

Prof. Dr. Ottmar Beucher

Il Teorema del Limite Centrale

Lezione
all'Università degli Studi di Parma
2009

Buongiorno!

Mi chiamo Ottmar Beucher, sono Professore di matematica ed di teoria dei segnali all'università delle Scienze Applicate a Karlsruhe in Germania.

Dal 2007 c'è una cooperazione Erasmus fra il Dipartimento di Ingegneria della vostra università e il Dipartimento di Costruzione Meccanica et Meccatronica della mia università e oltre il cambio di studenti è anche previsto un cambio di docenti.

Dopo che sono già stati a Karlsruhe due colleghi italiani, i professori Cagnoni e Guarino Lo Bianco, io sono molto lieto oggi di aver l'occasione di dare una lezione a Parma e tengo a ringraziare il professore Bononi di avermi dato la possibilità di parlare sul tema del „Teorema del Limite Centrale“ nel ambito del suo corso di probabilità.

Tengo a dare il corso in Italiano e spero che le mie conoscenze della vostra lingua bastino almeno per farmi capire, se no vi prego di segnalarmelo.

Introduzione

La probabilità è una branca delle scienze matematiche che cerca a spiegare fenomeni casuali e risultati di esperimenti casuali.

Per questa ragione ho pensato che fosse una buona idea di motivare il tema della presente lezione con l'aiuto di un esperimento casuale.

Ovviamente non volevo e non potevo portare dalla Germania attrezzature complicate e così sono limitato ad esperimenti che si possono effettuare col computer e dei mezzi che permettono di salvarci direttamente i risultati.

Esperimenti

Così vi propongo due esperimenti acustici ed un esperimento „virtuale“ basata su una simulazione che ho trovato nella rete.

In tutti e due i casi lo scopo sarà l'analisi di un rumore nel senso che esamineremo la ripartizione della sonorità o più precisamente la ripartizione della *pressione del suono* perché è proprio questo che stiamo misurando quando registriamo un suono con un microfono.

Il rumore di un getto d'acqua

In un primo esperimento guardiamo il rumore che fa l'acqua quando esce di un rubinetto e cade nel lavandino.

Siccome non sapevo di sicuro che fosse possibile effettuare l'esperimento direttamente in questa sala di lezione ho filmato questo esperimento.

La seguente foto e il film mostrano l'impostazione dell'esperimento.

Il suono viene registrato da un microfono ed salvato numericamente con l'aiuto di un piccolo programma di MATLAB.

Slide 1

Film
acqua.MOD

programma
make_audio.m



Grafica 1: Impostazione dell'esperimento „getto d'acqua“

Ascoltiamo quello che è stato registrato.

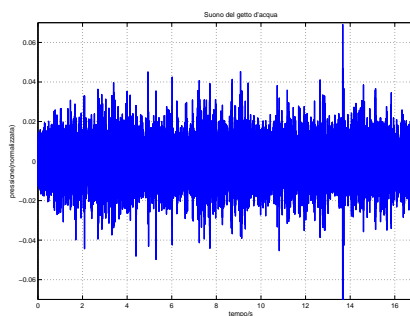
Quanto al suono ci interessiamo alla ripartizione della pressione *momentanea* del suono. Con l'aiuto di MATLAB possiamo visualizzare il suono in una grafica (2).

ascolto del
suono
acqua1A10.wav

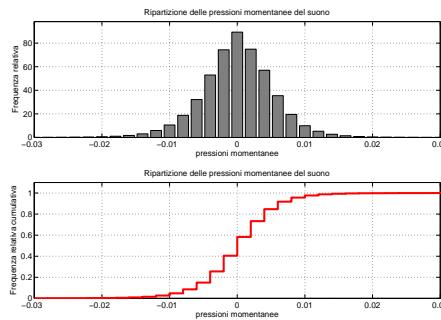
calcolo della
grafica
(calgrafica.m)

Si vede che si tratta di un suono molto irregolare un cosiddetto rumore bianco. Gli amplitudini non rappresentano i valori veri della pressione del suono perché il segnale viene normalizzato durante la registrazione. Però per il fenomeno che vogliamo studiare questo non importa.

La ripartizione delle amplitudini si lascia calcolare facilmente con l'aiuto di un piccolo programma di MATLAB (Grafica 3).



Grafica 2: Suono „getto d'acqua“ registrato



Grafica 3: Ripartizione delle pressioni momentanee del suono „getto d'acqua“

Si riconosce facilmente una ripartizione ben nota. Le grafiche ovviamente somigliano a la funzione densità e la funzione cumulativa della distribuzione Gaussiana!

Un vecchio strumento di calibrazione acustico

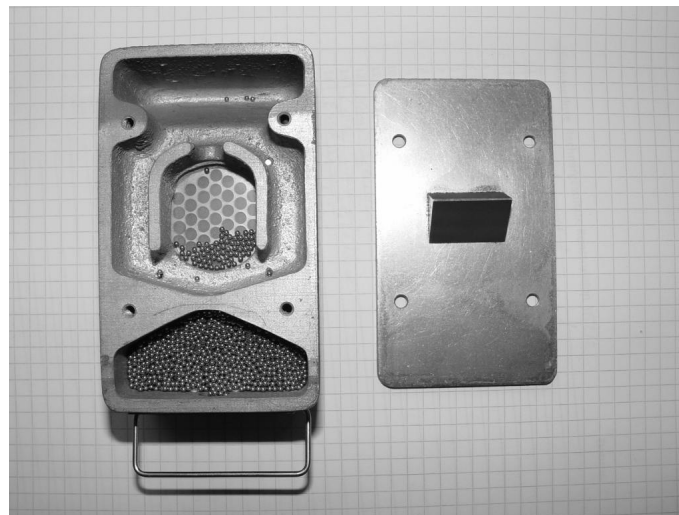
Il secondo esperimento lo posso effettuare direttamente in sala perché potevo portare lo strumento necessario.

Si tratta di questo vecchio ma nonostante geniale strumento per calibrare la sonorità.

Prima ve lo faccio sentire.

Questo strumento fa un rumore detto „bianco“ di esattamente 127 phon.

Vediamo come lo fa. Il principio di funzionamento è semplice e si può vedere su questa foto (Grafica 4).



