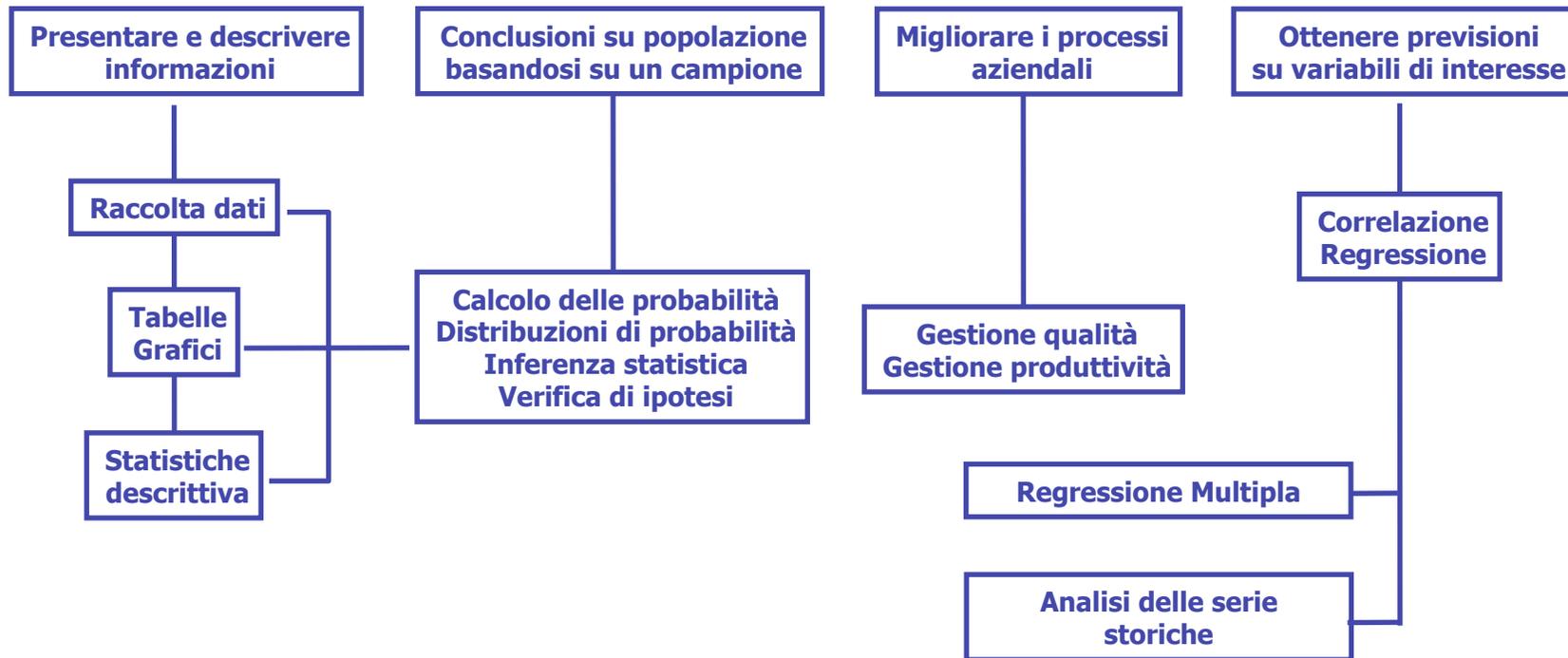
# Motivazione

- Legge di Antoine  $P = 10^{A - \frac{B}{C+T}}$



# Motivazione

## La statistica in ambito industriale



# Motivazione

- Durante il corso verranno presentati:
  - Metodi statistici
  - Applicazione di metodi statistici
- L'uso di pacchetti statistici o di fogli elettronici elimina il problema del calcolo
  - **ATTENZIONE!**: Non mette al riparo l'utilizzatore da interpretazioni o conclusioni che possono essere fuorvianti qualora egli non sia consapevole delle ipotesi di base e dei limiti degli strumenti utilizzati
- Nel campo dell'ingegneria chimica è abbastanza comune modellare matematicamente processi di interesse.
- **La descrizione matematica di processi implica l'introduzione di parametri.**
  - **Vantaggio**: Se il modello descrive adeguatamente il processo può essere utilizzato per analizzare condizioni diverse con semplici operazioni
  - **Problema da risolvere**: Come si può procedere per la determinazione del valore dei parametri?

# Software statistici

- Software commerciali
  - SAS (<http://www.sas.com/>)
  - SPSS (<http://www.spss.com/>)
  - SYSTAT (<http://www.systat.com/>)
  - Minitab (<http://www.minitab.com/>)
  - ...
- Funzioni statistiche incluse in software matematici (e.g. MATLAB, Mathematica,...)
- GAMS (Guide to Available Mathematical Software)
  - <http://gams.nist.gov/>

# Variabilità delle grandezze misurate

- Lavorate al processo di produzione di un polimero.
- Il polimero viene raccolto a valle del reattore di polimerizzazione.
- **Dovete controllare la qualità del polimero prodotto, ovvero che il suo peso molecolare rispetti le specifiche richieste dall'acquirente.**
- Il controllo di qualità viene eseguito misurando la viscosità del prodotto che esce dal reattore.
- Il prodotto è soddisfacente se la viscosità misurata in certe condizioni ha un certo valore.

