La Simulazione

di Valter Di Dio

Con la scorsa puntata abbiamo praticamente esaurito lo studio dei processi aleatori che riguardano le liste di attesa. Ma non tutto ciò che accade nella realtà può essere simulato con l'uso delle file d'attesa. Ci sono alcuni problemi che si evolvono nel tempo e i cui elementi godono di una specie di 'ricordo' di quanto accaduto in precedenza; ovvero le condizioni al tempo t sono una diretta conseguenza di fatti accaduti al tempo t-1. Ad esempio leggendo un libro possiamo coprire una parola e cercare di indovinarla, il più delle volte ci si riesce e questo deriva dal fatto che sebbene la parola in esame possa teoricamente essere una qualsiasi, lo studio delle parole che la precedono ci permette di escludere molti vocaboli e limitare la scelta solo ad alcuni termini altamente probabili. Se vediamo le parole del testo come una serie molto lunga di termini, possiamo dire che l'ennesimo termine della serie dipende, in modo abbastanza deterministico, da quelli che lo precedono. Infatti se si prova a rifare il giochetto con un testo di sole tre parole, la probabilità di indovinare quella giusta diminuisce di molto.

Una serie di termini o più esattamente di dati che segua un andamento del genere prende il nome di serie storica. Generalmente infatti si tratterà di dati raccolti nel tempo e tra loro legati da alcune relazioni; potranno ad esempio essere le temperature in una certa località o la quantità di petrolio estratta da un pozzo o anche il flusso di macchine in una autostrada.

Le serie storiche

Una delle caratteristiche comuni alla maggior parte dei fenomeni fisici o sociali è il fatto che il loro evolversi può essere rappresentato su un grafico cartesiano la cui ascissa indica lo scorrere del tempo. Su un grafico del genere si possono effettuare già ad occhio molte scoperte sulla natura dell'evento e, spesso, è possibile anche sviluppare un minimo di previsione sull'andamento generale. Se osservate la figura 1 che mostra l'andamento della temperatura per alcuni giorni, si vede chiaramente l'evoluzione sinusoidale del fenomeno e, da questo, si può prevedere che nelle prossime ore la temperatura tenderà a scendere. Osservate ora la figura 2 che indica la produzione di alcool da un distillatore; in questo caso l'andamento 'bizzarro' del fenomeno non permette di valutare ad occhio i valori relativi ai prossimi istanti, nemmeno di quelli più vicini. Si nota però una certa stazionarietà del fenomeno che oscilla (piuttosto violentemente) intorno ad un valore medio stabile (circa 50).

Tutti e due i grafici che abbiamo visto sono esempi di serie storiche: serie in quanto successioni di valori, storiche perché riferite allo scorrere del tempo. È comunque evidente che, nel caso in cui la variabile non fosse il tempo, tutto il discorso resta valido lo stesso: si tratta di cambiare solo i nomi.

Le serie storiche si dividono in due categorie: continue e discrete. Per lo studio teorico si considerano di solito le serie continue perché in matematica le funzioni continue godono di particolari proprietà che consentono in genere maggiori libertà. Ma nella realtà di tutti i giorni le serie sono spesso ottenute come estrazione periodica di dati da una serie continua: come la misura della temperatura giornaliera effettuata ad ore prestabilite, oppure come raccolta di dati in un periodo di tempo fisso: ad esempio il numero di nascite giornaliere in un ospedale.

Per la simulazione si useranno allora sempre serie discrete, sia perché il computer le digerisce meglio, sia perché sono principalmente queste, quelle con cui si lavora nella pratica.

Lo studio dell'andamento di un fenomeno, fisico o sociale che sia, descritto da una serie storica si sviluppa in tre fasi:

1) l'analisi della serie per individuare le caratteristiche salienti;

2) la costruzione di un modello matematico in grado di rappresentare la serie;

3) la previsione dei valori futuri che può essere effettuata sia dal modello descrittivo della serie, che da un modello previsionale il quale, pur non essendo in grado di interpolare i valori attuali, riesce tuttavia a generare i valori futuri.