Grafica 4: Lo strumento di calibrazione acustico

All'interno dello strumento si trova un magazzino che contiene delle palline di ferro. Se giro lo strumento queste palline centrano la rampa che vedete nella parte destra della foto. La funzione della rampa è semplicemente di deviare le palline affinché centrino una membrana che finalmente si mette a vibrare e fare il rumore che avete sentito.

Proviamo adesso a registrare il suono. In questo caso è un po' più complicato registrarlo perché il suono si sente soltanto pochi secondi. Dunque per avere un campionamento paragonabile a quello ottenuto per il getto d'acqua facciamo l'esperimento alcune volte.

Verifichiamo adesso quello che abbiamo appena registrato e ascoltiamo il suono.

Ancora una volta calcoliamo la ripartizione della pressione *momentanea* del suono registrato et ancora una volta visualizziamo il suono in una grafica con l'aiuto di MATLAB (Grafica 5).

esperimento
col strumento
acustico

slide 2

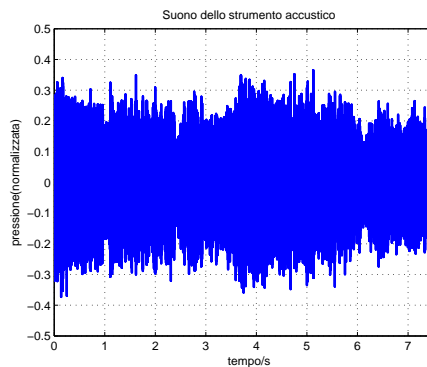
registrazione
del suono
con
make_audio.m

ascolto
del suono

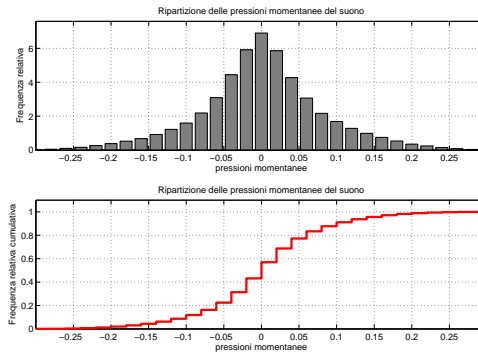
Anche in questo caso si tratta di un segnale tipo „rumore bianco“.

La ripartizione delle amplitudini momentanee si calcola con l'aiuto di un programma di MATLAB simile a quello che abbiamo utilizzato nell'ultimo esperimento (Grafica 6).

calcolo della
grafica
(calgrafstrto.m)



Grafica 5: Suono dello strumento di calibrazione acustico



Grafica 6: Ripartizione delle pressioni momentanee del suono dello strumento di calibrazione acustico

Si nota lo stesso risultato. Le grafiche somigliano di nuovo a la funzione densità di probabilità (PDF) e la funzione cumulativa di distribuzione (CDF) della distribuzione Gaussiana!

Un esperimento „virtuale“ - La macchina di Galton

L'ultimo esperimento è infatti un esperimento famosissimo di cui potete trovare realizzazioni addirittura in alcuni musei specializzati alle scienze naturali come il *Museo della Matematica* a Roma.

slide 3

Questa foto (Grafica 7) rappresenta la cosiddetta *Macchina di Galton*.

Si tratta di una costruzione in cui delle palline escono da un buco e cadono in recipienti posati al di sotto del buco attraversando una rete di chiodi.

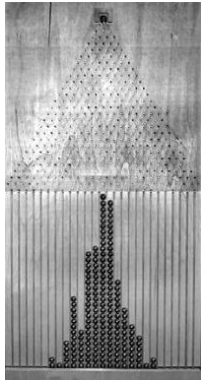
Questi chiodi sono sistemati in modo tale che i pallini possano appena passare fra due chiodi di una riga e poi cadano esattamente su un chiodo della riga seguente. Così in ogni riga le palline vengono a caso deviate sia a destra sia a sinistra con la stessa probabilità.

La macchina di Galton è uno strumento che potrebbe facilmente essere realizzato da una persona artigianalmente dotata.

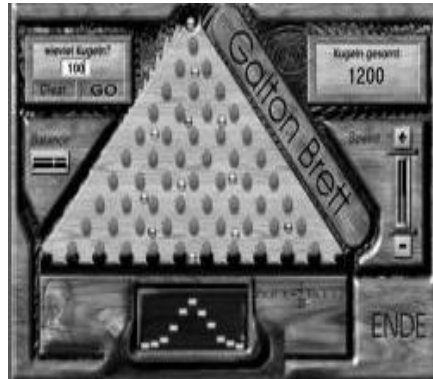
Siccome io non sono per niente dotato per fare bricolage ho rinunciato a provarlo per non far ridere i polli. Però essendo più dotato per trovare delle cose sull'internet vi ho portato un piccolo programma di simulazione (Grafica 8).

simulazione
galton.exe

Si vede come i pallini vengono deviate. Nella parte inferiore vengono contati i numeri delle palline che cadono in ogni recipiente.



Grafica 7: La macchina di Galton del Museo della Matematica a Roma



Grafica 8: La macchina di Galton - Programma di simulazione

Si nota anche in questo caso che la grafica sviluppandosi al do sotto del pannello dopo un certo tempo comincia a somigliare a la funzione densità di probabilità (PDF) della distribuzione Gaussiana!

Caratteristiche comuni degli esperimenti

Abbiamo appena visti tre esperimenti casuali differenti in cui la distribuzione della grandezza osservata era simile alla distribuzione Gaussiana! Certo i due primi esperimenti non sono tanto differenti perché si tratta di un suono in entrambi i casi.

Però bisogna chiedersi se il fatto che seguono la stessa distribuzione sia basato su caratteristiche comuni.

Ovviamente il rumore che viene prodotto dallo strumento di calibrazione acustico è *la somma* dei rumori prodotti da ciascuna delle palline.

Anche il rumore dell'acqua si lascia concepire come una *sovrapposizione* di tanti effetti acustici. In effetti la seguente rappresentazione (Grafica 9) in cui vediamo una fotografia istantanea del getto d'acqua ci dimostra che il getto consiste di una varietà di gocce singole.

Siccome ogni goccia fa un rumore quando cade nel lavandino il rumore che ascoltiamo è infatti *la somma* dei rumori prodotti da ciascuna delle gocce.

