

Ulteriori Conoscenze di Informatica elementi di Statistica

Carlo Meneghini

Dip. di fisica - via della Vasca Navale 84,
st. 83 (I piano) tel.: 06 55 17 72 17

meneghini@fis.uniroma3.it

Scopi del corso

-1- Interpretazione dei risultati sperimentali

Quanta fiducia posso riporre nei risultati di un esperimento? Quale è la probabilità che il risultato di un'esperimento/misura rifletta il reale stato delle cose?

-2- Utilizzo "avanzato" del PC per l'analisi e la rappresentazione dei dati sperimentali:

EXCEL: funzioni avanzate (strumenti di analisi dati),

Gnuplot: rappresentazione dei dati e fit nonlineari)

Piano del corso: 3 c.f.u.

Lezioni frontali ~ 1.5 c.f.u = 12h = 6/7 lez. (mart. 9-11 aula 3)

- L.1) Introduzione. Sicurezza, configurazione e ottimizzazione del proprio PC.
- L.2) a. Rappresentazione dei dati: plot e istogrammi.
b. Calcolo numerico: derivate, integrali, equazioni differenziali.
- L.3) Richiami di statistica descrittiva: distribuzioni di probabilità, media, varianza, covarianza e correlazione, intervalli di confidenza. Uso di funzionalità avanzate di EXCEL.
- L.4) Fit dei dati: regressione lineare, metodo dei minimi quadrati, fit nonlineare.
- L.5-6/7) test statistici (t-test, c2, F-test, tabelle di contingenza, etc...

Esercitazioni ~ 1.5 c.f.u = 6/7 lez. (lab. informatica, dip. di fisica Via della vasca navale 84, orario 14-17, giorni: da definire)

- E.1) Rappresentazione grafica di dati: istogrammi e plot. Calcolo numerico: derivate, integrali, eq. differenziali.
- E.2) Statistica descrittiva utilizzando un foglio elettronico (EXCEL). Media, varianza, covarianza e correlazione. T-test, intervalli di confidenza
- E.3) Fit dei dati: regressione lineare, linearizzazione dei dati.
- E.4) Fit dei dati: fit non lineare (Gnuplot)
- E.5-6/7) test statistici (t-test, χ^2 , F-test, tabelle di contingenza, etc...

Esami:

prova pratica di analisi dati con discussione
riguardo al metodo utilizzato e sui risultati
ottenuti.

Testi consigliati

S. A. GLANTZ: Statistica per discipline
biomediche - McGraw Hill

Introduzione all'uso di un foglio elettronico (EXCEL)
per applicazioni statistiche. Testo a cura dell'ISTAT
scaricabile on-line dal sito:

<http://www.istat.it/servizi/studenti/binariodie/scuola.htm>

Appunti, lucidi, esercizi, comunicazioni e altro:
www.fis.uniroma3.it/~meneghini

Ivo B.-Birula, Iwona B.-Birula *Modeling reality. how
computers mirror life* - Oxford University press

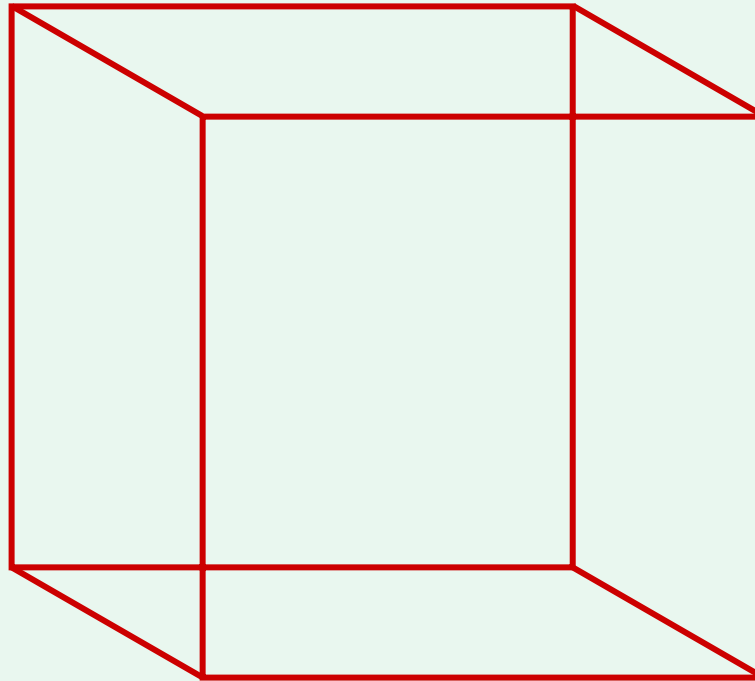
Prendere una decisione richiede informazione