Può capitare anche il caso in cui ci siano più serie legate tra loro da particolari relazioni, in tal caso occorre anche evidenziare in che modo l'evolversi di una serie influisce sullo sviluppo delle altre e costruire quindi le relative 'funzioni di trasferimento' che esprimono, in forma matematica, le relazioni funzionali tra gli eventi.

Le serie storiche "prezzi al dettaglio" e "domanda" sono un classico esempio di serie correlate: le variazioni di una inducono variazione sui prossimi risultati dell'altra e viceversa.

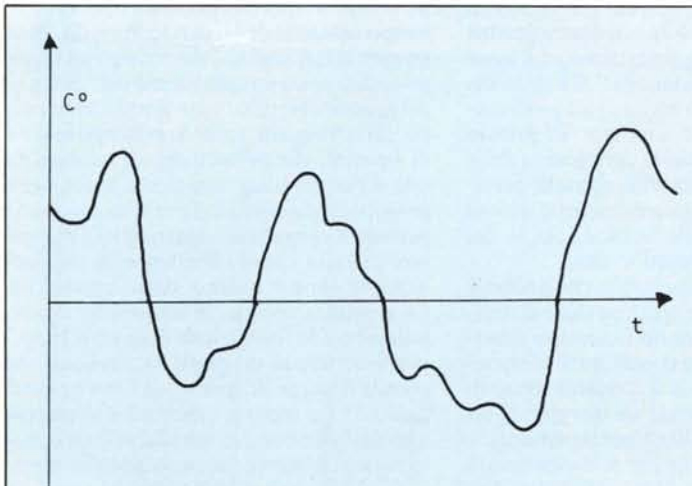

*Figura 1
Andamento della temperatura nell'arco di alcuni giorni.*

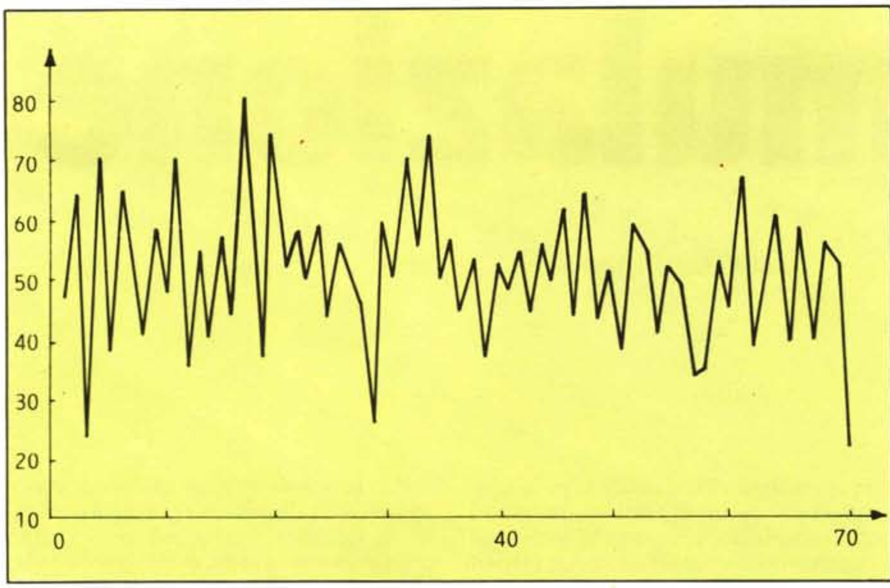

Figura 2 - Andamento dei valori di produzione di una colonna di distillazione.