Se definiamo nel linguaggio matematico la pressione momentanea del suono una variabile aleatoria X e se chiamiamo X_k la pressione momentanea di una singola pallina rispettivamente di una singola goccia k , possiamo allora in entrambi i casi concepire la pressione momentanea del suono come una somma di variabili aleatorie di cui le caratteristiche statistiche (vuol dire le distribuzioni) sono uguali oppure approssimativamente uguali.



Grafica 9: Fotografia istantanea dell'getto d'acqua

Modello matematico dei rumori

X : la variabile aleatoria della pressione momentanea del suono,

lavagna

X_k : la pressione momentanea di una singola pallina (goccia) k .

$$X = \sum_{k=1}^N X_k, \quad N \gg 1. \quad (6.1)$$

Anche l'esperimento famoso della macchina di Galton si lascia vedere così.

Se definiamo la distanza orizzontale dal centro (che è la posizione del buco) del recipiente in cui cade la pallina il valore di una variabile aleatoria X e se definiamo la deviazione orizzontale suscitata da un chiodo nella riga k il valore di una variabile aleatoria X_k , si vede bene che la distanza finale X è la somma delle deviazioni X_k , dunque X è la somma di un gran numero di variabili aleatorie come negli esempi precedenti.

Modello matematico della macchina di Galton:

X_k : la deviazione orizzontale suscitata da un chiodo nella riga k :

lavagna

$$X_k = \begin{cases} 1 & \text{se pallina cade a destra,} \\ -1 & \text{se pallina cade a sinistra.} \end{cases} \quad (6.2)$$

Funzione Massa di Probabilità:

$$f_{X_k}(n) = \begin{cases} 0.5 & \text{se } n = 1, \\ 0.5 & \text{se } n = -1 \\ 0 & \text{altrimenti.} \end{cases} \quad (6.3)$$

X : la distanza orizzontale dal centro (posizione del buco) del recipiente finale:

$$X = \sum_{k=1}^N X_k, \quad N \gg 1. \quad (7.1)$$

Nel caso della macchina di Galton abbiamo inoltre il vantaggio che conosciamo bene le distribuzioni delle variabili X_k . Questi variabili sono identicamente distribuite e hanno una Funzione Massa di Probabilità (FMP) ben conosciuta: la distribuzione discreta di Bernoulli.

Infatti anche la somma è una variabile discreta, però l'esperimento ci lascia sperare che la distribuzione della somma possa essere approssimata da una distribuzione Gaussiana.

Vedremo più tardi che, in effetti, abbiamo visto in modo sperimentale il risultato di un caso speciale del teorema generale che vogliamo dimostrare oggi.

Nei due casi degli esperimenti acustici non sappiamo di preciso a quale distribuzione segue la pressione del suono suscitato da *una sola pallina di ferro* rispettivamente da *una sola goccia d'acqua*. Però possiamo almeno presupporre che queste distribuzioni, anche se non sono perfettamente uguali, hanno caratteristiche paragonabili e non sono così differenti l'una l'altra.

Vediamo infatti più tardi che la presupposizione di un'uguaglianza degli distribuzioni che partecipano alla somma non è davvero per forza necessario anche se nella dimostrazione del risultato principale lo richiederemo per facilitare le cose.

Ricapitolazione di Nozioni e Risultati

Abbiamo visto con gli esempi precedenti che possiamo concepire i processi aleatori osservati come una *somma* di variabili aleatorie con caratteristiche simili. Si può anche ammettere negli esempi che queste variabili sono statisticamente indipendenti.

Dunque in un'analisi matematico dei fenomeni osservati si tratterà di studiare delle somme di variabili aleatorie indipendenti e le loro distribuzioni.

Avete già visto nel corso un metodo molto elegante che permette di calcolare distribuzioni di somme di variabili aleatorie indipendenti che era l'uso delle *funzioni generatrici dei momenti (MFG)* oppure delle *funzioni caratteristiche (CF)*.

Funzioni gener. dei momenti (MFG) e funzioni caratteristiche (CF)

Ripetiamo brevemente le definizioni di queste nozioni:

lavagna

Sia X una variabile aleatoria continua con una funzione densità di probabilità $f_X(x)$. Allora chiamiamo

$$\phi(s) = \mathbb{E}(e^{sX}) = \int_{-\infty}^{\infty} f_X(x) \cdot e^{sx} dx, \quad s \in \mathbb{C} \quad (8.1)$$

slide 5

la sua *funzione generatrice dei momenti (MFG)*.
Invece chiamiamo

$$\Phi(\omega) = \mathbb{E}(e^{j\omega X}) = \int_{-\infty}^{\infty} f_X(x) \cdot e^{j\omega x} dx, \quad \omega \in \mathbb{R} \quad (8.2)$$

la sua *funzione caratteristica (FC)*.

La funzione generatrice dei momenti (MFG) corrisponde infatti essenzialmente alla *trasformata di Laplace* che normalmente non esiste su tutto il piano complesso bensì soltanto su una strisciata verticale di \mathbb{C} .

La funzione caratteristica (CF) corrisponde essenzialmente alla *trasformata di Fourier* che esiste per tutti i valori $\omega \in \mathbb{R}$, se l'integrale del valore assoluto

$$\int_{-\infty}^{\infty} |f_X(x)| dx \quad (8.3)$$

esiste. Infatti le definizioni di queste trasformazioni sono leggermente differenti di quelle date in (8.1) e (8.2) però le proprietà rimangono le stesse.

Siccome nel nostro caso si tratta della trasformazione di una funzione densità di probabilità per cui

$$\int_{-\infty}^{\infty} |f_X(x)| dx = \int_{-\infty}^{\infty} f_X(x) dx = 1, \quad (8.4)$$

la funzione caratteristica (CF) *esiste per ogni* $\omega \in \mathbb{R}$.

