Riassunto Computational Statistics

# Moto Browniano

Il Moto browniano è un fenomeno al centro della teoria dei processi stocastici, ciò avviene ad esempio quando si mettono piccole particelle nell’acqua, il loro movimento è infatti casuale.

Secondo Einstein, il moto casuale è dovuto all’azione causata dagli urti tra le particelle che, anche se di bassa intensità, sono in gran numero e il loro insieme lo rende tale. Di conseguenza non si può sapere dove andrà la particella in base all’urto.

La quantità di una particella non è quindi il semplice spostamento x(t) di una particella ma la media di questi, tuttavia questa fa 0 dato che gli spostamenti sono opposti, quindi si considerano solo quelli positivi:

< x(t)2>= 2D2t

dove D è il coefficiente di diffusione calcolabile con la seguente formula:

dove:

* Kb è la costante di Boseman;
* T è la temperatura;
* n è la viscosità;
* R è il raggio.

Un modo per descrivere questi moti in funzione della velocità e della posizione è l’equazione di Langevin.

# Random Walk

La random walk è un processo stocastico (casuale) in cui viene ricavata una quantità s lanciando una moneta: se è testa, s è positiva, altrimenti è negativa.

Questo processo è discreto dal momento che i suoi intervalli e i suoi valori lo sono, si può rendere continuo diminuendo ogni volta gli intervalli, ottenendo il cosiddetto Processo di Wiener.

<grafico random walk>

Con l’equazione di Lengevin si studia la diffusione delle particelle come segue:

dove indica l’incremento casuale di un processo di Wiener.

# 

# Variabili casuali

Una variabile casuale x^ dipendente da z (il risultato di un esperimento), è un numero che indica tutte le possibili uscite, queste ultime corrispondono allo spazio degli eventi S.

Da ciò si possono ricavare i concetti di:

* Distribuzione di probabilità: probabilità che una variabile casuale sia minore o uguale di un valore dato:
* Densità di probabilità: derivata della distribuzione:

per le variabili continue

(x) =piδ(x−xi) per le variabili discrete (dove pi = P{x^ = xi)

Nel caso in cui si voglia calcolare la probabilità che x^ cada in un intervallo x1-x2:

## Funzioni di variabili casuali

Data una variabile casuale x^ e una funzione g, y^=g(x^) corrisponde a una nuova variabile casuale, quindi:

* La distribuzione di y^ equivale alla distribuzione di x^ in cui si applica la funzione g al parametro:
* Per quanto riguarda la densità di y^, si procede nel seguente modo:

dati n valori di x, allora , quindi

.

## Note sulla densità di probabilità

* Se la densità fx è costante, allora lo è anche fy, anche se solo in un dato intervallo;
* Con e quindi
* se e quindi se, allora
* se con e quindi se fx è gaussiana standard,

## Valore atteso

Data una variabile casuale x^, E{x^} è il suo valore atteso:

* nel caso continuo è l’integrale del prodotto di ogni valore per la sua densità;
* nel caso discreto, è la somma di questi prodotti:

## Media condizionata

La media condizionata è il valore atteso condizionato da un evento M:

* Nel caso continuo:
* Nel caso discreto:

Esempio:

## Varianza

La varianza di una variabile x^ si calcola nel seguente modo:

* Nel caso continuo:
* Nel caso discreto:

Esempio: x^ assume valore 1 con probabilità p e 0 con q=1-p

quindi

Nel caso in cui non si conoscano p e q, si fa una stima con la media degli scarti:

dove n è il numero di gradi di libertà, ovvero il numero di variabili indipendenti.

## Momenti

Il momento di ordine n di una variabile x^ è il valore atteso di x^n:

Il momento centrato di ordine n è il valore atteso di x^n meno la sua media:

Il momento assoluto di ordine n è il valore atteso del valore assoluto di x^n:

Tra questi momenti vi sono le seguenti relazioni:

* m0 = u0 = 1;
* u1 = 1;
* m= m1;
* =u2;

## Sviluppo in serie per il valore medio:

1. Decidere intorno a quale valore effettuare lo sviluppo, in questo caso m;
2. E{g(x^)}= g(m)+0+g[2](m)\*sigma^2/2+ g[n](m)\*un/n!

E{g(x^)}=sum[i=0,n](g[i](m)\*ui/i!=g(m)+g[2](m)\*sigma^2/2=my

## Sviluppo in serie per la varianza:

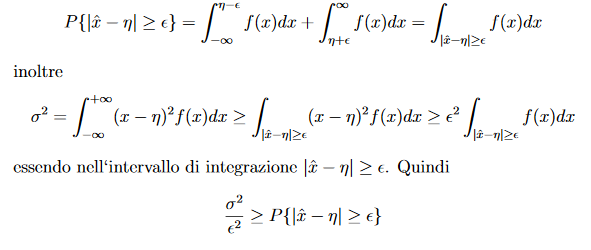
<Sviluppo per la varianza>

## Disuguaglianza di Chebychev

La disuguaglianza di Chebychev è la misura di concentrazione di una variabile x^ intorno al suo valore atteso.

Infatti, se si considera la probabilità , questa definisce l’intervallo [m-, m+].

### Dimostrazione



### Note:

Se la varianza è 0, la probabilità che x^ sia fuori dall’intervallo citato prima è 0 per ogni , di conseguenza x^ = con probabilità 1.

## Probabilità Gaussiana

La probabilità gaussiana è la probabilità P{|ˆx−η|≥3σ}= 2[1−G(3)] = 0.0027 dove G è la distribuzione normale.

## Funzione caratteristica

La funzione caratteristica è la funzione



Questa funzione è legata ai momenti, ciò significa che possono essere ricavati da qui.

Ponendo infatti s=i\*omega, oi(s) è la funzione generatrice di momenti, in cui oi(i\*omega)=OI(omega)

La seconda funzione caratteristica è





Da queste definizione si può scrivere la funzione caratteristica come valore atteso:



Data una gaussiana G[m,sigma]:



Questa funzione è anche detta trasformata.