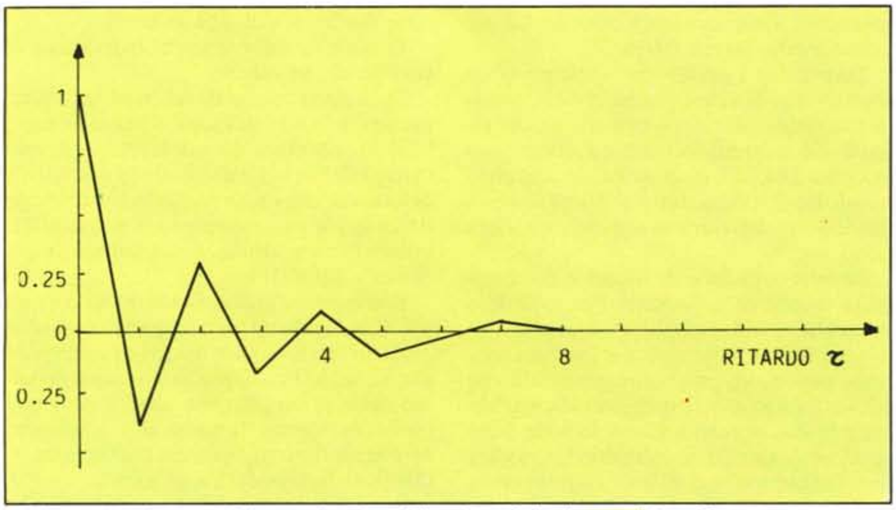

Figura 3 - Correlogramma della serie di prodotti della colonna di distillazione di figura 2

Classificazione delle serie

La prima suddivisione delle serie storiche riguarda il modo con cui sono generati i dati: questi possono essere deterministici od aleatori. Le serie deterministiche sono ad esempio la serie dei numeri di Fibonacci oppure la tensione di uscita di un generatore ideale di corrente alternata; in pratica tutte quelle serie in cui la conoscenza della funzione è sufficiente alla corretta previsione. Nelle serie deterministiche infatti non compaiono, nelle funzioni che le descrivono, elementi legati al caso.

Va comunque considerato che anche le serie deterministiche, qualora siano il risultato di misurazioni di un fenomeno deterministico, si possono considerare come serie aleatorie in quanto il concetto stesso di misurazione sottintende un margine di errore e quindi una distribuzione casuale.

Le serie deterministiche si dividono ancora in periodiche e non periodiche. Appunto la tensione ai capi di un generatore di corrente alternata è una serie storica deterministica e periodica, la sua funzione è infatti una sinusoide. I valori di una serie periodica si ripetono dopo un intervallo di tempo costante detto perciò Periodo. Non sempre le funzioni che descrivono una serie periodica sono semplici come nell'esempio del generatore; in tal caso si utilizzano particolari strumenti, come lo sviluppo in serie di Fourier, che permettono di estrarre da una serie periodica complessa le sue componenti fondamentali. In pratica ogni serie periodica comunque ingarbugliata può essere pensata come l'inviluppo di più funzioni, in seno e coseno, dette armoniche. Le ampiezze con cui le armoniche contribuiscono alla formazione della serie si possono mettere in un grafico cartesiano che prende il nome di spettro di Fourier (vedi figura 5). Lo spettro, a meno di una piccola approssimazione dovuta allo sviluppo matematico, descrive univocamente la serie.

Anche le serie non periodiche con funzioni complicate possono essere semplificate con una speciale funzione che si chiama integrale di Fourier e che può essere considerato come un caso particolare dello sviluppo in serie di Fourier, ma con periodo infinito. Non ci occupiamo oltre delle serie deterministiche perché non interessano ai fini della simulazione, ma bisogna tener presente che molte serie storiche sono il risultato di una serie deterministica cui si sovrappone una specie di indecisione, sia dovuta ad errori di misura che all'influenza di molti parametri esterni che vengono riuniti e descritti da un'unica variabile casuale chiamata, di solito, rumore.

La figura 4 mostra una suddivisione generale delle serie storiche. Come si vede le serie casuali si dividono in stazionarie e non stazionarie. Sono stazionarie quelle serie che tendono a mantenere media e varianza costanti nel tempo, come quella già vista in figura 2. Le serie non stazionarie si dividono a loro volta in serie ad incrementi stazionarie e serie di altro tipo. In generale si ha a che fare con serie stazionarie o ad incrementi stazionari; queste ultime si possono però ricondurre alle serie stazionarie con semplici operazioni. È quindi la classe delle serie stazionarie quella che interessa il maggior numero di applicazioni pratiche.