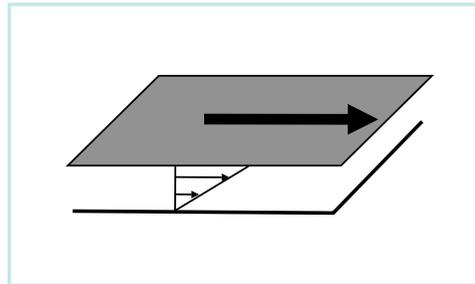
Esiste un modello  
che collega  
viscosità a peso  
molecolare

# Variabilità delle grandezze misurate

- **STRUMENTAZIONE**

- Viscosimetro da campo a banda

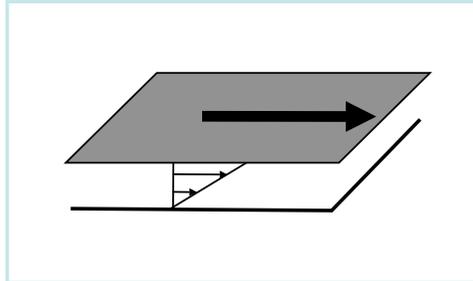
- **Principio di funzionamento**



- Il flusso è uno scorrimento: è possibile individuare una famiglia di superfici materiali (cioè fatte di particelle di fluido) che si mantengono inalterate durante il moto.
  - **Le superfici sono in moto relativo l'una rispetto all'altra.** Il moto di ciascuna di esse è però rigido (traslatorio e/o rotatorio).
  - Per effetto dello scorrimento si genera nel liquido uno sforzo tangenziale.

# Variabilità delle grandezze misurate

## Principio di funzionamento



- Se il liquido è Newtoniano allora:  $\tau = \eta \gamma$ 
  - Se conoscete lo sforzo tangenziale e la velocità di deformazione potete ricavare la viscosità del liquido.
- Serve un modello per il viscosimetro.
- Obiettivo: dare un valore plausibile al parametro viscosità  $\eta$

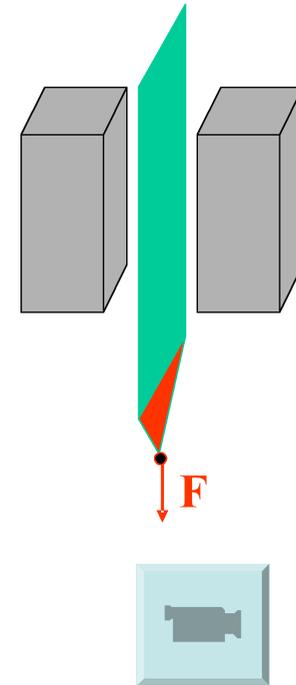
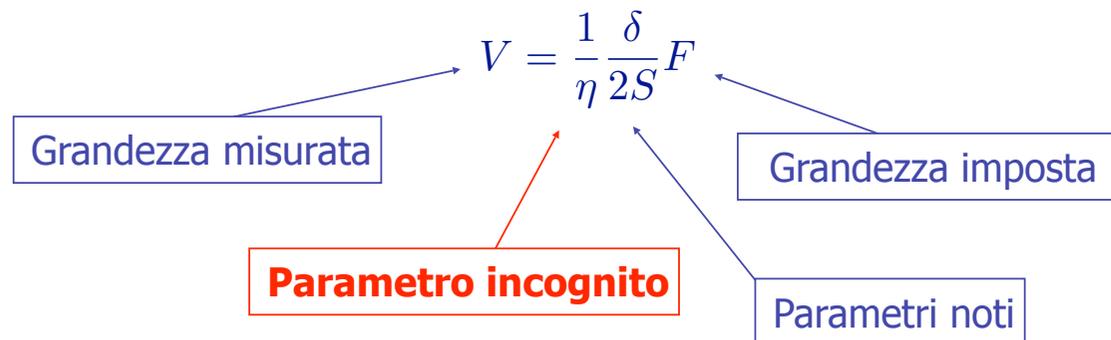
**MODELLO**

# Variabilità delle grandezze misurate

## Viscosimetro a banda

- Si applica un peso ad un nastro che viene fatto scorrere all'interno del fuso polimerico tra due pareti.
- Si misura la velocità con cui il nastro scorre.

**MODELLO**  
(si basa su ipotesi)



# Variabilità delle grandezze misurate

- L' esperimento si basa sulla misura simultanea di due grandezze:
  - Una cinematica:  $V$  la velocità con cui si muove il piatto
  - Una dinamica:  $F$  la forza da applicare per tenere in moto il nastro alla velocità  $V$

- Possiamo fare un ulteriore passo  $V = \frac{L}{\Delta t}$

$$\Delta t = \eta \frac{2LS}{\delta} \frac{1}{F}$$

# Variabilità delle grandezze misurate

- Facciamo alcune ipotesi sulle quali poi ritorneremo:
  - Conosciamo perfettamente la geometria dell'apparato, ovvero conosciamo perfettamente  $S$ ,  $L$  e  $\delta$ .
  - Applichiamo una forza nota perfettamente (un peso calibrato) e misuriamo la velocità con cui si muove il nastro.
  - $F$  quindi rappresenta la condizione sperimentale scelta.
  - $\Delta t$  è la variabile misurata. **Nota questa dall'equazione del viscosimetro ricaveremo la viscosità**

$$\eta = \frac{\delta}{2LS} F \Delta t = K \Delta t$$

- Proviamo a simulare l'esperimento. Misurate il tempo 
- Se consideriamo i risultati che avete ottenuto ci accorgiamo che la grandezza misurata non è univoca, anzi, sembra essere **distribuita**.