## Teorema del momento

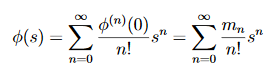
Dato oi(s), se si deriva questa funzione n volte si ottiene che:



ciò significa che



Si può quindi generalizzare lo sviluppo in serie nel seguente modo:



In conclusione, oi genera i momenti per via di quest’ultima affermazione, ciò significa che la funzione caratteristica può essere considerata come la somma di tutti i momenti:

mn=sum(x^n)/N

## Chiarimenti della disuguaglianza di Chebychev

### Cumulanti

Un cumulante è una derivata di ordine n di ui(s) intorno al valore 0.



Dato che





In generale:



Dato che





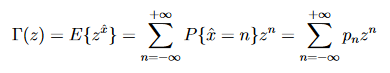
In generale:

oi[n](s)= sum( ui[j](s)\*e^ui(s))

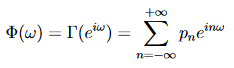
Nel caso discreto, data una variabile x^ con valori discreti xi con probabilità pi:



Si definisce



Quindi



Derivando k volte la funzione L si ottiene che:

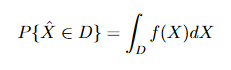


Ponendo z uguale a 1:



# Successione di variabili aleatorie

Dato un vettore di variabili X^ =[x1^,...,xn^] e uno spazio D n-dimensionale, la probabilità che X^ appartenga a D è l’integrale rispetto a D della densità:



Quindi f(X)=f(x1,...,xn) (che dipende dai valori assunti dalle n variabili) equivale alla derivata parziale rispetto a una variabile ( tutte le altre sono di conseguenza costanti):



Ogni xi cade in un intervallo [xi - dxi, xi + dxi].

La distribuzione di probabilità F(X) si calcola come segue:



Esempio:



se x2 e x4 tendono a infinito, quindi:



Si suppone ora che il vettore X^ e k funzioni g1(X),...,gk(X), con queste ultime è possibile creare nuove variabili, quindi:

Y1^=g1(X),...,Yk=gk(X)

Per determinare le statistiche di questi valori, vi sono tre possibili casi:

* Quando k<n, si determina la densità congiunta delle k variabili y;
* Con k>n, si possono esprimere le k-n variabili in più come le prime n, si hanno quindi k-n soluzioni indipendenti.
* Per k=n, si può scrivere un sistema di equazioni:

g1(X)=y1

.

.

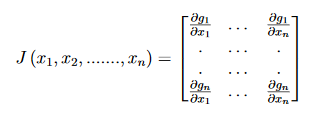
.

gn(X)=yn

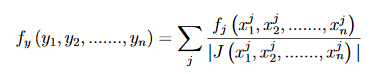
Se il sistema non ha soluzioni, allora fy(Y)=0, altrimenti



dove j(X) è la Jacobiana, equivalente a:



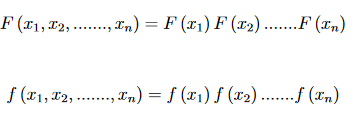
Nel caso in cui vi siano più soluzione, si fa la sommatoria tra queste:



## Indipendenza tra variabili

Date n variabili x1^,...,xn^, esse sono mutuamente indipendenti se gli eventi {x1^ <= x1},...,{xn^ <= xn} sono indipendenti.

Se è così, la distribuzione, allora:



Un caso interessante riguarda gli esperimenti indipendenti e ripetuti:

Dato S^n=S1 x … x Sn come una combinazione di n esperimenti singoli, generalmente una variabile xi^ può dipendere da tutti gli esperimenti.

Si assume che xi^ dipende solamente dall’esperimento i, in tal caso se gli esperimenti Si sono indipendenti, allora lo sono anche tutti gli xi^.

Si suppone ora una funzione g(x1^,...,xn^), da esse si può calcolare il valore atteso:



Inoltre si può aggiungere che:



La media della variabile xi^ mi si calcola nel seguente modo:

mi=E{xi^}= integral(...integral(xi\*f(x1,...,xn)dx1)...dxn)=integral(xi\*f(xi)dxi)

Ciò succede perchè l’integrazione di variabili diverse da xi dà come risultato 1, quindi, date delle coppie di valori (xi^, mi), la covarianza Cij si ottiene facendo:



Nel caso in cui i==j, si ottiene la varianza:



## Correlazione

Date due variabili xi^ e xj^, la correlazione Cij si calcola nel seguente modo:

Cij=E{xi^ \* xj^} - mi\*mj

Se Cij è uguale a zero, xi^ e xj^ sono scorrelate, quindi è possibile costruire una variabile x^ =sum(xi^) avente le seguenti proprietà:

* la varianza di x^ è la somma delle varianze;
* la media di x^ è la somma delle medie.

## Statistiche campionarie

* Media campionaria: xc=sum[i=0,N](xi^)/N
* Varianza campionaria: vc=sum[i=0,N](xi^-xc)/(n-1)

### Dimostrazione

Il valore atteso di xc è la media m

E{xc}=m

Ciò non è altro che la media dei valori attesi delle singole variabili xi:

E{xc}=sum(E{xi^})/n=n\*m/n = m

La varianza invece è sigma^2/n, la quale si ricava nel seguente modo:

E{(xc - m)^2} = E{ (sum(xi^)/n - m)^2} = E{ sum(xi^-n\*m)}/n^2=E{(x1 - m)^2+...+(xn - m)^2}=n\*sigma^2/n^2=sigma^2/n

La varianza di vc si calcola nel seguente modo:

E{vc}=E{(xi^ -m)(xc - m)} = E{(xi^-m)\*[(x1^-m)+...+(xn^-m)]}

Dal momento che le variabili sono tra loro scorrelate, tutte tranne (xi - m) si annullano

E{(xi^-m)\*[(x1^-m)+...+(xn^-m)]}=E{(xi^-m)(xi^-m)}/n= sigma^2/n

## Dipendenza e indipendenza

Date delle variabili x1^,..., xn^, se questo sono indipendenti, allora sono anche scorrelate (non è vero il contrario).