Inoltre se l'integrale $\int_{-\infty}^{\infty} x \cdot f_X(x) dx$ converge assolutamente ed uniformemente, cioè se *esiste il primo momento*, la funzione caratteristica (CF) è *derivabile* per ogni $\omega \in \mathbb{R}$ e la derivata si lascia calcolare cambiando derivazione ed integrale:

$$\begin{aligned}
\frac{d}{d\omega} \Phi(\omega) &= \frac{d}{d\omega} \mathbb{E}(e^{j\omega X}) \\
&= \int_{-\infty}^{\infty} f_X(x) \cdot \frac{d}{d\omega} e^{j\omega x} dx \\
&= j \cdot \int_{-\infty}^{\infty} x \cdot f_X(x) \cdot e^{j\omega x} dx \quad \forall \omega \in \mathbb{R}.
\end{aligned} \tag{9.1}$$

lavagna

Di questa equazione si deduca un'identità importantissima:

$$\frac{d}{d\omega} \Phi(0) = j \cdot \int_{-\infty}^{\infty} x \cdot f_X(x) dx = j \cdot \mathbb{E}(X) \tag{9.2}$$

In generale, se esiste l'n-esimo momento $\mathbb{E}(X^n)$, possiamo dedurre nello stesso modo:

$$\frac{d^n}{d\omega^n} \Phi(0) = (j)^n \cdot \mathbb{E}(X^n). \tag{9.3}$$

slide 6

Dunque i momenti si lasciano calcolare a partire dalla funzione caratteristica e le sue derivate.

Inoltre calcoliamo per una variabile $Y = \lambda \cdot X$ graduate da un fattore $\lambda \in \mathbb{R}$ che

$$\Phi_Y(\omega) = \mathbb{E}(e^{j\omega Y}) = \mathbb{E}(e^{j(\omega \cdot \lambda) X}) = \Phi_X(\lambda \cdot \omega). \tag{9.4}$$

slide 6

Per una variabile $Y = X + b$ traslate da una costante $b \in \mathbb{R}$ deduciamo dalla definizione e dalle proprietà del valore atteso:

$$\Phi_Y(\omega) = \mathbb{E}(e^{j\omega Y}) = \mathbb{E}(e^{j\omega(X+b)}) = \mathbb{E}(e^{j\omega b} \cdot e^{j\omega X}) = e^{j\omega b} \cdot \Phi_X(\omega). \tag{9.5}$$

slide 6

Funzioni caratteristiche (CF) di somme di VA indipendenti

Quanto alle somme di variabili *indipendenti* ci interessa un'altra proprietà importante delle funzioni generatrici dei momenti (MFG) oppure delle funzioni caratteristiche (CF).

Per prima cosa ripetiamo che le variabili aleatorie

$$g_i(X_i), \quad i = 1, \dots, n \quad (10.1)$$

sono indipendenti, se le $g_i(x)$ sono qualunque funzioni reali oppure a valori complessi e se le variabili aleatorie X_i sono indipendenti.

lavagna

Se guardiamo specialmente le funzioni

$$g_i(x) = e^{j\omega \cdot x}, \quad i = 1, \dots, n, \quad x \in \mathbb{R} \quad (10.2)$$

ne deduciamo che le variabili

$$Y_i := e^{j\omega \cdot X_i} \quad (10.3)$$

sono indipendenti, se variabili aleatorie X_i lo sono.

Inoltre sappiamo già che il valore atteso di un prodotto di variabili aleatorie indipendenti è il prodotto delle valori attesi delle variabili:

$$\mathbb{E} \left(\prod_{i=1}^n Y_i \right) = \prod_{i=1}^n \mathbb{E} (Y_i) \quad (10.4)$$

di cui deduciamo per la funzione caratteristica (CF) di una **somma** $Z = X_1 + \dots + X_n$ **di variabili indipendenti**:

slide 7

$$\begin{aligned} \Phi_Z(\omega) &= \mathbb{E} (e^{j\omega Z}) = \mathbb{E} \left(e^{j\omega \left[\sum_{i=1}^n X_i \right]} \right) \\ &= \mathbb{E} \left(\prod_{i=1}^n e^{j\omega X_i} \right) = \prod_{i=1}^n \mathbb{E} (e^{j\omega X_i}) \\ &= \prod_{i=1}^n \Phi_{X_i}(\omega). \end{aligned} \quad (10.5)$$

Dunque la funzione caratteristica (CF) di una **somma di variabili indipendenti** è il **prodotto delle funzioni caratteristiche (CF)** degli addendi.

Nel caso speciale che tutte le variabili sono identicamente distribuite, quindi avendo una CF comune $\Phi_X(\omega)$, abbiamo allora:

slide 7

$$\Phi_Z(\omega) = \Phi_X^n(\omega). \quad (10.6)$$

Funzioni caratteristiche (CF) e funzioni cumulative (CDF)

Gli esempi precedenti ci suggeriscono che le distribuzioni di una somma di variabili aleatorie indipendenti siano molto vicino alla distribuzione Gaussiana.

L'ipotesi si impone che questa somiglianza è più precisa più variabili vengono sommate. Nel senso matematico stiamo allora studiando *la convergenza* delle distribuzioni delle somme.

Prima di continuare dobbiamo dunque definire di preciso che cos'è questa convergenza. Definiamo allora:

Una successione $F_n(x)$ di funzioni cumulative di distribuzione (CDF) *converge*, se esiste una funzione cumulativa di distribuzione (CDF) $F(x)$ per cui *in ogni punto di continuità* si ha

$$\lim_{n \rightarrow \infty} F_n(x) = F(x) \quad (11.1)$$

slide 8

Ovviamente avremmo anche potuto definire la convergenza di una successione di *funzioni densità di probabilità (PDF)* però preferiamo le funzioni cumulative di distribuzione (CDF) prima perché questa nozione è più generale sapendo che esiste per variabili continue e discrete e in secondo luogo perché è la nozione adeguata per quello che vogliamo dimostrare.

Abbiamo già detto che nella dimostrazione del teorema che stiamo per formulare le funzioni caratteristiche (CF) giocheremo un ruolo cruciale.

Con l'aiuto di funzioni caratteristiche (CF) si lascia studiare in un modo univoco la convergenza di successioni di funzioni cumulative di distribuzione (CDF).

Il risultato su di cui ci baseremo in quello che segue è il *teorema di LÉVY-CRAMÈR* che dice:

Sia X_n una successione di variabili aleatorie e siano $F_n(x)$ e $\Phi_n(\omega)$ le successioni corrispondenti delle funzioni cumulative di distribuzione (CDF) e delle funzioni caratteristiche (CF) di X_n .

