

ALMA MATER STUDIORUM · UNIVERSITÀ DI BOLOGNA

FACOLTÀ DI SCIENZE MATEMATICHE, FISICHE E NATURALI
Corso di Laurea Magistrale in Matematica

**CORREZIONI ASINTOTICHE PER
IL TEOREMA DEL LIMITE
CENTRALE GENERALIZZATO**

Tesi di Laurea in Metodi matematici per la Finanza

Relatore:
Chiar.mo Prof.
Vincenzo Grecchi

Presentata da:
Giovanni Rabitti

Seconda Sessione
Anno Accademico 2012/2013

*Ai miei nonni,
che avrebbero tanto voluto vedere questo giorno.*

Indice

1	Introduzione	3
1.1	Esempio introduttivo: calcolo del VaR	4
2	Preliminari	8
2.1	Definizioni di distribuzione stabile	9
2.2	Probabilità delle code e momenti	10
2.3	Sulla somma di variabili aleatorie	12
3	Oltre il Limite Centrale: gli sviluppi asintotici	17
3.1	Funzioni generatrici per \bar{S}_N	17
3.2	Polinomi di Hermite	19
3.3	Espansioni asintotiche di Edgeworth	20
3.4	Serie di Gram-Charlier di Tipo A	21
4	Correzioni asintotiche razionali	23
4.1	Espansioni classiche per $F_n(x)$	30
A	Approssimanti di Padé	31
A.1	Proprietà omografiche	33
	Bibliografia	36

Capitolo 1

Introduzione

"Il diavolo sta nei dettagli"
– Detto popolare

Gli eventi rari, come si deduce dal nome, sono di difficile previsione oltre che di (quasi) impossibile modellizzazione matematica. I modelli maggiormente utilizzati infatti si basano sulle proprietà (comodissime) della distribuzione Gaussiana (altreimenti detta Normale) per variabili aleatorie, di cui si può andare a considerare solo la parte ‘centrale’, dimenticandosi delle code che asintoticamente vanno a zero e che quindi rappresentano la probabilità di verificarsi di eventi molto improbabili. Sia chiaro, questa non è un’operazione matematicamente scandalosa: la giustificazione di questa semplificazione risiede infatti in un risultato ‘pesante’: la Disuguaglianza di Chebyshev ¹.

Questa comodità dell’uso della Normale fa sí che i modelli applicati in Economia e Finanza siano piú facili da trattare e oltretutto da parte di quei professionisti che, ad esempio, non possiedono un background matematico consistente (quindi operatori di mercato, banchieri, ecc). Da parte di molti autori infatti, specialmente a partire dal 2009, c’è stata una crescente critica alla matematica e ai modelli matematici semplificati, usati con disinvoltura da parte dell’industria finanziaria.

In particolare, due sono stati i principali articoli che hanno individuato le cause della crisi dei sub-prime e del successivo crollo di banche come Lehman Brothers: “Recipe for disaster: the formula that killed Wall Street” (Salmon, Wired Magazine), pubblicato online, e il piú fortunato “Of couples and copulas: the formula that felled Wall St” (Jones, Financial Times 2009 e Wiley Publications 2012) con 157 fino ad oggi (un solo anno).

¹Sia X una variabile aleatoria reale con valore atteso μ e varianza σ^2 , sia $\lambda \in \mathbb{R}$, allora:

$$P(|X - \mu| \geq \lambda \cdot \sigma) \leq \frac{1}{\lambda^2}$$

cioé la probabilità di trovare una realizzazione di X distante λ volte dal valore centrale atteso va a zero come $\frac{1}{\lambda^2}$.

Questa formula a cui si fa riferimento nei titoli é la formula di Li per la copule Gaussiane, che vengono modellizzate appunto in un mondo dove non vengono considerati eventi improbabili (azzeccato in tal senso é il titolo dell'articolo "The devil is in the tails: actuarial mathematics and the subprime mortgage crisis", [1]). La domanda cruciale a cui quindi si arriva é la seguente: dato un processo aleatorio, quanta informazione si perde considerando solo la parte centrale di una distribuzione (fondamentale in forza del Teorema del limite centrale) e scardando regioni nelle code? C'è la possibilità di migliorare il risultato, di correggerlo, a partire da una Gaussiana per generalizzarlo?