### Dimostrazione

E{x1^, x2^}=integral( integral( z1\*x2\*f(x1,x2)dx1)dx2) dove f(x1,x2)=f(x1)\*f(x2) dato che indipendenti, quindi:

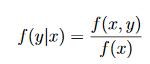
integral(x1\*f(x1)dx1) \* integral(x2\*f(x2)dx2)= m1 \* m2

Cov{x1,x2}=E{x1,x2} - m1\*m2=0

Ciò dice anche che E{g(x1^)\*...\*g(xn^)}=E{g(x1^}\*...\*E{g(xn^)}

## Densità di probabilità condizionata

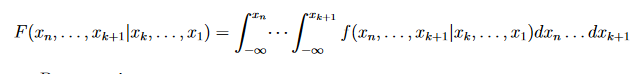
Date x^ e y^, la densità condizionata f(x^|y^) si calcola come segue:



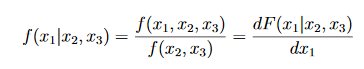
più in generale:



Allo stesso modo si può ricavare la distribuzione:

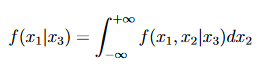


Si può quindi costruire una catene di probabilità:



Tutto ciò si semplifica in f(xi^|xi-1), le cosiddette catene di Markov.

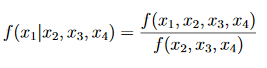
Si può osservare che, dato f(x1,x2,x3), è possibile ricavare f(x1|x3) integrando su x2:



Per calcolare f(x1|x4) invece:

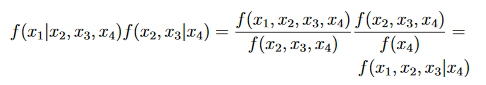


Dato che





si ha che:

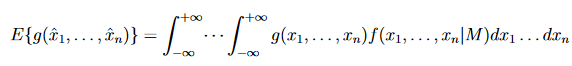


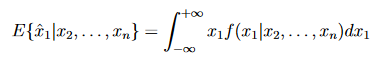
Quindi, per generalizzare il tutto:

* Se si è nella parte sinistra, si integra per portarlo a destra;
* Se invece si è a destra, si moltiplica per la congiunta e si integra per esse per andare a sinistra.

Tutto ciò può essere trasposto nel caso discreto sostituendo gli integrali con le sommatorie.

## Valori attesi condizionati





Come prima, il caso discreto sostituisce le sommatorie agli integrali.

## Funzione Caratteristica

Come già detto in precedenza, la funzione caratteristica OI(omega) si calcola nel seguente modo:

OI(omega) = E{ e^(i\*omega\*x^)

Nel caso delle successioni di variabili, questa funzione diventa:

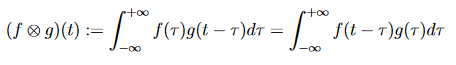


Nel caso in cui le variabili x1^,..., xn^ siano indipendenti e aventi la propria densità fi(xi), si può creare la variabile z come la somma di tutte le xi la cui densità fz è la convoluzione di tutte le fi:





Una convoluzione tra due funzioni f e g è l’integrale rispetto a tau di f(tau) per g(t-tau):



Nel caso in cui le variabili siano indipendenti, il valore atteso di xj^ è il prodotto dei singoli valori attesi:



Quindi la funzione OI per z si trova le seguente modo:

OIz(omega)=E{ e^i\*sum(omega[j]\*xj^)} = E{ e^(i\*omega[1]\*x1^)} \* … \* E{ 

## Teorema della convoluzione

Il teorema della convoluzione indica che la trasformata di una convoluzione è il prodotto di una trasformata:



## Vettori di variabili normali

Dato un vettore X^=[x1^,...,xn^] , i vari xi^ sono congiuntamente normali se e solo se:

AX^t=a1\*x1^+...+an\*xn^ è normale per ogni A, quindi si può dire che:



dove c è la matrice di covarianza e omega è un vettore [omega[1],...,omega[n]].

La densità f(X) si calcola nel seguente modo:



dove delta è il determinante della matrice c.

## Convergenza stocastica

Data la variabile x^ con media a e varianza sigma^2, applicando la disuguaglianza di Chebychev si nota che:





Se sigma<<eps, allora P{|x^ - a| < eps} tende a 1.

L’osservazione di x^(z) si trova nell’intervallo [a -eps, a+ eps]

Se x^ = a + v^ (dove v^ è una variabile a media nulla), allora la singola misura è una soddisfacente stima di a se sigma<<a.

Qualora si voglia rendere più accurata questa misura, si ripete più volte, quindi l’i-esima lettura xi^=a+vi^ si può considerare come la media campionaria.

Se il numero di letture n è grande tale che sigma^2<<n\*a^2, allora x^(z) è una stima soddisfacente di a.

Ora si assume n tale che sigma^2/(n\*a^2) sia uguale a 10^-4, in tal caso la probabilità che xcsia in [0.9\*a, 1.1\*a] (con eps che si conseguenza è uguale a 0.1\*a) é:



da qui si ricava a^2 = sigma^2 / (n\*10-4).

La variabile xc diventa quindi un buon estimatore teorico.

## Legge dei grandi numeri

La legge dei grandi numeri indica che, dato un numero di tentativi n di cui k vinti, la probabilità che |k/n - p|< eps tende a 1:



dove p è la probabilità di vincere un tentativo.

Questo risultato si può ottenere facendo il limite di successione di variabili:

Data xi^ binaria (0→ successo e 1→ insuccesso), per ottenere xc[n] (con n che tende a infinito), si fa la media di tutti gli xi^:

xc[n]=sum(xi^)/n → p

Ciò è possibile perchè E{xi^} = E{xc[n]} = p.