Allora la successione $F_n(x)$ *converge* verso una funzione cumulativa di distribuzione $F(x)$ *se e solo se* la successione delle funzioni caratteristiche (CF) $\Phi_n(\omega)$ di X_n *converge uniformemente* verso una funzione $\Phi(\omega)$ in un intervallo $[-\omega_0, +\omega_0]$ per un $\omega_0 \in \mathbb{R}$.

In questo caso la funzione $\Phi(\omega)$ è la funzione caratteristica (CF) della funzione cumulativa di distribuzione $F(x)$ e la convergenza di $\Phi_n(\omega)$ verso $\Phi(\omega)$ è *uniforme* in ogni intervallo *finito* dell'asse reale.

slide 9

Questo risultato ci dà la giustificazione di studiare la convergenza delle funzioni cumulative di distribuzione (CDF) delle nostre somme di variabili aleatorie con l'aiuto delle sue funzioni caratteristiche.

Abbiamo adesso tutti gli ingredienti per preparare „il pasto“!

Il Teorema del Limite Centrale

Quando parliamo del *Teorema del Limite Centrale* parliamo infatti di tutta una famiglia di teoremi.

Ci sono alcuni risultati che ci dicono su quali condizioni e con quali presupposizioni converge una successione $F_n(x)$ di funzioni cumulative di distribuzione verso la funzione cumulativa della distribuzione gaussiana (oppure normale).

Ci restringiamo in che segue a una versione che si chiama il *Teorema di LINDENBERG-LÉVY* e che copre essenzialmente i casi degli esempi con cui siamo avviati.

slide 10

Sia X_n una successione di variabili aleatorie *indipendenti* che hanno tutte la stessa distribuzione e che hanno una varianza *finita* $\sigma^2 > 0$ e di cui μ il valore atteso comune.

Sia

$$Z_n = \frac{\sum_{k=1}^n X_k - n \cdot \mu}{\sigma \sqrt{n}} \quad (12.1)$$

la successione di variabili *normalizzate* delle *somme* delle prime n variabili $(X_k)_{1 \leq k \leq n}$ e $F_n(z)$ la successione delle funzioni cumulative di distribuzione corrispondenti.

In questo caso segue per ogni $z \in \mathbb{R}$

$$\lim_{n \rightarrow \infty} F_n(z) = \phi(z) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{t^2}{2}} dt \quad (12.2)$$

cioè la successione delle funzioni cumulative di distribuzione $F_n(z)$ converge verso la funzione cumulativa della *distribuzione normale standardizzata*.

Dimostrazione: A causa di

Lavagna

$$\begin{aligned} Z_n &= \frac{1}{\sigma \sqrt{n}} \left(\sum_{k=1}^n X_k - n \cdot \mu \right) \\ &= \frac{1}{\sigma \sqrt{n}} \sum_{k=1}^n (X_k - \mu) := \frac{1}{\sigma \sqrt{n}} \sum_{k=1}^n Y_n \end{aligned} \quad (12.3)$$

deduciamo dall'equazione (9.4) e dall'equazione (10.6) che

$$\Phi_{Z_n}(\omega) = \Phi_Y^n \left(\frac{\omega}{\sigma \sqrt{n}} \right) \quad (12.4)$$

dove $\Phi_Y(\omega)$ è la funzione caratteristica comune delle Y_n .

Dalla presupposizione che le variabili X_n hanno una varianza **finita** $\sigma^2 > 0$ et un valore atteso comune μ sappiamo che le variabili Y_n hanno la stessa varianza **finita** $\sigma^2 > 0$ et un valore atteso comune 0.

Inoltre sappiamo che esiste almeno il secondo momento (perché esiste σ^2) è dalla definizione (8.2) e dall'equazione (9.3) estraiamo che esistono

Lavagna

$$\begin{aligned}\Phi_Y(0) &= 1 \\ \frac{d}{d\omega}\Phi_Y(0) &= (j) \cdot \mathbb{E}(Y) = 0, \\ \frac{d^2}{d\omega^2}\Phi_Y(0) &= (j)^2 \cdot \mathbb{E}(X^2) = -(\mathbb{V}(Y) + \mathbb{E}^2(Y)) = -\sigma^2.\end{aligned}\quad (13.1)$$

Siccome $\Phi_Y(\omega)$ è derivabile almeno due volte possiamo particolarmente calcolare lo sviluppo di Taylor¹ di secondo grado al punto $\omega = 0$:

$$\begin{aligned}\Phi_Y(\omega) &= \Phi_Y(0) + \frac{d}{d\omega}\Phi_Y(0) \cdot \omega + \frac{1}{2} \frac{d^2}{d\omega^2}\Phi_Y(0) \cdot \omega^2 + R_2(\omega) \\ &= 1 - \frac{\sigma^2 \cdot \omega^2}{2} + R_2(\omega)\end{aligned}\quad (13.2)$$

col resto di Lagrange

$$R_2(\omega) = \frac{1}{6} \frac{d^3}{d\omega^3}\Phi_Y(\gamma_\omega) \cdot \omega^3, \quad \gamma_\omega \in [0, \omega] \text{ (oppure } \in [\omega, 0]). \quad (13.3)$$

Dalla (12.4) risulta:

$$\begin{aligned}\Phi_{Z_n}(\omega) &= \left[1 - \frac{\sigma^2 \cdot \left(\frac{\omega}{\sigma\sqrt{n}}\right)^2}{2} + R_2\left(\frac{\omega}{\sigma\sqrt{n}}\right) \right]^n \\ &= \left[1 - \frac{\omega^2}{2n} + R_2\left(\frac{\omega}{\sigma\sqrt{n}}\right) \right]^n.\end{aligned}\quad (13.4)$$

Oltre abbiamo per il resto di Lagrange

$$R_2\left(\frac{\omega}{\sigma\sqrt{n}}\right) = \frac{1}{6} \frac{d^3}{d\omega^3}\Phi_Y(\gamma_{\frac{\omega}{\sigma\sqrt{n}}}) \cdot \left(\frac{\omega}{\sigma\sqrt{n}}\right)^3 \quad (13.5)$$

di cui possiamo dedurre che

$$\lim_{n \rightarrow \infty} n \cdot R_2\left(\frac{\omega}{\sigma\sqrt{n}}\right) = \lim_{n \rightarrow \infty} \frac{1}{6} \frac{d^3}{d\omega^3}\Phi_Y(\gamma_{\frac{\omega}{\sigma\sqrt{n}}}) \cdot \left(\frac{\omega}{\sigma}\right)^3 \cdot \frac{1}{\sqrt{n}} = 0 \quad \forall \omega \in \mathbb{R}. \quad (13.6)$$

¹ per potersi basare ad una rappresentazione più prammatica e conosciuta del teorema di Taylor mettiamo in quello che segue che il terzo momento esiste, cioè che $\Phi_Y(\omega)$ sia derivabile addirittura **tre** volte. Questo accorgimento ci permette di utilizzare il resto di Lagrange. Però in generale basta la presupposizione data nel tenore del teorema.