# Variabilità delle grandezze misurate

- Sembra non esistere un valore per la viscosità ma tanti valori.
- Presumiamo di avere fatto 50 misure, avremo 50 valori diversi della viscosità.
- Come possiamo razionalizzare questo risultato?
- **Errori di misura**

**Errori di  
misura**

# Variabilità delle grandezze misurate

- Analizziamo i risultati che abbiamo ottenuto.

- 50 numeri reali (il cronometro digitale riportava i centesimi di secondo

(3.91, 4.29, 4.01, 3.96, 4.45, 4.18, 4.26, 4.39, 3.65, 3.79, 4.38, 4.07, 3.56, 4.04, 3.38, 4.07, 3.98, 4.3, 3.74, 4.08, 4.08, 3.66, 4.26, 3.71, 3.99, 4.05, 4.11, 3.85, 4.08, 3.76, 3.99, 3.95, 4.45, 4.04, 4.17, 4.01, 4.11, 3.71, 4.52, 4.06, 4.04, 3.71, 4.55, 4.41, 4.08, 3.99, 4.47, 4.14, 4.00, 3.81)

- Possiamo stabilire quale sia il minimo ed il massimo dei dati sperimentali

Min=3.38 Max=4.55

- Quindi consideriamo l'intervallo tra minimo e massimo e lo suddividiamo in sottointervalli (p.e. di passo 0.125)

- Quindi contiamo quanti dati entrano in ciascun sottointervallo

(1, 1, 6, 4, 8, 15, 3, 4, 6, 2)

- Quanto vale la somma?

- Infine costruiamo un istogramma.

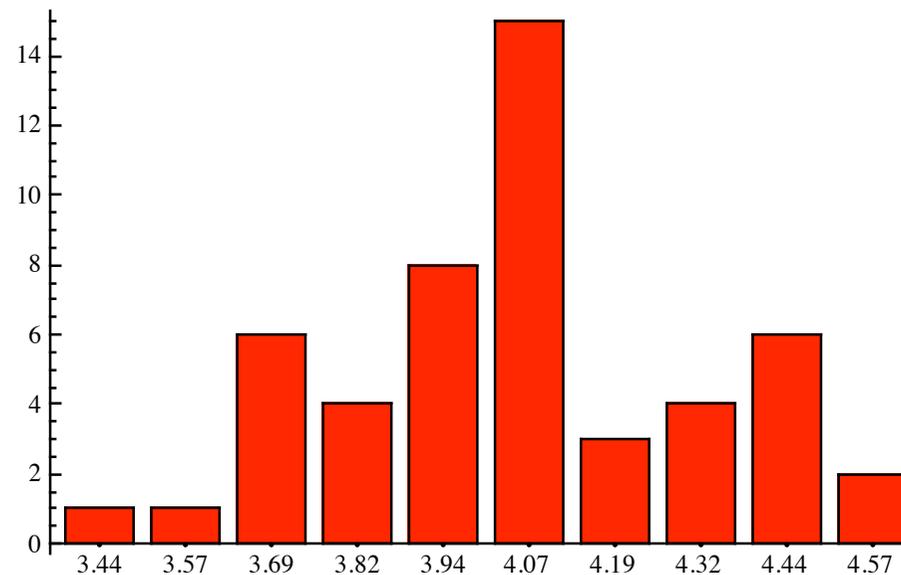
Variabilità

Range

Istogramma  
o bar chart

# Variabilità delle grandezze misurate

- **Istogramma:** utile a catturare facilmente informazioni su una **distribuzione**
  - Si divide l'intervallo di variabilità dei dati raccolti in un certo numero di sottointervalli e si contano quanti dati “cadono” in ciascun sottointervallo (frequenza). Questi conteggi si possono rappresentare sia attraverso una tabella che in un grafico.