Se invece si fa:



Ciò significa che la varianza di xi^ = uguale a p\*q, quindi:

sigma^2=p\*q → sigma^2[xc] = p\*q/n con p\*q<=¼ dato che p+q=1 e 0<p,q<1

La legge dei grandi numeri può essere quindi espressa nel seguente modo:



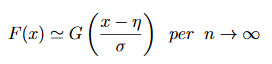
## Teorema del limite centrale

Il teorema del limite centrale indica che, data una generica distribuzione di una variabile z^ che è somma di n variabili x^ indipendenti ( con n tendente a infinito), z tende a una gaussiana:



e avrà mz= sum(mxi) e sigma(z)^2 = sum(sigma(xi)^2)

Se F(x) è la distribuzione di probabilità, allora:



Nel caso in cui le variabili xi siano continue, la densità f(x) tende al seguente valore:



Per semplicità si suppone che le medie mi siano uguali a 0.

Dato che gli xi sono indipendenti, la funzione caratteristica OI(omega) di z^ è il prodotto di quelli degli xi:



Inoltre, prendendo e^( - (sigma^2\*omega^2)/2)



# Processi Stocastici

Un processo stocastico x^(z,t) è una variabile che, oltre a dipendere dall’esperimento z, dipende anche dal tempo t.

Il tempo t è definito su un insieme R, se quest’ultimo è su asse reale, allora x^(t) è continuo, altrimenti è discreto se R è un insieme di numeri interi.

Un processo può essere a stati/tempi discreti o continui, quindi ci possono essere quattro combinazioni:

* stati e tempi discreti: non ci sono valori intermedi tra i vari tempi;
* stati continui e tempi discreti:
* stati discreti e tempi continui:molto simile al precedente con la differenza che vi sono vari intermedi tra i vari tempi;
* stati e tempi continui:

Il processo x^(z,t) può essere considerato come un insieme di funzioni in cui z e t sono le variabili:

* con z e t fissi si ha un semplice numero;
* con z variabile e t fisso si ha una variabile casuale;
* con t variabile e z fisso si che ogni z ha una dipendenza da t;
* con z e t variabili si ha un processo vero e proprio.

Esempio: moto browniano

Si considera x^(t) come il moto delle particelle e x^(z) come il moto della singola particella:

Se x^(t)=r^ é cos(omega\*t+ phi^)

la singola realizzazione è: x^(t,zi) = r^(z)\*cos(omega\*t + phi^(zi))

## Uguaglianza tra processi

Due processi x^(t) e y^(t) sono uguali ovunque se x^(t,z) e y^(t,z) sono uguali per ogni esperimento z.

Se questa condizione è soddisfatta, allora esiste un processo z^(t)=x^(t)+y^(t) tale che z^(z,t)=x^(z,t)+y^(z,t)

Due processi x^(t) e y^(t) sono uguali in media quadratica se vale la seguente condizione:

per ogni t, E{ ( x^(t) - y^(t) )^2 } = 0

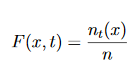
## Statistiche per processi





Un processo stocastico si può vedere come un insieme di n realizzazione, infatti se si fissa t si ottengono n valori dati dei diversi esperimenti.

La media di un processo stocastico si fa per il tempo su un realizzazione, da ciò si può calcolare la distribuzione di probabilità:



Quindi le normali statistiche non bastano e di conseguenza servono quelle del secondo ordine:

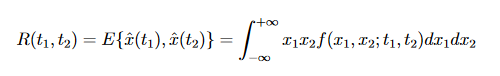




E’ anche possibile ottenere statistiche di ordine n, tutte queste valgono sia per i processi continui, sia per quelli discreti.

## Autocorrelazione

L’autocorrelazione si ottiene nel seguente modo:



Nel caso in cui t1=t2,



## Autocovarianza

L’autocovarianza si ottiene nel seguente modo:



I processi stocastici possono essere definiti con variabili complesse: dato un processo x^(t), il processo x\*(t) è detto complesso coniugato, ovvero un processo con la stessa parte reale ma con quella immaginaria negata.

Da ciò si può dire che:

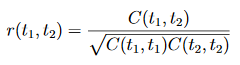








Il coefficiente di correlazione si calcola nel seguente modo:



## Processo centrato

Dato un processo x^(t), il processo centrato xp^(t) equivale al processo x^(t) meno la sua media:

xp^(t)=x^(t)-m(t)

Questo processo ha infatti media nulla.

## Cross-correlazione e cross-covarianza

Dati due processi x^(t) e y^(t), la cross-correlazione si calcola nel seguente modo:



La cross-covarianza si calcola nel seguente modo:



Due processi sono ortogonali quando la correlazione è nulla per ogni t1, t2, sono scorrelati se la covarianza è nulla sempre per ogni t1, t2.

## Processo di rumore bianco

Dato un processo v^(t) a media nulla, per ogni t1,t2 con t1!=t2, v^(t1) e v^t2) sono scorrelati.

Si può definire un processo di rumore bianco col concetto di indipendenza:

Dati v^(t1) e v^(t2), essi sono indipendenti per ogni t1, t2 con t1!=t2, di conseguenza sono anche scorrelati.

Questo processo è detto processo di rumore bianco in senso stretto.

## Incrementi di un processo

Dati t1<t2<t3<t4, si può costruire un incremento dato da x^(t2) - x^(t1).

Se questi incrementi sono scorrelati tra loro, allora per ogni t1<t2<t3<t4, x^(t) è un processo a incrementi scorrelati.

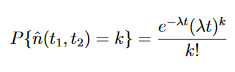
Si può dimostrare che il processo w^(t)= integral(v^(t)dt) è un processo a incrementi scorrelati, quindi dw^(t)=v^(t) indica che w^(t) è un processo di rumore bianco.

Se due processi x^(t1),...,x(tn) e y^(t1),...,y^(tn) sono tra loro indipendenti, allora x^(t) e y^(t) sono mutuamente indipendenti.

## Processo di Poisson

Si considera l’intervallo [t1,t2], si può quindi definire n^(t1,t2) data dal numero di punti ti che cadono nell’intervallo.