Dalla (13.4) risulta infine

$$\begin{aligned}\ln \Phi_{Z_n}(\omega) &= n \cdot \ln \left(1 - \frac{\omega^2}{2n} + R_2 \left(\frac{\omega}{\sigma\sqrt{n}} \right) \right) \\ &:= n \cdot \ln(1 + \rho_\omega)\end{aligned}\quad (14.1)$$

Di un formulario possiamo ricavare che

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots, \quad \forall |x| < 1. \quad (14.2)$$

Per un n sufficientemente grande possiamo ammettere alla base delle equazioni (13.4) e (13.6) che per un ω qualunque fisso vale $|\rho_\omega| < 1$ nella (14.1) e dunque

Lavagna

$$\begin{aligned}\ln \Phi_{Z_n}(\omega) &= n \cdot \rho_\omega - n \cdot \frac{\rho_\omega^2}{2} + n \cdot \frac{\rho_\omega^3}{3} - n \cdot \frac{\rho_\omega^4}{4} + \dots \\ &= -n \cdot \frac{\omega^2}{2n} + n \cdot R_2 \left(\frac{\omega}{\sigma\sqrt{n}} \right) \\ &\quad - n \cdot \frac{\omega^4}{8n^2} - n \cdot 2 \frac{\omega^2}{2n} R_2 \left(\frac{\omega}{\sigma\sqrt{n}} \right) - n \cdot R_2^2 \left(\frac{\omega}{\sigma\sqrt{n}} \right) + \dots \\ &= -\frac{\omega^2}{2} + n \cdot R_2 \left(\frac{\omega}{\sigma\sqrt{n}} \right) + \Omega(n, \omega).\end{aligned}\quad (14.3)$$

Si vede che il resto $\Omega(n, \omega)$ converge verso 0 se $n \rightarrow \infty$ (per un ω qualunque fisso) e dunque risulta con (13.6) che

$$\lim_{n \rightarrow \infty} \ln \Phi_{Z_n}(\omega) = -\frac{\omega^2}{2}. \quad (14.4)$$

Siccome la funzione esponenziale è continua risulta

$$\lim_{n \rightarrow \infty} \Phi_{Z_n}(\omega) = e^{-\frac{\omega^2}{2}}. \quad (14.5)$$

Nelle lezioni precedenti abbiamo visto che $e^{-\frac{\omega^2}{2}}$ è nient'altro che la funzione caratteristica della distribuzione Gaussiana standardizzata.

Abbiamo allora dimostrato che le funzioni caratteristiche delle variabili normalizzate Z_n convergono per ogni $\omega \in \mathbb{R}$ verso la funzione caratteristica della *distribuzione normale standardizzata*:

$$\lim_{n \rightarrow \infty} \Phi_{Z_n}(\omega) = e^{-\frac{\omega^2}{2}} \quad \forall \omega \in \mathbb{R}. \quad (14.6)$$

Se guardassimo più attentamente alla dimostrazione potremmo vedere che la convergenza è addirittura uniforme in un intervallo intorno a zero.

Dunque possiamo utilizzare il *teorema di LÉVY-CRAMÈR* per concludere infine che la successione delle funzioni cumulative di distribuzione $F_n(z)$ converge verso la funzione cumulativa della *distribuzione normale standardizzata*.

Applicazioni, Casi particolari e Generalizzazioni

Il risultato precedente ha ampie conseguenze teoriche e pratiche.

Applicazioni

In tutti i casi in cui la distribuzione di una variabile aleatoria

$$X = \sum_{k=1}^n X_i, \quad X_i \text{ iid, } n \text{ grande} \quad (15.1)$$

è sconosciuta oppure difficilmente computabile, si può adesso ammettere (se n grande) che X abbia una distribuzione Gaussiana coi parametri $\mu = n \cdot \mathbb{E}(X_i)$ e $\sigma^2 = n \cdot \mathbb{V}(X_i)$. Una probabilità si lascia dunque calcolare approssimativamente secondo

$$\begin{aligned} P(a < X \leq b) &= P\left(\frac{a - n \cdot \mu}{\sqrt{n} \cdot \sigma} < \frac{X - n \cdot \mu}{\sqrt{n} \cdot \sigma} \leq \frac{b - n \cdot \mu}{\sqrt{n} \cdot \sigma}\right) \\ &\approx \Phi\left(\frac{b - n \cdot \mu}{\sqrt{n} \cdot \sigma}\right) - \Phi\left(\frac{a - n \cdot \mu}{\sqrt{n} \cdot \sigma}\right). \end{aligned} \quad (15.2)$$

Lavagna

dove $\Phi(z)$ segna la funzione cumulativa di distribuzione (CDF) della Gaussiana normalizzata.

Caso particolare

Nell caso particolare della *macchina di Galton* si sa che la distribuzione della variabile

$$X = \sum_{k=1}^n X_i \quad (15.3)$$

in cui X rappresenta la distanza orizzontale dal centro (posizione del buco) del recipiente finale è una variabile *discreta binomiale* perché è una somma di variabili di *Bernoulli* indipendenti.

Dunque la sua funzione massa di probabilità (PMF) è conosciuta esattamente:

$$\begin{aligned} f_X(k) = P(X = k) &= \binom{n}{k} p^k (1-p)^{n-k} \\ &= \binom{n}{k} \left(\frac{1}{2}\right)^n, \quad \text{perché } p = \frac{1}{2}. \end{aligned} \quad (15.4)$$

slide 11

Però se il numero n è grande diventa difficile calcolare queste probabilità sul computer perché sia numeri estremamente grandi come il coefficiente binomiale sia numeri estremamente

piccoli come la n -esima potenza di p non possono essere calcolati senza errori numerici importanti.