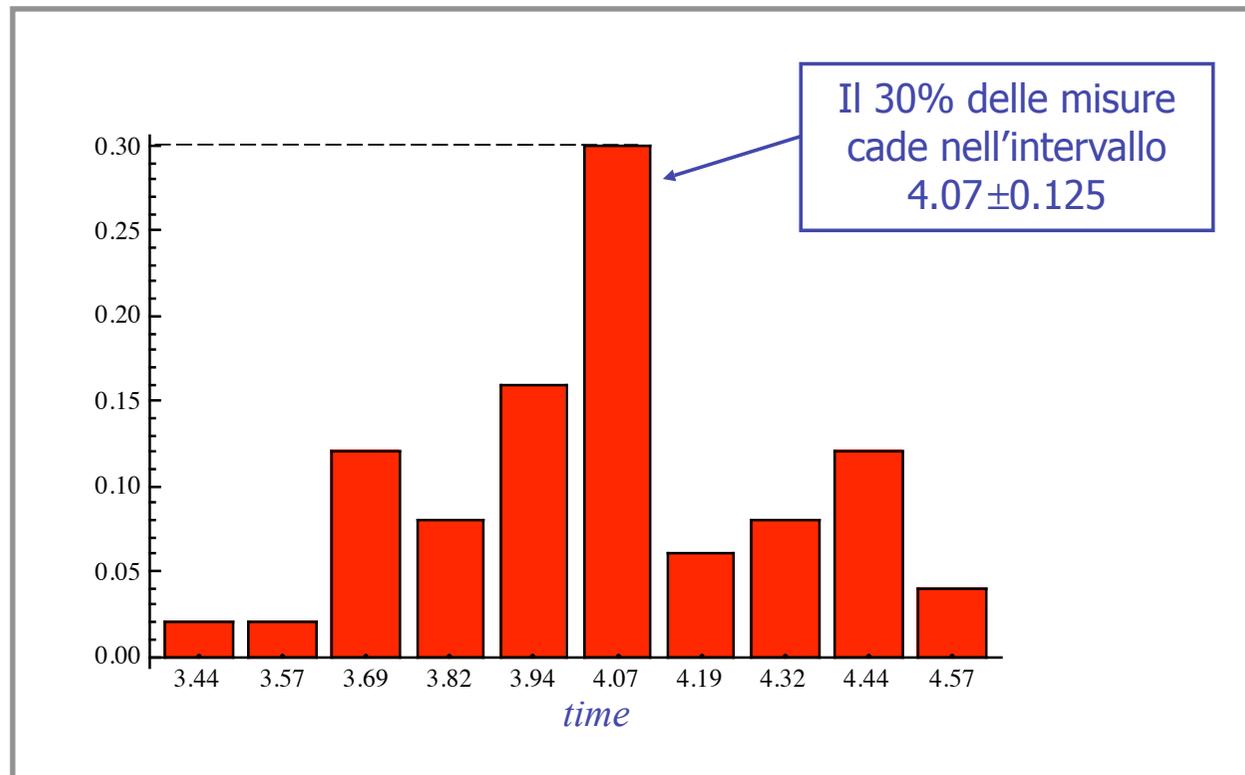


- Cosa possiamo dire?

# Variabilità delle grandezze misurate

- Ovviamente possiamo usare le **frequenze relative** dividendo le frequenze assolute per il numero totale di dati (50 in questo caso)

Frequenze  
relative



# Struttura dei dati

- Non abbiamo ottenuto un unico valore per la viscosità ma notiamo che esiste una **struttura nei dati**.
  - Secondo voi come mai esiste questa struttura?
- Avete qualche proposta su come utilizzare questi risultati?
- Riconoscete una tendenza centrale nell'istogramma?
  - Avete qualche suggerimento su come misurarla?
- Riconoscete una dispersione attorno alla tendenza centrale?
  - Avete qualche suggerimento su come misurarla?

# Errore nella misura sperimentale

- **Errore di misura:** differenza tra il risultato di una misura e il valore vero del misurando (in maniera intuitiva il valore vero è il risultato di una misura perfetta)
  - Essendo il valore vero generalmente ignoto, ignoto è anche l'errore (a parte i casi in cui si disponga di un campione di riferimento assunto come convenzionalmente vero)
  - Gli errori di misura sono inevitabili e anche inconoscibili, quindi anche il valore vero è non conoscibile, per questo si introduce il concetto di incertezza:
- **Incertezza:** è un parametro che caratterizza la dispersione dei valori che possono essere ragionevolmente attribuiti al misurando
- Nel seguito si parlerà di errore come sinonimo di incertezza nel senso appena espresso

# Incertezze ed analisi degli errori

- Un modo corretto di quantificare il risultato di una misura consiste nel dare la migliore stima della grandezza misurata ed un intervallo entro il quale è compresa con una certa fiducia tale grandezza.
  - Esempi con misure analogiche: righello, voltmetro etc.
  - Esempi con misure ripetute

$$(\text{valore misurato } x) = x_{best} \pm \delta x$$

- Il numero  $\delta x$  è detto incertezza o errore della misura.

- **Attenzione alle cifre significative:**

- Regola per l'incertezza: deve essere arrotondata ad una cifra significativa
- Regola per la stima migliore: L'ultima cifra significativa deve essere dello stesso ordine di grandezza dell'incertezza.

~~$$g = 9.82 \pm 0.02385 \frac{m}{s^2}$$~~

$$g = 9.82 \pm 0.02 \frac{m}{s^2}$$

~~$$v = 6051.78 \pm 30 \frac{m}{s}$$~~

$$v = 6050 \pm 30 \frac{m}{s}$$

# Errore nella misura sperimentale

## Cause degli errori di misura

– lo sperimentatore è parte integrante del processo di misura

- Incompleta definizione del misurando
- Incompleta realizzazione della definizione del misurando
- Imperfetta conoscenza delle condizioni ambientali e dei loro effetti sulla misura
- Errore di lettura di uno strumento
- Risoluzione finita dello strumento (es. Strumenti digitali)
- Valori inesatti dei campioni di riferimento

# Errore nella misura sperimentale

- Ogni misura sperimentale è influenzata da errori.
- Esperimento replicabile
  - Un esperimento si dice replicabile se è possibile ripeterlo più volte nelle stesse identiche condizioni.
- **In esperimenti replicabili le misure ottenute non sono generalmente uguali per effetto dell'errore nella misura.**