Abbiamo detto che una serie storica può essere rappresentata da una funzione che lega tra loro realizzazioni successive, questo legame viene evidenziato dalle autocovarianze che indicano come il dato del tempo t sia legato ai dati che lo precedono, le autocovarianze si calcolano quindi per ritardi via via crescenti cominciando da t-1 poi t-2, t-3 e così via, la figura 3 mostra il diagramma delle auto correlazioni del fenomeno di figura 2; questo tipo di grafico prende il nome di correlogramma. Se si verifica che tutte le autocovarianze di una serie sono nulle, e che quindi non c'è alcuna correlazione tra eventi successivi, la serie prende il nome di rumore bianco. Un generatore di numeri casuali è un tipico esempio di serie stazionaria non correlata (almeno teoricamente). Una bella definizione di rumore bianco è stata data da Benoit Mandelbrot: "se noi registriamo su un nastro del rumore bianco e lo riascoltiamo a varie velocità il suono che sentiamo è sempre lo stesso", cosa che invece non accade con il suono di uno strumento in cui gli eventi sono correlati tra loro e una variazione della velocità, e quindi del tempo, modifica la correlazione e perciò il timbro. Se noi andiamo a vedere lo spettro di frequenza del suono di un oboe ci troveremo di fronte a qualcosa tipo la figura 5. I segmenti verticali rappresentano le com-

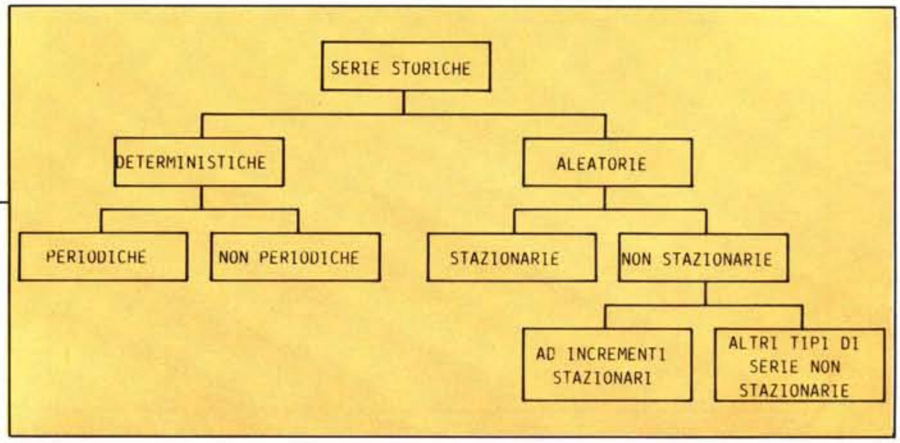

Figura 4 - Classificazione delle serie storiche.

ponenti elementari del suono, il rapporto in ampiezza tra queste componenti determina il fatto che il suono sia quello di un oboe e non di un violino. Quando si riascolta una registrazione a velocità diversa dall'originale le componenti spettrali si sono traslate e il suono è diverso. Per il rumore bianco invece l'analisi spettrale rivela la presenza di tutte le possibili frequenze e tutte della stessa ampiezza; per cui qualsiasi traslazione o restringimento si effettui sullo spettro, il suono risultante non cambia.

Il rumore bianco più comune è quello generato dall'agitazione termica delle molecole in un conduttore, e che noi ascoltiamo alla radio come fruscio o vediamo sullo schermo del televisore come una fitta e turbolenta nevicata quando non ci sono emittenti sintonizzate.

Per generare con un computer una serie che abbia le caratteristiche del rumore bianco basta utilizzare il seguente programma:

```
10 R=RND(1)
20 PRINT R
30 GOTO 10
```

All'estremità opposta del rumore bianco si trova il rumore scuro, nome preso dal fatto che si riferisce ai moti Browniani (brown = scuro), in cui ogni dato corrisponde ad un incremento o decremento casuale del dato precedente. I moti Browniani si usano per descrivere il moto delle molecole di un gas all'interno di un contenitore, mentre si agitano e si urtano continuamente a causa della loro energia termica. Il tracciato di una pallina che rotola su un fondo sconnesso è un altro esempio di moto browniano o rumore scuro. Non possiamo sapere dove sarà la pallina tra molto tempo, ma possiamo prevedere con una discreta approssimazione dove si troverà nel prossimo istante.