In questo caso il teorema del limite centrale ci aiuta. Sappiamo che le variabili X_i sono variabili di Bernoulli con un valore atteso di $\mu = \mathbb{E}(X_i) = p$ e una varianza di $\sigma^2 = \mathbb{V}(X_i) = p(1-p) = \frac{1}{4}$ nel nostro caso speciale.

Dunque per esempio la probabilità che una pallina cade in un recipiente a distanza ± 2 dal centro (posizione del buco) è approssimativamente

$$\begin{aligned} P(-2 < X \leq 2) &= P\left(\frac{-2 - n \cdot \mu}{\sqrt{n} \cdot \sigma} < \frac{X - n \cdot \mu}{\sqrt{n} \cdot \sigma} \leq \frac{2 - n \cdot \mu}{\sqrt{n} \cdot \sigma}\right) \\ &\approx \phi\left(\frac{2}{\sqrt{n} \cdot \frac{1}{2}}\right) - \phi\left(\frac{-2}{\sqrt{n} \cdot \frac{1}{2}}\right) \\ &= 2 \cdot \phi\left(\frac{4}{\sqrt{n}}\right) - 1. \end{aligned} \quad (16.1)$$

Se la macchina ha per esempio $n = 16$ righe si ha

slide 11
$$P(-2 < X \leq 2) \approx 2 \cdot \phi\left(\frac{4}{\sqrt{16}}\right) - 1 = 0.6827. \quad (16.2)$$

A questo punto bisogna notare che la formula (16.1) può essere migliorata aggiungendo un termine di correzione per tenere conto del fatto che stiamo approssimando una distribuzione discreta con una distribuzione continua:

$$P(a < X \leq b) = P\left(\frac{a - n \cdot \mu - 0.5}{\sqrt{n} \cdot \sigma} < \frac{X - n \cdot \mu}{\sqrt{n} \cdot \sigma} \leq \frac{b - n \cdot \mu + 0.5}{\sqrt{n} \cdot \sigma}\right). \quad (16.3)$$

Per la macchina nel esempio $n = 16$ risulta

slide 12
$$P(-2 < X \leq 2) \approx 2 \cdot \phi\left(\frac{5}{\sqrt{16}}\right) - 1 = 0.7887 \quad (16.4)$$

che corrisponde meglio ai risultati degli esperimenti.

esperimento
col
programma
galton.exe

Quanto al numero n necessario affinché l'approssimazione ottenuto sia accettabile si usa una regola approssimativa che dice che bisogna avere $np(1-p) > 9$ per una approssimazione ammissibile. Nel esempio questa regola non poteva essere rispettato perché nel programma non si poteva scegliere più di 16 righe. Eppure i risultati erano soddisfacenti.

Il Teorema di Moivre-Laplace

Nel caso come quello precedente in cui la successione delle funzioni cumulative di distribuzione (CDF) è infatti una successione di funzioni cumulative *binomiali* $F_{B(n,p)}(x)$ si parla del *Teorema di Moivre-Laplace*:

Sia X_n una successione di variabili aleatorie *binomiali indipendenti* coi parametri n e $0 < p < 1$.

Sia

$$Z_n = \frac{\sum_{k=1}^n X_k - n \cdot p}{\sqrt{n \cdot p \cdot (1-p)}} \quad (17.1)$$

la successione di variabili *normalizzate* delle *somme* e $F_n(z)$ la successione delle funzioni cumulative di distribuzione (CDF) corrispondenti.

In questo caso segue per ogni $z \in \mathbb{R}$

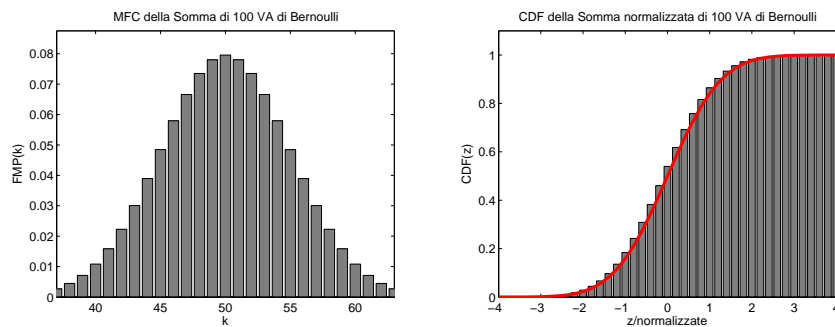
$$\lim_{n \rightarrow \infty} F_n(z) = \phi(z) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z e^{-\frac{t^2}{2}} dt \quad (17.2)$$

cioè la successione delle funzioni cumulative di distribuzione $F_n(z)$ converge verso la funzione cumulativa della *distribuzione normale standardizzata*.

Con l'aiuto di MATLAB possiamo in modo sperimentale verificare il risultato precedente.

Coi programmi `EempSomBernoulli.m` et `sommaBernoulli.m` possiamo calcolare con l'aiuto della funzione di MATLAB `conv`, realizzando la convoluzione discreta, la MPF e la CDF normalizzata di una somma di n VA di Bernoulli col parametro p .

Nella figura 10 vediamo per esempio il risultato della simulazione per i parametri $n = 100$ e $p = 0.5$.



Grafica 10: Il Teorema di Moivre-Laplace

Si vede che la CDF normalizzata della funzione cumulativa *binomiale* $F_{B(n,p)}(x)$ coincide benissimo con la curva della funzione cumulativa *Gaussiana* standardizzata.

slide 13

simulazione
MATLAB
EempSom-
Bernoulli.m
somma-
Bernoulli.m

Generalizzazioni

Non in tutti i casi le presupposizioni che le variabili sommate siano indipendenti oppure identicamente distribuite possono essere verificate.

Né nel caso del getto d'acqua né nel caso dello strumento acustico conosciamo la distribuzione della pressione del suono che fa *una sola goccia* oppure *una sola pallina di ferro*.

Sappiamo neanche se le gocce o le palline sono identicamente distribuite (oppure statisticamente indipendenti).