**Esperimento  
replicabile**

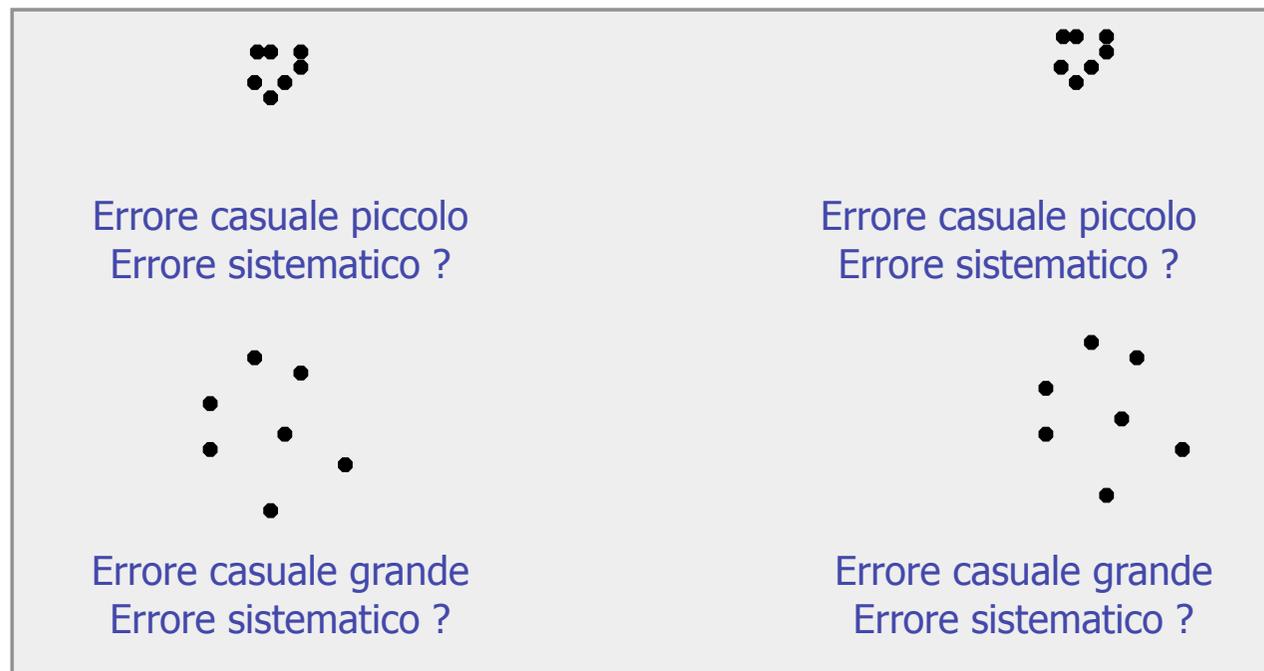
# Errori sistematici ed errori casuali

- Esempio: tiro al bersaglio. Gli esiti dei tiri sono influenzati da eventi casuali (mano poco ferma del tiratore, vento etc.) oppure dalla sbagliata calibrazione del mirino (errore quindi sistematico, perché porta una deviazione del tiro sempre nello stesso verso)



# Errori sistematici ed errori casuali

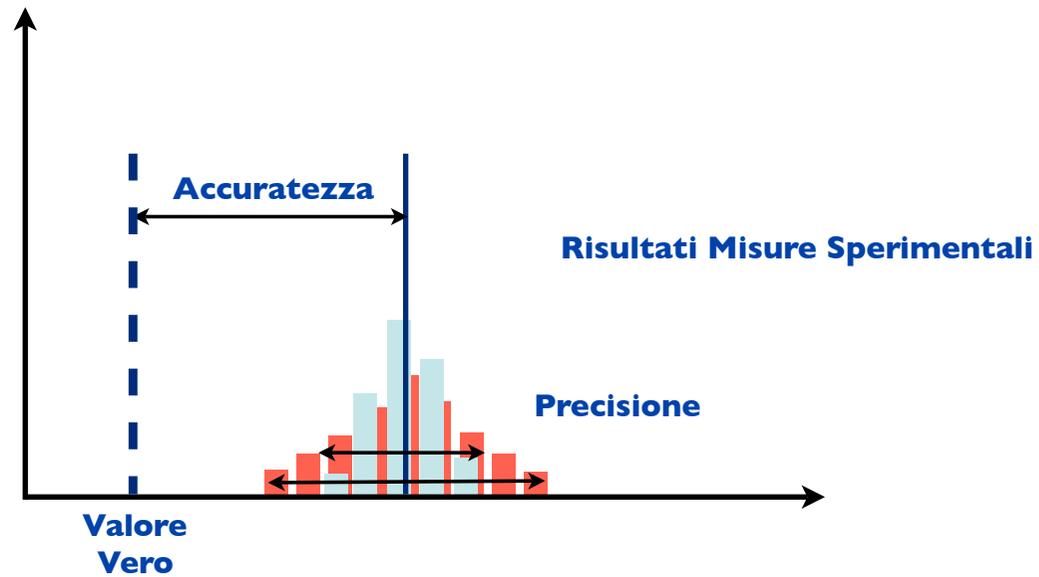
- Se rimuoviamo il segno del bersaglio abbiamo qualcosa di simile a quanto abbiamo fatto con il viscosimetro
  - Non conoscevamo a priori il valore della viscosità del polimero.
- E' possibile individuare gli errori casuali, ma non gli errori sistematici



# Accuratezza e Precisione

- **Accuratezza:** Una misura è tanto più accurata quanto più la media delle misure si approssima al valore vero della grandezza di un esperimento.
- **Precisione:** Una misura è tanto più precisa quanto più i singoli valori misurati in condizioni di ripetibilità si concentrano intorno alla media della serie di misure effettuate
  - E' una misura della riproducibilità del risultato
- L'accuratezza di un esperimento dipende da quanto bene possiamo controllare o compensare gli errori sistematici.
  - D'ora in avanti non ci occuperemo più di errori sistematici
- La precisione di un esperimento dipende da quanto bene riduciamo gli errori casuali.