Cerchiamo di introdurre la questione con un esempio teorico (modificato a partire da [2]):

1.1 Esempio introduttivo: calcolo del VaR

Vogliamo calcolare la probabilità di perdere tutto il capitale investito in borsa in una sola seduta.

Il valore a rischio (Var), cioè la massima perdita che *concretamente* si può verificare con una certa probabilità, viene indicato con Λ .

La legge dei grandi numeri ci dice che un evento, avente probabilità p in una singola giocata, appare in media

$$\langle \nu_n \rangle = n \langle f_n \rangle = np$$

volte in una serie di n giocate indipendenti, essendo le probabilità le frequenze medie. Ci si aspetta quindi di osservare eventi che hanno probabilità

$$p \sim 1/n$$

in modo che $\langle \nu_n \rangle \sim 1$: sarebbe molto sorprendente trovare un evento di probabilità molto inferiore a $1/n$.

L'ordine di grandezza dell'evento massimo Λ_{max} , osservato in una successione di n variabili aleatorie iid (indipendenti e identicamente distribuite), é data quindi da:

$$P_{>}(\Lambda_{max}) = P(x_i > \Lambda_{max}) = 1 - P(\Lambda_{max}) = \frac{1}{n}.$$

Piú precisamente, definito il valore massimo $x_{max} = \max\{x_i\}$, la misura di probabilità $P_n(x_{max} < \Lambda)$ é ottenuta notando il fatto che, se il massimo di tutte le x_i é minore di Λ , allora tutte le x_i sono minori di Λ . Ma sono iid e quindi:

$$P_n(x_{max} < \Lambda) = P_n(x_1, \dots, x_n < \Lambda) = \prod P(x_j < \Lambda) = P(\Lambda)^n$$

da cui si ricava, per $P_n(x_{max} < \Lambda_p) = p$:

$$P_{>}(\Lambda_p) = 1 - P_n(x_{max} < \Lambda)^{1/n} = 1 - p^{1/n} = 1 - \exp(\log p/n) \sim -\frac{\log p}{n}$$

Si noti che questo risultato é generale e non dipende dalla scelta specifica della distribuzione di probabilitá $F(x)$; cio' nonostante, serve una sua forma analitica per poter fare previsioni numeriche.

Caso 1 Dall'analisi statistica delle serie storiche viene suggerito l'uso di $F(x)$ con andamento esponenziale, cioé $P_{>}(x) \sim \exp(-\alpha x)$, e quindi si trova:

$$\exp(-\alpha \Lambda_{max}) = \frac{1}{n} \iff \Lambda_{max} \sim \frac{\log n}{\alpha}$$

Tuttavia, varie inefficienze nel modello della predizione di Λ_{max} sono da anni documentate: nella tabella successiva vengono confrontate le previsioni della massima perdita giornaliera nel caso di code esponenziali, fatte nel periodo 1989 – 1993, con la peggiore perdita giornaliera osservata nel 1994, che é quindi un valore empirico.

I dati sono presi da [4].

Paese	Λ_{max} previsto	Λ_{max} osservato
Gran Bretagna	1.59	2.08
Germania	0.60	1.44
Giappone	0.65	1.08
USA	1.85	2.26
Italia	1.31	4.18

Risulta chiaro che la previsione esponenziale risulta troppo ottimistica e, pertanto, poco attendibile.

Quello che si cerca di fare, é usare un approccio diverso per determinare il comportamento asintotico di $F(x)$ per eventi estremi (cioé a probabilitá asintoticamente nulla).

Caso 2 Anticipiamo quindi un piccolo risultato delle distribuzioni stabili, che in qualche modo generalizzano il mondo Gaussiano, come spiegheremo successivamente.

Da un teorema ² (che si può trovare ad esempio in [3]), segue come corollario che:

$$P_{>}(x) \sim x^{-\alpha}$$

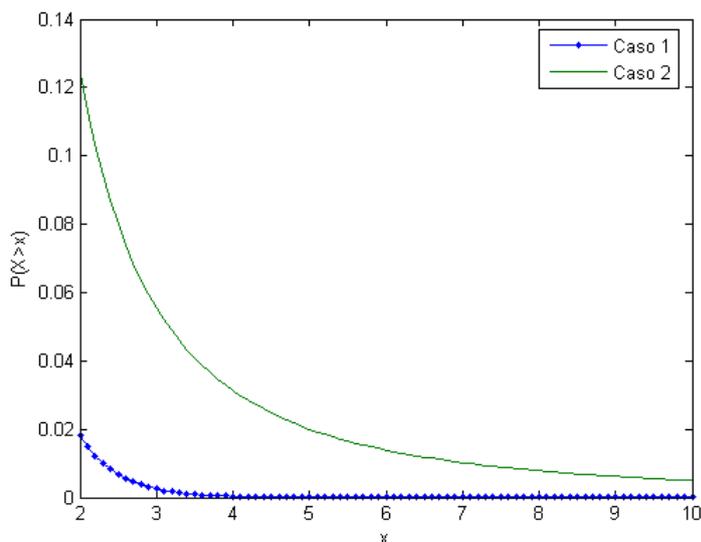


Figura 1.1: Comportamento asintotico delle code

Confrontandole, salta all'occhio che chiaramente entrambe vanno a 0, ma il modo con cui ci vanno è fondamentale: c'è più probabilità nella coda stabile. In altre parole, eventi più rari vengono tenuti in considerazione nella modellizzazione 'generalizzata' perché viene assegnata a loro più probabilità di verificarsi.

Che fare ora quindi? Si cerca un metodo per migliorare le stime asintotiche della coda, che appunto si chiama 'correzioni asintotiche'.

Alla fine della tesi scopriremo che si potrà adottare il termine correttivo asintotico del **Caso 1**:

$$P_{>}(x) \sim \left(\exp(-\alpha x) + \frac{1}{\sqrt{2}x^2} \right) (1 + o(1))$$

²Sia $\{X(t), t \geq 0\}$ un moto Browniano. Sia T_a il primo momento in cui raggiunge il livello $a > 0$. Allora, per $x > 0$, T_a ha distribuzione stabile di Lévy:

$$P(X \leq x) = 2 \left(1 - \Phi \left(\frac{a}{\sqrt{x}} \right) \right)$$

dove Φ indica la funzione cumulativa della distribuzione $N(0, 1)$.

e si otterrà:

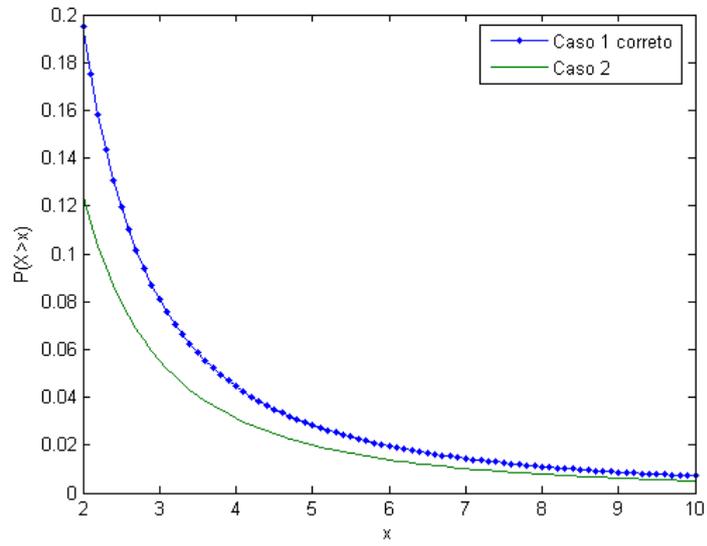


Figura 1.2: Comportamento asintotico con correzione

Capitolo 2

Preliminari

Con il termine 'distribuzioni stabili' si indica una ricca classe di distribuzioni di probabilità utili allo studio dei fenomeni a *heavy tails*. Questa classe è stata per prima caratterizzata da Paul Lévy nel suo studio sulla somma di variabili aleatorie indipendenti e identicamente distribuite (iid).

Il loro uso è spesso proposto come modello per molti tipi di sistemi fisici ed economici. Ci sono, infatti, molte ragioni per usare le distribuzioni stabili, ad esempio:

- la modellizzazione di un fenomeno che ci si aspetta essere non-Gaussiano;
- la formalizzazione del Teorema del Limite Centrale Generalizzato (GCLT), che afferma che l'unico limite non-triviale possibile di una somma normalizzata di variabili aleatorie iid è stabile;
- empiricamente è stato osservato che molti data set mostrano un comportamento a *heavy tail*

Tuttavia, la mancanza di formule chiuse per le densità e le funzioni di distribuzione per la maggior parte delle distribuzioni stabili (fanno eccezione la Gaussiana e le distribuzioni di Cauchy e di Lévy), è il maggior deterrente per il loro uso.