Nel rumore scuro esiste un'alta correlazione tra i dati, in quanto ciascun dato si ottiene direttamente dal precedente sommando o sottraendo una quantità casuale; ma lo stesso discorso si è fatto per il dato precedente e per il precedente ancora, in pratica esiste un ricordo della strada percorsa dall'inizio della serie fino al momento attuale. Un programma per la generazione di rumore scuro è il seguente:

```
10 R=RND(1)
20 IF RND(1) < .5 THEN X=X+1
                  ELSE X=X-1
30 PRINT X
40 GOTO 10
```

Un esempio grafico di rumore scuro è quello di figura 6.

In mezzo tra il rumore bianco e il rumore scuro, con vari gradi di autocorrelazione, si trovano tutte le serie storiche "vere", quelle cioè che dovremo andare a simulare nel lavoro quotidiano.

Le passeggiate aleatorie

Se immaginate ora di seguire i numeri che appartengono a una serie non su un diagramma cartesiano, ma come un punto che si muove a balzi su una retta, potete capire perché le serie storiche prendono anche il nome di passeggiate aleatorie. Nei casi in cui la serie è composta da due variabili il concetto di passeggiata si fa ancora più evidente. Il tracciato di una formica o i salti di una pulce sono esempi lampanti di passeggiate aleatorie e, come per le serie, le differenze tra i due tipi di moto dipendono dall'autocorrelazione dei dati. Nel caso della formica l'autocorrelazione è molto alta e il tracciato corrisponde ad una serie 'scura'; la formica cammina in una certa direzione per un periodo casuale poi cambia direzione a caso e riprende a camminare per un tratto casuale e così via: ogni nuovo percorso dipende strettamente dai precedenti avendone in comune un punto.

Per la pulce invece il moto è molto più vicino al rumore bianco che ad una passeggiata aleatoria con un minimo di correlazione: se potessimo sporcare d'inchiostro i piedi della pulce i punti neri riempirebbero il piano uniformemente.

Anche molte delle attività umane generano passeggiate aleatorie, una delle più tipiche è la musica. Le note che compongono una sinfonia sono una serie storica correlata (altrimenti si chiamerebbe cacofonia) di note musicali estratte a 'caso' dal compositore. In una sinfonia si trovano molti tipi di autocorrelazione sia tra eventi prossimi (melodia) che tra eventi molto lontani (ritornelli) e anche tra le varie note che compongono gli accordi; infatti, perché siano tali, ci devono essere certi rapporti matematici tra le note che li formano. Il grado di bellezza di una composizione musicale dipende in modo molto stretto dalla autocorrelazione delle varie sequenze di note e dalla varianza locale che genera quella imprevedibilità che sorprende ed entusiasma l'ascoltatore.

Il modo con cui le note di una data composizione si susseguono può essere rappresentato, ovviamente in modo generale e un po' restrittivo, da una matrice di transizione.

La matrice di transizione descrive la probabilità che all'Xesimo dato del tempo t segua al tempo t+1 il dato Y. Facciamo

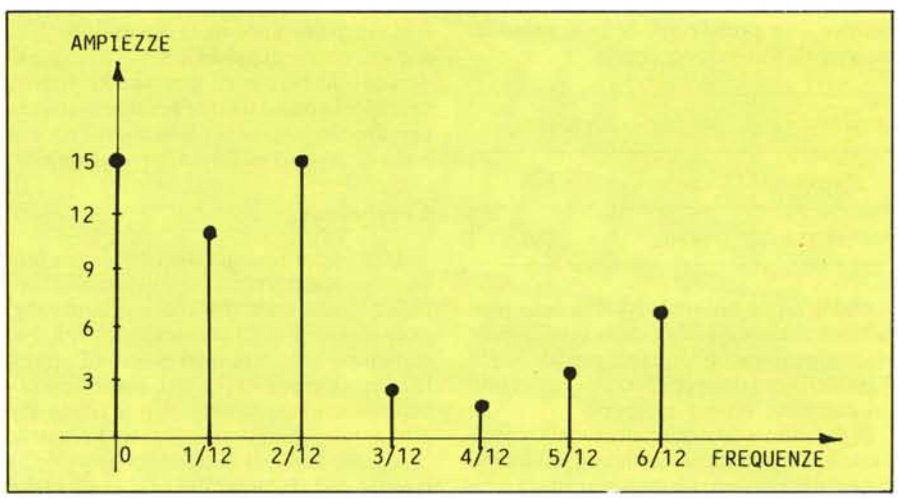

Figura 5 - Spettro di frequenza del suono di uno strumento musicale.