# Accuratezza e Precisione



# Formalizzazione

- Torniamo al problema della misura di viscosità.
- Può convenire riconsiderare il problema in modo un po' più astratto.
- Individuiamo la grandezza (nell'esempio considerato) che possiamo variare a nostro piacimento (l'apparecchiatura la consideriamo fissata).
  - Chiamiamo questa grandezza **Variabile Indipendente** o **Condizione Sperimentale**:  $X$
- Individuiamo la grandezza misurata.
  - Chiamiamo questa grandezza **Variabile dipendente o misurata**:  $Y$
- Individuiamo il parametro incognito e chiamiamolo  $p$
- Il modello che abbiamo costruito diventa quindi:  $Y = p X$

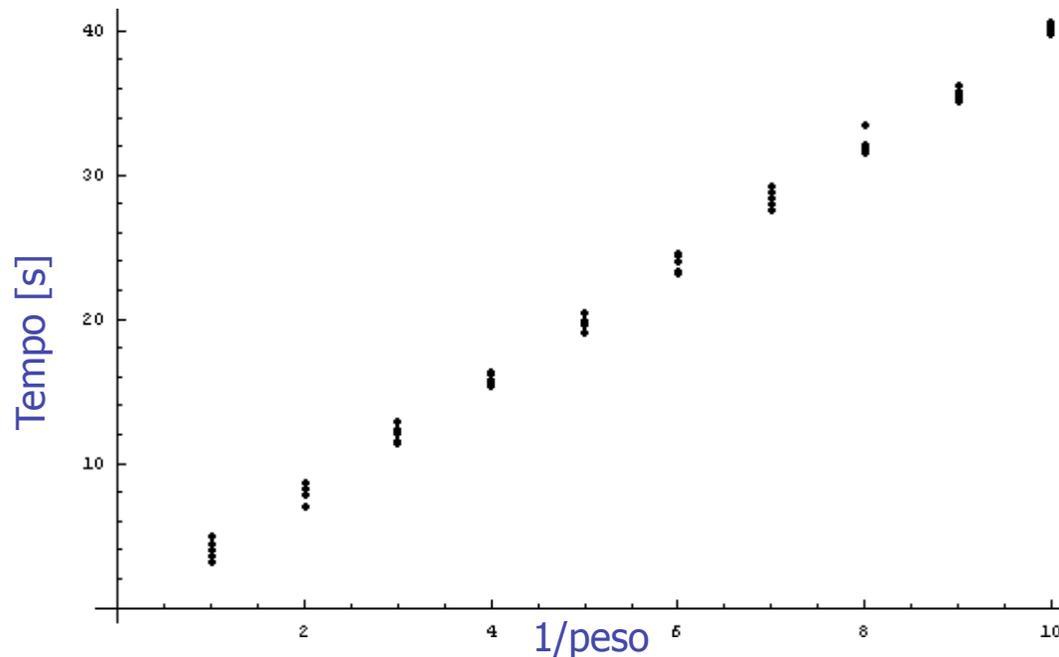
**Modello del  
processo**

# Dipendenza dalla condizione sperimentale

- Abbiamo fatto varie ipotesi per arrivare alla formalizzazione dell'equazione del viscosimetro.
- La più importante forse è quella che il liquido sia Newtoniano. Sulla base delle informazioni che abbiamo a disposizione possiamo dire qualcosa su questa ipotesi?
- Cosa succede se facciamo una nuova campagna sperimentale raddoppiando il peso applicato?
- Se il liquido è effettivamente Newtoniano cosa vi aspettate accada?

# Dipendenza dalla condizione sperimentale

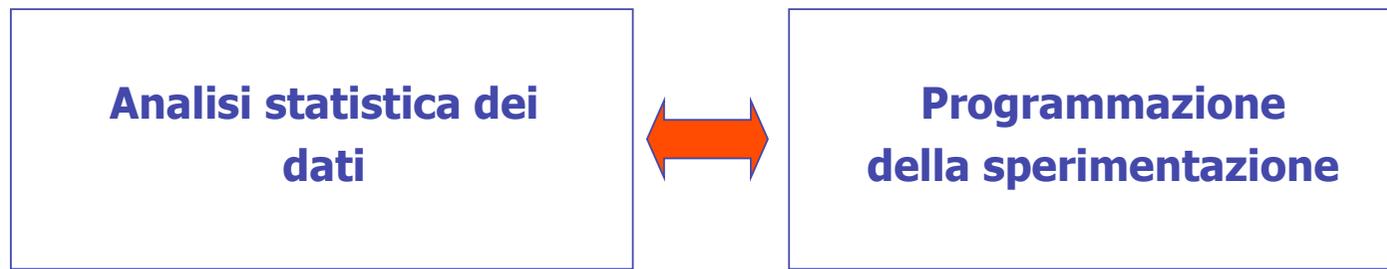
- Risultati con dieci diversi pesi e repliche



- Come analizzereste questi dati? Sono coerenti con quanto abbiamo fatto fin'ora?

# Formalizzazione del problema

- Si definisce progetto statistico degli esperimenti il processo che permette di raccogliere dati appropriati che possono essere analizzati con metodi statistici in modo da ottenere conclusioni obiettive e valide



- La ragione dell'analisi statistica sta nel fatto che tutti i dati raccolti sperimentalmente sono, in qualche modo, affetti da **errori casuali**.