Questo numero è descrivibile come la probabilità di Poisson:



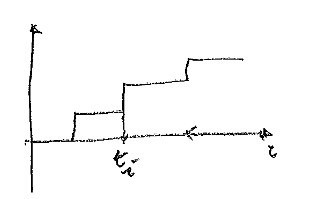
dove delta\*t è il parametro di distribuzione in cui:

* delta è l’ampiezza dei punti;
* t è la differenza tra t2 e t1.

Con due intervalli [t1,t2] e [t3,t4] non sovrapposti, il numero di ti che cade nel primo intervallo è indipendente da quello del secondo, quindi n^(t1,t2) è indipendente da n^(t3,t4).

Si può quindi costruire un processo x^(t)=n^(0,t) tale che:

* Ogni gradino è uno step, ovvero un punto che cade nell’intervallo;



* Il processo deve essere a stati discreti;
* il valore atteso di x^(t) deve essere delta\*t;

La cosa interessante è che x^(t) è a incrementi indipendenti.

## Processo normale

Un processo normale è un processo x^(t) tale che x^(t1),..., x^(tn) sono congiuntamente normali per ogni n e per ogni t1,...,tn.

Le statistiche del processo sono date dalla media m(t) e dalla covarianza C(t1,t2).

In termini di funzione caratteristica:



## Processi Stazionari

### Processo stazionario in senso stretto

Il processo x^(t) è stazionario in senso stretto se le proprietà statistiche non cambiano quando si sposta l’origine. In breve, x^(t) e x^(t+c) hanno le stesse statistiche.

Ciò vuol dire che:

per ogni costante c, f(x1,...,xn,t1,...,tn)=f(x1,...,xn,t1+c,...,tn+c)

Se si considera f(x,t)=f(x,t+c) in cui c è un qualunque valore, allora f(x,t)=f(x,t+c)=f(x).

Per quanto riguarda la densità f(x1,x2,t1,t2)=f(x1,x2,t1+c,t2+c) si può prendere c=-t2, quindi:

f(x1,x2,t1,t2)=f(x1,x2,tau) dove tau=t1-t2

La statistica del secondo ordine dipende solamente dal tempo trascorso tra t1 e t2, quindi:

E{x^(t)}=m

R(t1,t2)=integral(integral(x^(t1)\*x^(t2)\*f(x1,x2,tau)dx1)dx2)=R(tau)

Non è necessario avere statistiche troppo elevate, ci si può accontentare di quelle di primo e secondo ordine.

### Processo stazionario in senso ampio

Il processo x^(t) è stazionario in senso ampio se il suo valore atteso è m e la correlazione dipende solo da tau:

R(t,t+tau) = E{ x^(t+tau)\*x(t)}

## Sistemi a input stocastico

Dato un processo x^(t), si può ottenere y^(t)=T[ x^(t) ] in due modi:

* con T deterministico se agisce solo sul tempo;
* con T stocastico se agisce anche sulle variabili.

## Sistemi deterministici

### Sistemi senza memoria

Un sistema senza memoria è una funzione di x^(t). Se y^(t) = g(x^(t)) è senza memoria, vuol dire che y^(t1) dipende solo da x^(t1) e non dai precedenti.

Si possono quindi ricavare le statistiche nel seguente modo:

E{y^(t)} = integral( g(x) \* fx(x,t) dx)

E{y^(t1)\*y^(t2)} = integral(integral(g(x,t1)\*g(x,t2)\*fx(x1,x2,t1,t2)dx1)dx2)

Se il processo in input è stazionario (anche in senso stretto), allora lo è anche quello in output.

Per calcolare la densità di probabilità di y^(t), allora si fa un sistema in cui si scrivono le varie g(xi)=yi.

Se questo sistema ha soluzione, allora:

fy(y1,...,yn,t1,..,tn)=fx(x1,..,xn,t1,..,tn)/Jacob(g’(x1),...,g’(xn))

### Processi deterministici

Un processo deterministico è un processo y^[ x^(x) ], ad esempio: dato un sistema senza memoria g(x^(t)) e y>0, la densità di y^(t) = g( x^(t) ) si calcola nel seguente modo:

fy(y,t) = [ fx(sqrt(y),t) + fx’(-sqrt(y),t)] / ( 2\* sqrt(y))

Se g(x)= x^2, allora dati due valori y1 e y2:

y1=x1^2, y2=x2^2

La jacobiana sarà quindi jacob=+- 4\*Sqrt(y1\*y2) = +- 4\*x1\*x2, quindi:

fy(y1,y2,t1,t2)=sum(fx( +-sqrt(y1), +-sqrt(y2), t1, t2)/jacob

Se la densità di x^ è stazionaria allora:

fx(x,t)=fx(x) fx(x1,x2,t1,t2)=fx(x1,x2,tau) dove tau=t1-t2,

Se la densità di x^(t) è una distribuzione normale con varianza Rx(0), se vale y^(t) = (x^(t))^2), allora Il valore atteso di y^(t) è la correlazione di x^(t).

E{y^(t)} = E{(x^(t))^2)} =Rx(0)

Quindi:

fy(y)=e^(-y/(2\*Rx(0))/sqrt(2\*PI\*y\*Rx(0))

e di conseguenza:

fx(x)=e^(-x^2/(2\*sigma^2)/sqrt(2\*PI\*y\*sigma^2)

<grafici>

### Sistemi lineari

Dato un processo x^(t), y^(t)=L[ x^(t) ] è un sistema lineare.

I sistemi lineari hanno le seguenti proprietà:

* Linearità: L[ a1 \* x1^(t) + a2 \* x2^(t) ] =a1\*L[x1^(t)] + a2\*L[x2^(t) ];
* Invarianza nel tempo: per ogni c, x^(t+c) → y^(t+c)
* La risposta a un sistema lineare è una convoluzione:

y^(t)= convolution(x^(t),h(t)) = integral(x^(t)\*h(alpha) dalpha) dove h=L[delta(t)] è la risposta a un impulso.