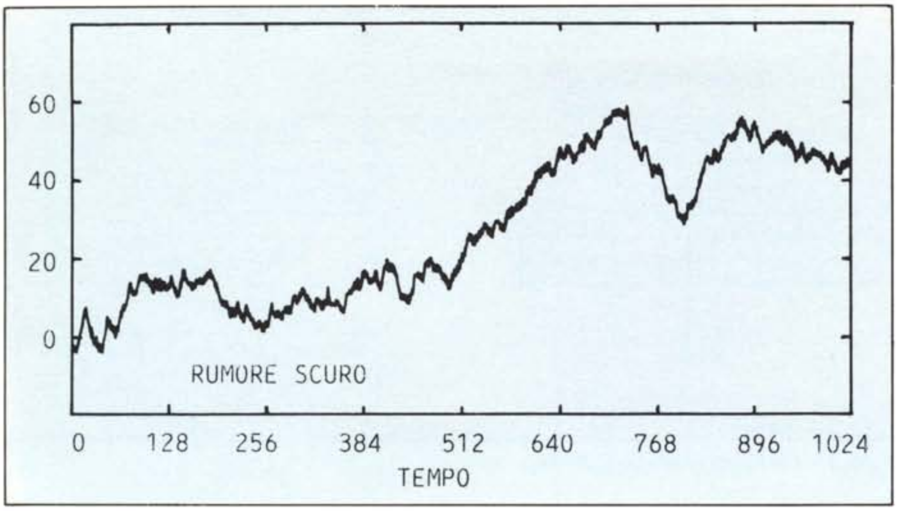

Figura 6 - Esempio di rumore scuro o Browniano.

```
10  GOSUB 200
20  IF CIELO$ = "SERENO" THEN 100
30  IF  RND (1) < T(1,0) THEN CIELO$ = "SERENO"
40  PRINT CIELO$
50  GOTO 20
90 :
100  IF  RND (1) > T(0,0) THEN CIELO$ = "COPERTO"
110  GOTO 40
150 :
200  REM  VARIABILI
210 CIELO$ = "SERENO"
220 T(0,0) = .640
230 T(0,1) = 1 - T(0,0)
240 T(1,0) = .420
250 T(1,1) = 1 - T(1,0)
260  RETURN
```

Figura 7
Minisimulazione di previsioni metereologiche.

un esempio. Per un periodo di quattro anni si rileva ogni giorno, in una data località (Tel Aviv), lo stato del cielo che può essere solo di due tipi: sereno o coperto. Trascorso il periodo di tempo dell'esame si conta quante volte si sia verificato ciascun tipo di transizione. Le transizioni possibili nel nostro caso sono quattro:

da cielo sereno a cielo sereno
da cielo sereno a cielo coperto
da cielo coperto a cielo sereno
da cielo coperto a cielo coperto

Si divide poi il numero di transizioni contate per il numero di giorni presi in considerazione e il risultato rappresenta la probabilità di un certo tipo di cambiamento. Si può allora scrivere il tutto in una matrice che prende appunto il nome di matrice di transizione: eccola

DA \ A	SERENO	COPERTO
SERENO	.640	.360
COPERTO	.420	.580

Come si può notare la somma delle probabilità di ciascuna riga della tabella deve essere uguale ad uno, questo perché il cielo il giorno dopo deve per forza essere in uno dei due stati: sereno o coperto.

Si può allora costruire un modello previsionale semplicemente partendo dal tempo di oggi e generando un numero casuale con le probabilità derivate dalla tabella.