G. Barbieri: Statistica Ufficiale, patrimonio di tutti (2004)



Giovanni Buridano ~1300

L'illusione della certezza



cubo di Necker (1832)

La propensione a ignorare l'incertezza è innata

LA STATISTICA (Trilussa)

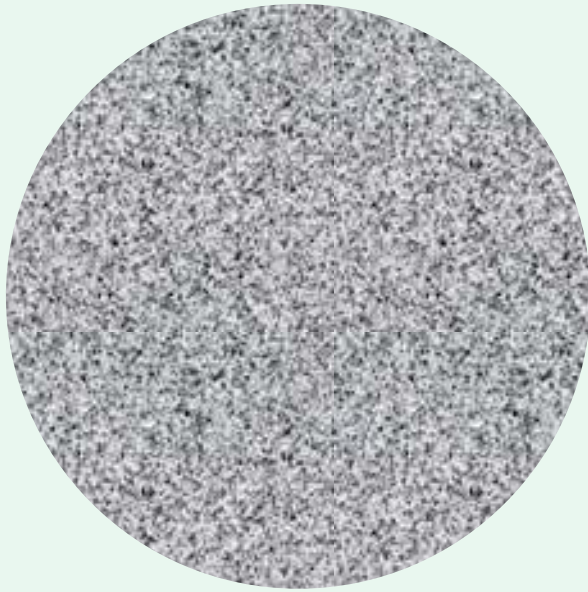
Sai ched'è la statistica? E' 'na cosa
che serve pe' fa' un conto in generale
de la gente che nasce, che sta male,
che more, che va in carcere e che sposa.

Ma pe' me la statistica curiosa
è dove c'entra la percentuale,
pe' via che, lì, la media è sempre eguale
puro co' la persona bisognosa.

Me spiego, da li conti che se fanno
seconno le statistiche d'adesso
risurta che te tocca un pollo all'anno:

e, se nun entra ne le spese tue,
t'entra ne la statistica lo stesso
perchè c'è un antro che se ne magna due.

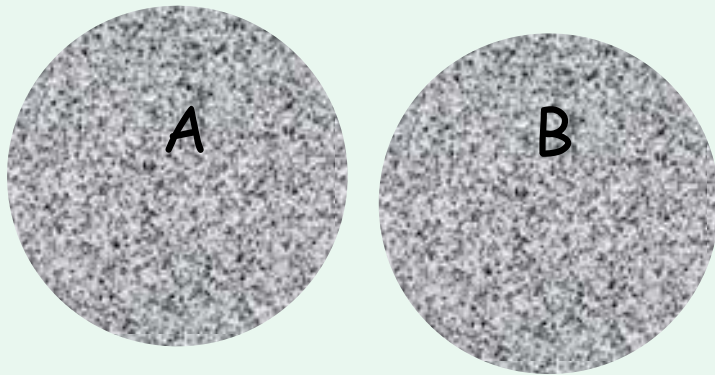
Coltura batterica: descrizione



Statistica descrittiva:

- Tasso di crescita,
- densità di batteri
 - mortalità,
 - vita media, ...

confronti



Analisi statistica:

quanto sono affidabili/significative le analogie e le differenze tra risultati di diversi esperimenti?