Ciò significa che:

L[delta(t)] = convolution(delta(t), h(t))= h(t)

* Se l’input è normale, allora per linearità lo è anche l’output;
* Se x^(t) è stazionario in senso stretto, per invarianza temporale lo è anche y^(t) per il seguente motivo:

y^(t+c)=L[ x^(t+c)] tende a x^(t), che ha le stesse statistiche di x^(t+c).

#### Teorema

Il valore atteso di un sistema lineare è il sistema lineare di un valore atteso:

E{ L[x^(t)] } = L [ E{x^(t)} ]

Ciò implica che my[t] = L [ mx[t] ]

Vi sono due possibili interpretazioni:

* Il valore atteso di y^(t) è il valore atteso di una convoluzione:

E{y^(t)} = E{ convolution(x^(t),h(t)) } = mx[t]

* il processo y^(t,z) è il sistema lineare di x^(t,z):

E{y^(t)} = sum[i=0,n] ( y^(t,zi) )/n =

= sum[i=0,n]( L[ x^(t,zi) ]/n = L[ sum[i=0,n](x^(t,zi) )/n ] = L[ mx[t] ]

* Dato il processo centrato xp^(t) = x^(t) - mx[t], si può ottenere il processo centrato yp^(t) nel seguente modo:

yp^(t)=L[ xp^(t) ] = L[x^(t)] + L[mx[t]]

* Dato un processo x^(t) = f(t) + v^(t) ( dove v^(t) è a media nulla), il valore atteso di x^(t) è la somma dei valori attesi degli operandi:

E{x^(t)} = E{f(t)} + E{v^(t)} = E{f(t)} dato che v^(t) è a media nulla.

Quindi se y^(t)=L[ x^(t) ], allora:

my[t]= L[ mx[t] ] =convolution(f(t),h(t))

Per ottenere l’autocorrelazione Ryy(t1,t2) da Rxx(t1,t2), si eseguono due passaggi:

1. Si ricava la cross-correlazione:

Rxy(t1,t2)=L2[Rxx(t1,t2)]= convolution(Rxx(t1,t2),h(t))=

=integral(Rxx(t1,t2-alpha)\*h(alpha)dalpha)

1. Si ricava Ryy(t1,t2) da Rxy(t1,t2):

Ryy(t1,t2)=L1[ Rxy(t1,t2) ] = convolution(Rxy(t1,t2),h(t)) =

= integral(Rxy(t1+alpha,t2)\*h(alpha)dalpha)

* Nel caso in cui t=t2:

y^(t)\*y^(t2)= Lt[ x^(t)\*y^(t)]

E{ yì(t)\*y^(t2)} = Lt[ E{ x^(t)\*y^(t)} ]

Nel caso delle variabili complesse si utilizza il complesso coniugato:

Rxy(t1,t2)= convolution(Rxx(t1,t2),h\*(t2))

Ryy(t1,t2)=convolution(Rxy(t1,t29, h(t1))

Rxy(t1,t2)=E{ x^(t1) \* y\*(t2)} =E{ x^(t1) \* integral(x^(t+alpha \* h\*(alpha)dalpha)}=

=integral( Rxx(t1,t-alpha) \* h\*(alpha) dalpha=convolution(Rxx(t1,t2),h\*(t2))

## Caso particolare: Processo di rumore bianco con sistema lineare

<vedere dispense prof>

## Spettri di potenza

Un spettro di potenza è la trasformazione matematica di una correlazione, calcolabile nel seguente modo:

S(omega)= integral(R(tau) \* e^(i\*omega\*tau) dtau) =

= integral( R(tau) \* ( cos(omega\*tau) + i\*sin(omega\*tau))

In campo complesso tutto diventa come segue:

S\*(omega)=integral(R\*(tau) \* e^(i\*omega\*tau) dtau)=

=integral(R(-tau) \* e^(i\*omega\*tau) dtau)=

=integral(R(tau) \* e^(-i\*omega\*tau) dtau)=S(omega)

Allo stesso modo si può antitrasformare a partire da S(omega) per ottenere R(tau):

R(tau)=integral(S(omega)\*e^(i\*omega\*tau) domega)/(2\*PI)

Nel caso in cui il processo x^(t) sia reale, allora R(tau) è pari, quindi:

R(tau)=R(-tau) S(omega)=integral(R(tau)\*cos(omega\*tau)dtau)

dal momento che la funzione sin è dispari.

## Cross-Spettro

Il cross-spettro è la trasformata di una cross-correlazione:

Sxy(omega)=integral(Rxy(tau)\*e^(i\*omega\*tau) dtau)

Syx\*(omega)=integral(Ryx\*(tau)\*e^(i\*omega\*tau)dtau)=

=integral(Rxy(-tau)\*e^(i\*omega\*tau)dtau)=Sxy

Come per la cross correlazione, per ottenere Syy da Sxx si effettuano due passaggi:

1. Sxy(omega)=Sxx(omega)\* H\*(omega)
2. Syy(omega)=Sxy(omega)\*H(omega)

dove H(omega) è la trasformata di h(tau)

<continua sulle dispense>

### Media mobile

La media mobile è un metodo utilizzato in analisi dati in cui, dato un processo x^(t), si ricava y^(t) nel seguente modo:

y^(t)=integral[t-T,t+T](x^(alpha)dalpha)/(2\*T)

Tutto ciò si può considerare come un sistema lineare,quindi:

Dato che integral[a,b] ( delta(t) dt) = 1 se a<0<b e 0 altrimenti, si può dire che

y^(t)=integral[t-T,t+T](delta(alpha) dalpha)/(2\*T)

Si può quindi calcolare H(omega):

H(omega)=transform(h(t)) = integral[-T,T](e^(-i\*omega\*t) dt)=sin(omega\*T)/(omega+T)

Quindi:

Syy(omega)=Sxx(omega)\*|H(omega)|^2 = Sxx(omega)\*sin(omega\*T)/(omega\*T)

Nel campo delle frequenze la media mobile funzione come un filtro passa-basso.