Il programma listato in figura 7 permette di simulare l'evolversi del tempo di una certa località in base alla storia precedente. Ovviamente la matrice di transizione non permette di utilizzare tutte le informazioni della serie, manca infatti qualsiasi riferimento ad autocorrelazioni, oscillazioni stagionali (riconoscibili nello spettro) e rumori vari di sottofondo. Un minimo di autocorrelazione è possibile recuperarlo con una matrice di transizione in cui si considerano sequenze di eventi anziché eventi elementari. Nel caso precedente si dovrebbero, ad esempio, segnare i passaggi tra due giornate serene e una coperta o la probabilità di avere una giornata coperta dopo una serena seguita da una coperta e così via. Naturalmente la dimensione della matrice cresce di parecchio, già solo considerando le coppie di giornate la matrice precedente passa da due per due a quattro per due con le terne si passa a nove per due e via di seguito in forma esponenziale.

Conclusioni

Terminiamo qui questo articolo per problemi di spazio, ma continueremo la trattazione delle serie storiche e delle passeggiate aleatorie nella prossima puntata. Nel frattempo chi fosse interessato ad approfondire gli elementi trattati, molto superficialmente, in questi articoli può rifarsi alla bibliografia pubblicata a fianco. Per quello che riguarda i testi di statistica generale ho inserito nella bibliografia solo quelli utilizzati nella stesura dell'articolo. MC

Bibliografia

Statistica generale

H. Cramer:
Mathematical Methods of Statistics
Princeton University Press, 1946

M. Fisz:
Probability Theory
and Mathematical Statistics
Wiley, 1963

A. Mood - F. Graybill:
Introduction to the Theory of Statistics
Mc Graw Hill, 1963

In italiano

A. Rizzi - T. Salvemini:
Lezioni sulla analisi della regressione
La Goliardica Editrice, 1975

A. Serrecchia:
Lezioni di inferenza statistica classica
Istituto di calcolo delle probabilità
Università di Roma, 1977 (dispense)

Processi Aleatori

E. Parzen:
Stochastic Processes
Holden Day, 1962

D.R. Cox - H.D. Miller:
The Theory of Stochastic Processes
Chapman & Hall, 1965

J. Mc Kinsey:
Introduction to the Theory
of Communication
University of Illinois Press, 1949

In italiano

F. Carlucci:
Introduzione all'analisi delle serie storiche ed ai modelli autoregressivi a somma mobile
Istituto di calcolo delle probabilità, Università di Roma, 1975 (dispense)

Ricerca Operativa

R.A. Fisher:
Statistical Methods of Research Workers
Oliver & Boyd, 1954

F.A. Graybill:
An introduction to Linear Statistical Models
Mc Graw Hill, 1961

P. Morse:
Queues, Inventories, Maintenance
Wiley, 1958

J.R. Raser:
Simulation and Society
Allyn & Bacon, 1969

In italiano

A. Ruscitti:
Grafi probabilistici: tecniche reticolari GERT
La Goliardica Editrice, 1975

U. Colombo:
Applicazione della teoria dei grafi alla R.O.
Istituto di calcolo delle probabilità,
Università di Roma, 1970 (dispense)

P.L. Piccari:
Manuale di controllo della qualità
e di affidabilità
ISEDI, 1974

Autori Vari:
I limiti dello sviluppo:
verso un equilibrio globale
a cura del MIT e del Club di Roma,
Biblioteca EST, Mondadori, 1973

J.L. Taylor:
I giochi di simulazione nell'organizzazione del territorio
F. Angeli, 1976

5·9 settembre 1985
Fiera Milano

SIM HI-FI IVES

19° salone internazionale della musica e high fidelity
international video and consumer electronics show

padiglioni 16-17-19-20-21-41F-42

Segreteria generale SIM-HI-FI-IVES
Via Domenichino, 11 - 20149 Milano
Tel. 02/48.15.541 (r.a.)
Telex 313627

ASSOEXPO

Ingressi: Porta Meccanica (P.za Amendola)
Porta Edilizia (V.le Eginardo)
Orario: 9.00 - 18.00

Strumenti musicali, P.A. System, Apparecchiature Hi-Fi, Attrezzature per discoteche, Musica incisa, Broadcasting, Videosistemi, Televisione, Elettronica di consumo, Videogiochi, Home computers

Il più eccitante appuntamento europeo con la musica, l'hi-fi, il computer e il video è alle porte! Segnati le date: dal 5 al 9 Settembre!

Partecipa anche tu alla GRANDE CACCIA AL TESORO *con migliaia di premi ed un omaggio per tutti!*