## OBIETTIVI

- Stimare parametri
- Verificare una ipotesi
- Rappresentare una situazione fisica funzionalmente

# Formalizzazione del problema

- La **stima dei parametri** è il procedimento con cui dal campione osservato si traggono informazioni per assegnare un valore ad un parametro.
  - Nell'esempio che abbiamo considerato la stima era finalizzata a conoscere il valore della viscosità del polimero prodotto.
  - Si definisce stima puntuale la procedura che usa dati del campione per calcolare un singolo valore ottimale che stima un parametro della popolazione
    - La stima puntuale non fornisce informazioni sull'errore coinvolto nell'operazione di stima.
  - Una stima per intervallo fornisce un intervallo di valori che presumibilmente include il valore del parametro con un fissato grado di incertezza
    - Fornisce una misura dell'errore di stima.
    - Più largo è l'intervallo peggiore è la stima puntuale
- La **verifica delle ipotesi** è un procedimento che consiste nel fare una congettura sul valore di un parametro e nel decidere sulla base di un campione se essa sia accettabile.
  - Nell'esempio che abbiamo considerato il problema era il controllo di qualità. Vi era stato detto che se la viscosità era 10Poise il prodotto era buono. Dalla stima che avete effettuato (non sappiamo ancora bene come) ottenete un valore di 9.7Poise. Il polimero prodotto è OK o no?

# Introduzione

- ESPERIMENTO

- Riproducibile

- Protocollo

Deterministico

Aleatorio (non predicibile)

Esperimento  
aleatorio

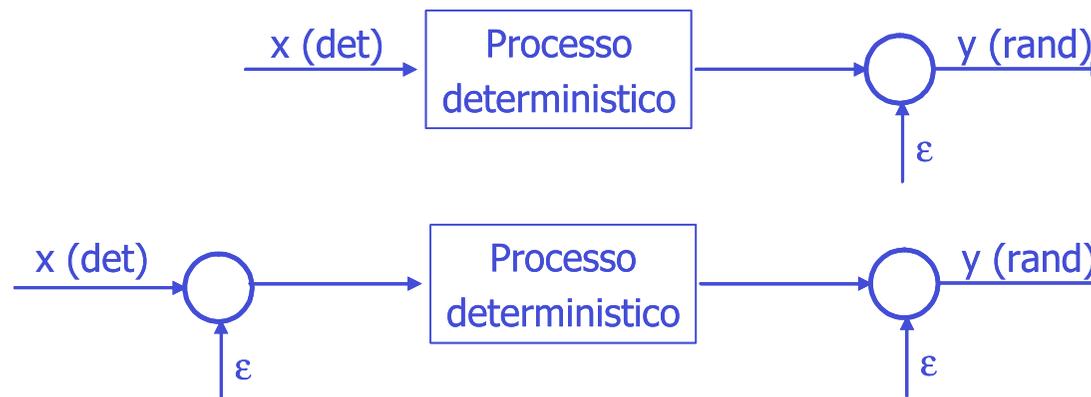
- Nella vita reale ogni esperimento o misura è affetto da errori: se si effettuano due esperimenti con lo stesso protocollo il risultato varia, ovvero le risposte sono **aleatorie**.
- Se si trascurano gli **errori sistematici** l'analista è in genere interessato a determinare sulla base di un numero finito di prove:
  - La tendenza centrale di una variabile dipendente
  - La dispersione dei dati intorno alla tendenza centrale
  - L'incertezza delle stime che si possono fare sulla base dell'analisi statistica dei dati sperimentali.

# Introduzione

- Modello di un processo - **Deterministico**



- Modello di un esperimento - **Aleatorio**



# Introduzione

- **Il modello del processo è deterministico:**
  - L'equazione del viscosimetro associa ad un peso un tempo di corsa in modo esatto, noti i parametri.
- Il risultato sperimentale è invece influenzato da errori casuali.
- Per poter stimare in modo efficiente i parametri dovremo saper costruire **modelli per l'esperimento.**
- Abbiamo visto che i risultati sperimentali sono caratterizzati da distribuzioni di valori e che possono essere razionalizzati in termini di frequenze.
- **Una frequenza può essere collegata ad una probabilità**, il modello dell'esperimento dovrà tenere conto di questo aspetto probabilistico.

# Prove ripetute

- Quando facciamo una misura  $y_1$  di una quantità  $Y$ , l'osservazione sperimentale approssima il valore di  $Y$ . Se ripetiamo la misura  $y_2 \neq y_1$  ma entrambe approssimano il valore vero.
- Facendo altre misure vediamo emergere una struttura nei dati che raccogliamo
  - Se siamo bravi e non commettiamo errori sistematici i valori tendono ad addensarsi attorno al valore vero
- Con una campagna sperimentale si ottengono sequenze di osservazioni che vengono riportate nell'ordine in cui si determinano.
- Per esempio in una sperimentazione abbiamo raccolto i trenta dati riportati nella tabella.

1.2	1.0
1.3	1.2
1.1	1.1
1.4	1.1
1.0	0.9
1.2	1.5
1.2	1.2
1.3	1.2
1.4	1.2
1.0	1.4
0.8	1.3
1.1	1.1
2.0	0.7
1.1	1.1
1.2	1.1

# Campione

- La tabella riporta un sottoinsieme, il **CAMPIONE**, di tutti gli infiniti valori che avremmo potuto raccogliere, la **POPOLAZIONE** o **SPAZIO CAMPIONARIO**.
- La nozione intuitiva di campione (parte di un tutto) rimanda a quella di popolazione (intendendosi con questa la totalità dei casi).
- La nozione di popolazione è importante perché è sui suoi elementi (unità statistiche) che vengono rilevate le informazioni (dati) necessari a comprendere i fenomeni di interesse.
  - In statistica le caratteristiche di interesse vengono definite **Variabili**
  - Solitamente le variabili sono indicate con lettere maiuscole ( $X$ ) ed i valori con lettere minuscole ( $x$ ).