Esempio:<dispense>

## Processi discreti

Un processo discreto x^[n] è un processo a tempo discreto n.

La media si calcola come segue:

m[n]=E{x^[n]}

La correlazione e la covarianza vengono calcolate nel seguente modo:

R[n1,n2]=E{x^[n1]\*x^[n2]} //Si usa il complesso coniugato nel secondo in campo complesso

C[n1,n2]=R[n1,n2]-m[n1]\*m[n2]

Il processo x^[n] è stazionario in senso stretto se invariante al variare dell’origine dei tempi, lo è in senso ampio se la media è costante e R[n+m,n]=E{x^[n+m,n]}=R[m]

Il processo x^[n] è un processo di rumore bianco in senso stretto se tutti gli x^[ni] sono indipendenti, è invece un processo di rumore bianco normale se tutte le x^[ni] sono scorrelate.

<dispense>

## Random Walk

Dato un processo x^(n\*T) equivalente alla somma di n esperimenti S, ogni xi^ del processo può assumere i valori +S o -S, quindi:

* Si lancia una moneta: se esce testa allora xi^ = +S, altrimenti fa -S;

Se su n lanci sono uscite k teste:

x^(n\*T)=k\*S + (n-k)\*(-S) = k\*S - (n - k)\*S =k\*S - n\*S + k\*S= 2\*k\*S -n\*S= (2\*k - n)\*S=m\*S

con m=(2\*k - n)

Quindi:

P{x^(n\*T)=m\*S}=1/(2^n)

Le statistiche dei singoli xi^ sono le seguenti:

E{xi^}=p\*s \*q\*(-s)=0

E{xi^2}= (p\*s)^2 \* (q\*(-s))^2 = s^2

Da esse si può calcolare le statistiche di x^(n\*T):

E{x^(n\*T)}=0

E{(x^(n\*T))^2}=n\*s^2

COl teorema del limite centrale, si può calcolare la distribuzione binomiale da quella normale (con un n grande a sufficienza):

(n k)\*p^k \*q^(n-k) = e^((k-n\*p)^2/(2\*n\*p\*q))/sqrt(2\*PI\*n\*p\*q)=

=e^(-m^2/(2\*n))/sqrt(n\*PI/2)

Quindi P{x^(n\*T)=m\*S} = e^(-m^2/(2\*n))/sqrt(n\*PI/2)

Questa distribuzione può essere scritta come G(m/sqrt(n)) =P{x^(t)<=m\*S} dove t è nell’intervallo [n\*T-T,n\*T].

Se si prendono più incrementi successivi, questi sono indipendenti tra loro.

Dati n1<n2<n3<n4, gli incrementi sono i seguenti:

[x^(n4\*T)-x^(n3\*T)], [x^(n3\*T)-x^(n2\*T)], [x^(n2\*T)-x^(n1\*T)]

Con m tendente a sqrt(n) e n tendente a infinito, T tende a 0,quindi:

E{(x^(t))^2}=n\*s^2=t\*s^2/T con t=n\*T e s^2=alpha\*T, quindi:

E{(x^(t))^2}=alpha\*t

Con T che tende a 0, il processo w^(t)=lim[T→ 0](x^(t), quindi:

f(w,t)=e^(-w^2/(2\*alpha\*t))/sqrt(2\*PI\*alpha\*t) dove w=m\*s

Quindi: P{w^(t) <=w} =G(w/sqrt(alpha\*t)) //Processo di Wiener

R(t1,t2)= alpha\*min(t1,t2)

### Dimostrazione

Supponendo che t<t2, incremento [w^(t2) - w^(t1) ] è indipendente da w^(t1), quindi:

E{ [w^(t2) - w^(t1) ] \* w^(t1) } =0

E{ [w^(t2) - w^(t1) ] }\* E{w^(t1) } =0

E{ w^(t2) w^(t1) - (w^(t1))^2 }=0

E{ w^(t2) w^(t1) } =E{(w^(t1))^2 } = alpha\*t1

## Moto Browniano

Ritornando a parlare del moto browniano, si possono trattare le equazioni differenziali come sistemi lineari:

M\*x^’’(t)+f\*x^’(t) +c\*x^(t)=F^(t)

dove f è il coefficiente di attrito e F^(t) è un processo di rumore bianco gaussiano a media nulla.

Si può definire lo spettro SF(omega)=2\*kB\*T\*f

Per determinare le statistiche, bisogna tenere conto dei moti delle particelle.

### Moto legato

Si sa che |H(omega)|^2 = 1/((c-m\*omega^2) + (f\*omega)^2)

Si può quindi riscrivere tutto nel seguente modo:

h’’(t) + f\*h’(t)/m +c\*h(t)/m = delta(t)

Quindi:

|H(omega)|^2 = (1/m)^2/((c-m\*omega^2) + (f\*omega)^2)

Sx(omega)=SF(omega)\*|H(omega)|^2 =(2\*kB\*T\*f)//((c-m\*omega^2) + (f\*omega)^2)

Si può quindi ottenere Rx(tau) → Rx(0)=kB\*T/c

e quindi:

fx(x)=e^(-c\*x^2/(2\*kB\*T))\*sqrt(c/(2\*PI\*kB\*T))

### Moto Libero

Alla precedente equazione differenziale si pone c=0, quindi:

m\*x^’’(t) +f\*x^’(t)=F^(t)

Quindi, data v^(t)=x^’(t), si ha che:

m\*v^’(t) +f\*v^’(t)=F^(t)

Si può riscrivere l’equazione nel seguente modo:

h’(t)+f\*h(t)/m = delta(t)/m

quindi:

|H(omega)|^2 = 1/((m\*omega)^2 + f^2)

Sv(omega)=SF(omega)\*|H(omega)|^2= (-2\*kB\*T\*f)/((m\*omega)^2 + f^2)

Rv(tau)=(kB/m)\*e^(-|tau|/m/f)

La variabile v^(t) indica la velocità della particella a tempo t, per trovare la posizione si integra:

x^(t)=integral(v^(t)dt)

E{(x^(t))^2} = 2\*kB\*T\*(t - (m/f) - (m/f)\*e^((f\*t)/m))

Quindi con t>> m/f:

E{(x^(t))^2} = 2\*kB\*T\*t/f=2\*D^2\*t

Il processo è quindi diffusivo per tempi t lunghi.