Pur nondimeno abbiamo osservato una distribuzione Gaussiana in tutti e due i casi.

Il Teorema del Limite Centrale di LINDENBERG-FELLER

Tuttavia ci sono risultati teorici (il teorema di LJAPUNOV, il teorema di LINDENBERG-FELLER) dicendo che anche in questi casi le funzioni cumulative (CDF) di una somma di variabili indipendenti convergono verso la funzione cumulativa Gaussiana se grosso modo nessun gruppo di variabili *domina in varianza* gli altri variabili.

Questo ci spiega in molti casi pratici l'osservazione della distribuzione Gaussiana anche se le distribuzioni esatte non sono conosciute.

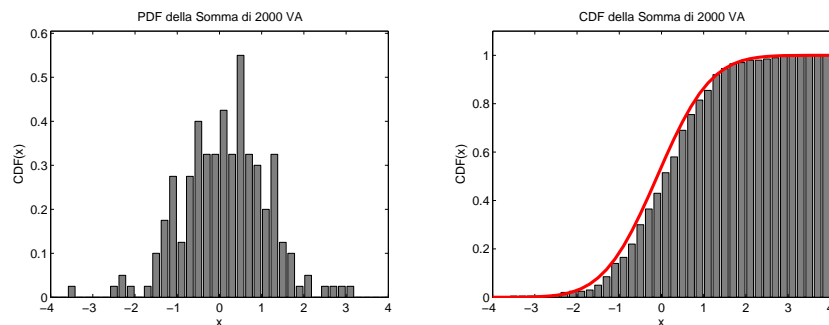
Non vorremmo entrare nei dettagli e riprodurre esattamente il contenuto di questi teoremi. Invece vorremmo ancora una volta verificare in modo sperimentale con l'aiuto di MATLAB che i risultati valgono.

Col programma `SomQualunque.m` possiamo calcolare una somma di n VA *normalizzate* indipendenti *senza dominanza della varianza*.

simulazione
MATLAB
Som-
Qualunque.m

I valori delle somme vengono generati in modo sperimentale. Da un gruppo di distribuzioni continue vengono generati e sommati n valori. Ogni valore proviene da una distribuzione scelta a caso da un generatore casuale. I parametri delle distribuzioni sono fissati per motivi di semplicità in modo tale che nessun gruppo di VA ha una varianza dominante.

Nella figura 11 vediamo per esempio il risultato della simulazione di 200 prove con $n = 2000$ VA scelte dal insieme di tutte le 10 distribuzioni predefinite.



Grafica 11: Il Teorema del Limite Centrale di LINDENBERG-FELLER

Si vede che la CDF normalizzata della funzione cumulativa coincide abbastanza bene con la curva della funzione cumulativa *Gaussiana* standardizzata.

Il Teorema del Limite di Poisson

Un caso speciale in cui le somme (15.1) *non convergono* verso una Gaussiana è il *teorema del limite di Poisson* che ci dice che la CDF della somma in (15.1) converge infatti verso una CDF di Poisson se le variabili X_k sono variabili di Bernoulli indipendenti col parametro p_k soddisfacendo alla condizione $\lim_{k \rightarrow \infty} k \cdot p_k = \lambda > 0$.

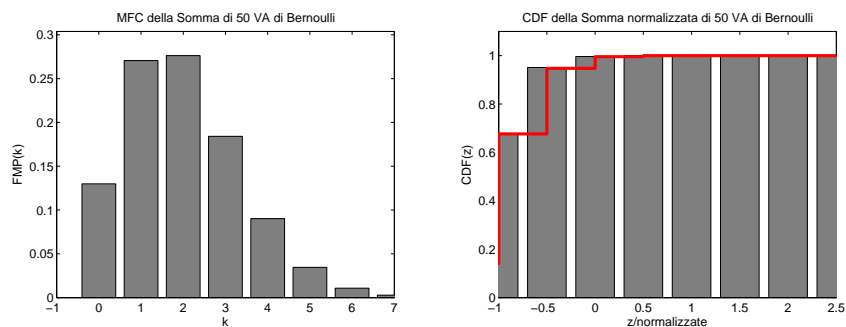
In questo caso viene violata la presupposizione della non-dominanza in varianza perché le varianze $p_k(1 - p_k)$ devono convergere verso 0 per soddisfare alla condizione. Dunque un gruppo finito delle variabile sommate ha delle varianze sufficientemente grandi è un gruppo infinito ha delle varianze troppo piccole.

Col programma `sommaPoisson.m` oppure con `somPoisson.m` possiamo calcolare una somma di n VA *normalizzate* indipendenti soddisfacendo alla condizione $\lim_{k \rightarrow \infty} k \cdot p_k = \lambda > 0$.

La funzione `sommaPoisson.m` calcola direttamente le distribuzioni delle somme (che sono infatti tutti distribuzioni binomiali) con l'aiuto delle funzioni di MATLAB `binopdf` e `binocdf`.

Nella funzione `somPoisson.m` i valori delle somme vengono ancora una volta generati in modo sperimentale da un generatore casuale.

Nella figura 12 vediamo per esempio il risultato della simulazione di $n = 50$ VA per cui vale la condizione $n \cdot p_n = \lambda = 2$.



Grafica 12: Il Teorema del Limite di Poisson

Si vede bene che in questo caso la CDF normalizzata della funzione cumulativa *non assomiglia* alla funzione cumulativa *Gaussiana* standardizzata bensì alla funzione cumulativa standardizzata *di Poisson*.

Epilogo

Come ce l'ho già detto all'inizio della lezione, dal 2007 esiste una cooperazione Erasmus fra il Dipartimento di Ingegneria della vostra università e il Dipartimento di Costruzione Meccanica et Meccatronica della mia università.

simulazione
MATLAB
somma-
Poisson.m
somPoisson.m

Quanto al cambio di studenti sono già state due studenti della nostra università a Parma di cui uno durante il semestre attuale.

Un vero cambio però non può essere unilaterale. Per questo caso questa lezione è per me anche un'occasione di fare un pò di pubblicità.

Se uno di voi ha voglia di passare un semestre o forse di più in Germania lo prego di rivolgersi direttamente a me oppure al professore Guarino Lo Bianco.

Grazie della vostra attenzione.