Campione e  
Popolazione

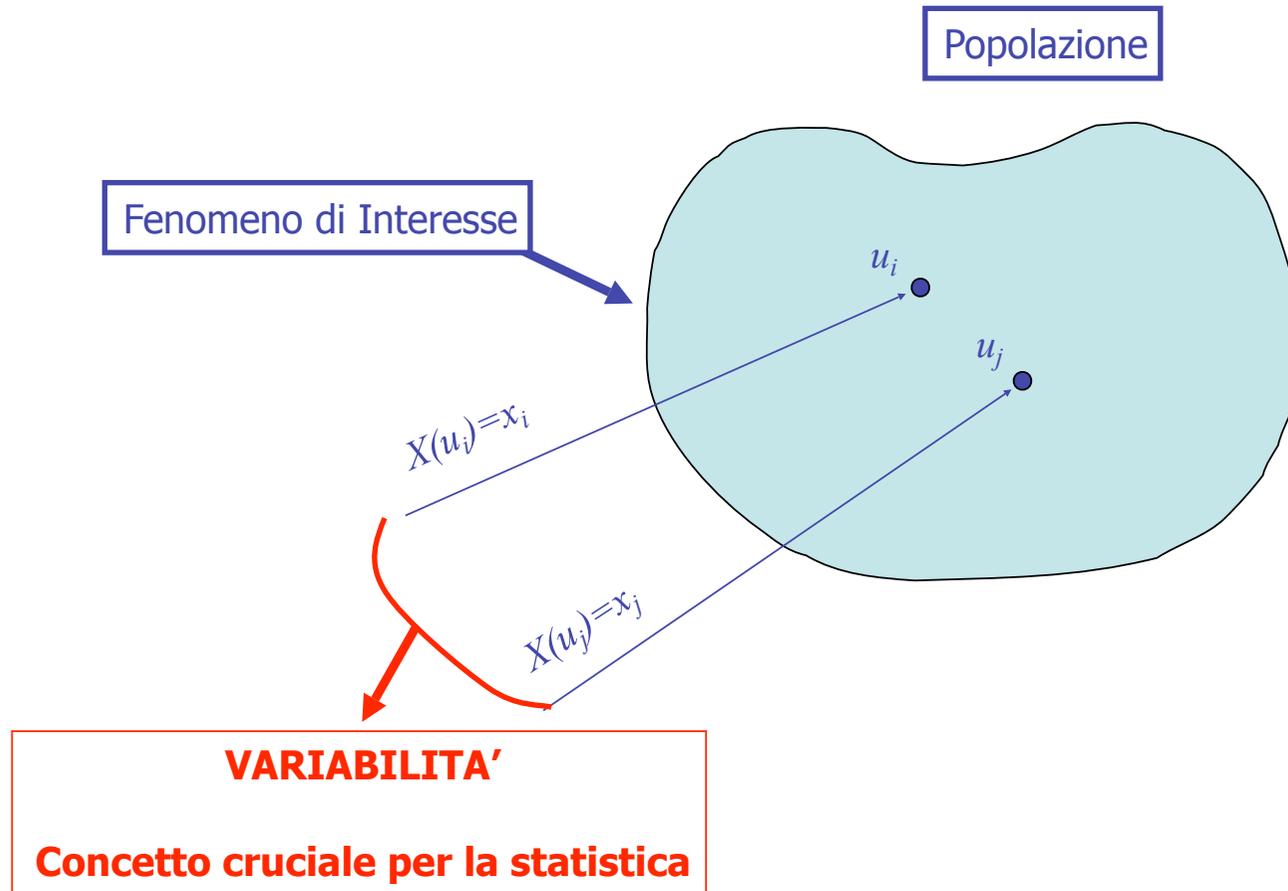
Variabili

# Campione

- Un primo aspetto significativo: come si genera il campione.
  - Nel caso di esperimenti riproducibili ci limiteremo a considerare Campioni Casuali.
  - Per esempio: Estrazione di elementi da un'urna
- Nel caso di **popolazioni finite** il ricorso al campionamento può essere dovuto a:
  - Riduzioni costi e tempi dell'indagine
  - Rapidità con cui si ottengono i risultati
  - Attendibilità delle informazioni
- Nel caso di **popolazioni virtuali**, ovvero composte da infiniti elementi, il ricorso al campionamento è ovviamente indispensabile

**Popolazioni  
virtuali**

# Campione



# Scopi statistici

- I metodi e gli scopi della statistica sono diversi a seconda che la il campione sia:
  - semplicemente caratterizzato **STATISTICA DESCRITTIVA**
  - usato per trarre informazioni sulla popolazione da cui è estratto **STATISTICA INFERENZIALE.**
- **Statistica descrittiva**
  - Metodi che comprendono la raccolta, la presentazione e la caratterizzazione di un insieme di dati allo scopo di descriverne in maniera essenziale le proprietà
- **Statistica inferenziale**
  - Metodi che consentono di valutare le proprietà di una popolazione sulla base delle proprietà di un campione.

**Statistica  
Descrittiva**

**Statistica  
Inferenziale**

# Concetti importanti

- Variabilità del campione
- Istogrammi
- Frequenza relativa ed assoluta
- Errore di misura – Incertezza
- Errori sistematici
- Accuratezza e precisione
- Modello del processo
- Modello dell'esperimento
- Stima dei parametri
- Verifica di ipotesi
- Popolazione e campione
- Statistica descrittiva
- Statistica inferenziale