Nel caso in cui c=0 e m\*x^’’(t)<<f\*x^’(t), si ha che f >> m/t, quindi:

f\*x^(t)=F^(t) x^(t)=integral(F^(alpha) dalpha)/f

Se F è un processo di rumore bianco con SF(omega)=2\*kB\*T\*f

Quindi:

C(t1,t2)= q(t1)\*delta(t1-t2)

E{(x^(t))^2} = alpha\*t con alpha=2\*Kb\*T/f

f(x)=e^(-(x^2)/(2\*alpha\*t))/sqrt(2\*PI\*alpha\*t)

Nel caso in cui gli incrementi sono indipendenti, dati t1>t2>t3>t4, il valore atteso del prodotto di due incrementi è uguale a 0:

E{[ x^(t4) - x^(t3)] \* [x^(t2) - x^(t1)]} = 0

## Probabilità condizionata tra processi

f(x|x^(t)=x0) = e^((x-x0)^2/(2\*alpha\*(t-t0))/sqrt(f\*PI\*alpha\*(t-t0))

Questo valore equivale alla probabilità di transizione da x0 a x.

<continua su dispense>

## Processi Markoviani

I processi markoviani sono particolari processi stocastici in cui la probabilità a tempo t dipende solo da quella a tempo t-1:

P{x^(t)=xn | x^(t) t<=tn-1} = P{x^(t)=x | x^(t-1)}

Quindi con t1<t2<...<tn

P{x^(tn)<=xn|x^(tn-1),...,x^(t1)} = P{x^(tn)<=xn|x^(tn-1)}

### Proprietà generali

* f(xn| xn-1,...,x1) = f(xn|xn-1)
* f(x1,...,xn)=f(xn,xn-1)\*f(xn-1|xn-2)\*...\*f(x2|x1)\*f(x1)

Quest’ultima proprietà viene anche detta catena di Markov in quanto calcola le probabilità sapendo solo lo stato precedente.

Se la prima proprietà vale, il processo è detto markoviano, questo per il seguente motivo:

f(xn|xn-1,...,x1)=f(x1,...,xn)/f(x1,...,xn-1) = f(xn|xn-1)

Dato un processo xn^ tale che xn^+1 - a(xn,n)=vn^ ( un processo di rumore bianco in senso stretto), ciò soddisfa la condizione di Markov:

Se x0=0, v0=0 e a(xn,n)=xn, allora xn=xn-1 + vn-1 =sum[i=1,n-1](vi)

Tutte le vi sono quindi indipendenti e la loro somma è la successione di Markov.

## Processi omogenei

Quando un processo è stazionario:

f(xn)=f(xn|xn-1)

f(x1,x2)=f(x2|x1)\*f(x1)

Un processo è omogeneo se f(xk|xk-1) non dipende da spostamenti nell’origine dei tempi mentre f(xk) può dipendere dal tempo k.

## Equazione di Chapman - Kolmogorov

Dati i tempi n>m>k, per andare da k a n si fa come segue:

f(xn|xk)= integral(f(xn|xm)\*f(xm|xk)dxm)

In parole semplici si va da k a n passando per tutti gli stati m.

### Notazione

xn è un processo di Markov con tempo e stati discreti.

pi[n]=P{ xn = ai} ( dove ai è lo stato i-esimo) è la probabilità a tempo n di essere nello stato ai.

PIij[n1,n2] = P{xn2=aj|xn1=ai}

è la probabilità di essere in aj a tempo n2 sapendo che a tempo n1 si era in ai.

Normalizzazione: la somma delle probabilità degli stati finali deve essere uguale a 1.

sum[j](PIij[n1,n2]=1

sum[i](pi[k]\*PIij[k,n]=pj[n]

Nel caso discreto:

PIij[n1,n2]=sum[t]PIir[n1,nr]\*PIrj[nr,n3]

## Processo Markoviano discreto e omogeneo

Nel caso in cui il processo markoviano sia discreto e omogeneo, si ha che PIij[n1,n2] è omogeneo e quindi si può definire m=n2-n1 tale che:

PIij[m]=P{xn+m =aj| xn =ai}

Nel caso in cui k=n2-n1 e n=n3-n2, si ha che:

PIij[n+k]=sum[r](PIir[k]\*PIrj[m])

tutto ciò si può riscrivere in forma di matrice e graficamente nel seguente modo:

<dispense>

Dal momento che vale la proprietà di normalizzazione, la somma degli archi deve essere 1, quindi è possibile determinare n-1 probabilità e calcolare n-esima sul momento.

Se si pone k=1, si ha che PI[n+1]=PI[1]\*PI[n]=PI[1]\*PI[1]\*PI[n-1], quindi si pul definire PI=PI[1] e PI[n]= PI^n.

In termini di stati, la probabilità è di questi ultimi è un vettore mentre in termini di matrici:

P[n]=P[n-1]\*PI=P[0]\*PI^n dove n indica i tempi e P[n] è un vettore di pi[n].

Se P[1]=P è il vettore degli stati iniziali, per omogeneità si può dire che P[2]=n e in generale P[n]=P.

Si può quindi concludere che P\*PI=P, quindi le probabilità degli stati non cambiano nel tempo e la loro somma fa 1.

## Esempio: Random walk con due pareti riflettenti

Questo processo markoviano è omogeneo in cui:

* o può andare in s oppure -s con provabilità 0.5;
* con la stessa probabilità, s e -s possono andare rispettivamente in +2s e -2s oppure tornare in o;
* Dal momento che non esiste un +-3s, da +2s e -2s si torna rispettivamente in +s e -s con probabilità 1.

<dispense>

## Equazione Forward

<dispense>