Statistica: analisi quantitativa dei fenomeni collettivi che hanno attitudine a variare, allo scopo di descriverli e di individuare le leggi o i modelli che permettono di spiegarli e di prevederli.
(dizionario Garzanti)

La **statistica descrittiva** studia i criteri di rilevazione, *classificazione* e di *sintesi* delle informazioni relative a una popolazione oggetto di studio. La statistica descrittiva raccoglie le informazioni sulla popolazione o su una parte di essa (**campione**) in distribuzioni e sintetizza i risultati attraverso **famiglie** di indici: **valori medi**, **variabilità**, **forma**, rapporti statistici, relazioni statistiche etc...

La **statistica inferenziale** è l'insieme dei procedimenti di generalizzazione che, avvalendosi dei metodi del calcolo delle probabilità, consentono di estendere all'intera popolazione le informazioni fornite da un campione.

La **test statistici** è una regola di decisione: permette di verificare Ipotesi sui parametri

Fenomeni collettivi:

moto delle particelle in un gas, previsioni metereologiche, andamento dei prezzi, evoluzione di un ecosistema, etc...

LA STATISTICA

riduce la fonte di incertezza creando modelli (basati sul calcolo delle probabilità) in grado di conservare l'informazione quantitativa riducendo la molteplicità dei fenomeni. Riduce l'incertezza e permette di quantificare il rischio

Rischio

Il rischio indica la probabilità del verificarsi di un evento

INCERTEZZA

Un risultato è incerto quando non è possibile, in base ai dati a disposizione, assegnare un valore di probabilità ai diversi esiti possibili

1) cerca di tenere sotto controllo l'incertezza che deriva dal trattare fenomeni collettivi (molti individui/molti eventi)

2) Semplificare la complessità mediante modelli basati sul calcolo delle probabilità. Conservare l'informazione riducendo la molteplicità dei fenomeni

3) Calcolo quantitativo del rischio

Definizioni di Probabilità

a) Frequenza relativa di un evento all'interno di una classe di riferimento, calcolata per un numero molto grande (infinito) di osservazioni.

Es.: Su 100 000 prove ottengo 95750 risultati positivi, quale è la probabilità di un risultato positivo per il 100 001 esperimento?

b) Propensione è legata alle caratteristiche fisiche.

Es.: la probabilità che esca testa lanciando una moneta è 0.5, la probabilità che esca 6 lanciando un dado è $1/6$, etc...

c) Probabilità soggettiva di un evento è il grado di fiducia che un individuo attribuisce al verificarsi dell'evento. Tipico nel passare dai modelli alla realtà:

In una simulazione ottengo una sopravvivenza pari al 95 % per un dato comportamento. Quale è la probabilità che la popolazione reale sopravviva?

Qual'è la probabilità che domani piova?

La decisione richiede informazione.

Piú l'informazione è completa piú la decisione è semplice:

Fenomeni deterministici: informazione esatta e completa

Fenomeni non deterministici: informazione incompleta e incerta

Quanto sono deterministici i fenomeni?

previsioni del tempo

orario dell'autobus

traiettoria di una pallina

temperatura di ebollizione
dell'acqua

voto di laurea

diffusione di un virus

Andamento dei prezzi

Mercato del lavoro

Stato delle risorse energetiche

evoluzione di una popolazione

Informazione, Statistica e democrazia

Statistica: da **statistico** = *che concerne il governo e la vita dello stato, che riguarda lo studio di fenomeni interessanti la collettività*

La diffusione dell'informazione statistica è strettamente associata alla diffusione delle democrazie (dopo 1830)

La statistica ufficiale in uno stato democratico è un servizio pubblico essenziale per garantire i diritti di eguaglianza (e di scelta libera) a tutti i cittadini.

Nota: l'illusione della certezza è ciò che guida la scelta soggettiva, l'alfabetizzazione statistica è requisito essenziale per una libertà di scelta consaspevole (valutazione del rischio)