Inoltre, non esiste una sola definizione: qui ne daremo quattro, che sono poi le più utilizzate. Le prime due riguardano le proprietà di stabilità: la famiglia delle distribuzioni stabili è invariante sotto convoluzione. La terza concerne la funzione caratteristica di una variabile casuale stabile. La quarta ed ultima riguarda il ruolo delle distribuzioni stabili nel contesto del teorema del limite centrale: come detto prima, le distribuzioni stabili approssimano la distribuzione delle somme normalizzate di variabili aleatorie iid.

La distribuzione stabile dipende in particolare da un parametro $\alpha \in [0, 2]$, detto anche indice di stabilità; per $\alpha = 2$ si ottiene la Gaussiana. Se $\alpha < 2$ le code della distribuzione vanno a 0 come una potenza negativa. In altre parole, una variabile aleatoria stabile mostra molta più variabilità di una variabile Gaussiana: è molto più probabile

che prenda valori lontano dalla mediana. Questa sua alta variabilità é il motivo centrale per cui sono importanti nelle modellizzazioni.

2.1 Definizioni di distribuzione stabile

La teoria delle distribuzioni stabili ha origine nella teoria sviluppata tra il 1920 e il 1930 da Paul Lévy. Successivamente, altri autori hanno approfondito ed ampliato l'argomento: ci si riferisce alle monografie di Gnedenko e Kolmogorov (1954) e Feller (1971). In questa tesi, per quanto riguarda le definizioni, si fá comunque riferimento a [3] e [5].

Definizione 1. *Una variabile casuale X ha una distribuzione stabile se, presi due numeri positivi A e B , esiste un numero positivo C e un numero reale D tali che*

$$AX_1 + BX_2 \stackrel{d}{=} CX + D$$

dove X_1 e X_2 sono copie identiche di X e $\stackrel{d}{=}$ significa uguaglianza in legge (cioé entrambe le espressioni hanno la stessa legge di probabilità).

Da notare che una variabile casuale X considerata puntualmente é sempre stabile. Inoltre, X si dice *strttamente* stabile se $D = 0$. X si dice infine stabile *simmetrica* se la sua distribuzione é stabile e simmetrica, cioé X e $-X$ hanno la stessa distribuzione.

Esempio 1. *Se X é una variabile casuale Gaussiana di media μ e varianza ν^2 , cioé $X \sim N(\mu, \nu^2)$, allora X é stabile con $\alpha = 2$ poiché*

$$AX_1 + BX_2 \sim N((A+B)\mu, (A^2 + B^2)\nu^2)$$

che é la definizione sopra con la scelta $C = (A^2 + B^2)^{1/2}$ e $D = (A + B - C)\mu$.

Diamo la seconda definizione:

Definizione 2. *Una variabile aleatoria X ha una distribuzione stabile se, $\forall n \geq 2$, esistono un numero positivo C_n e un numero reale D_n tali che*

$$X_1 + X_2 + \dots + X_n \stackrel{d}{=} C_n X + D_n$$

dove X_1, X_2, \dots, X_n sono copie indipendenti di X .

Le definizioni sopra sono equivalenti. Infatti, se X é stabile secondo la Definizione 1, allora per induzione é stabile anche secondo la Definizione 2. Il viceversa lo mostra Feller; quindi le due definizioni sono equivalenti.

La terza definizione specifica la funzione caratteristica di una variabile aleatoria stabile: puó essere quindi la forma analitica.

Definizione 3. Una variabile aleatoria X possiede una distribuzione stabile se esistono parametri $0 \leq \alpha \leq 2$, $\sigma \geq 0$ e M tali che la sua funzione caratteristica ha la seguente forma

$$\langle \exp i\theta X \rangle = \exp \{-\sigma^\alpha |\theta|^\alpha + iM\theta\}$$

Esiste in generale una formula di rappresentazione, detta di *Lévy-Khinchine*, a seconda dei valori di α : per $\alpha = 2$ si ottiene

$$\langle \exp i\theta X \rangle = \exp \{-\sigma^2 \theta^2 + i\mu\theta\}$$

che è la funzione caratteristica di una variabile Gaussiana di media μ e varianza $2\sigma^2$. Per $\alpha < 2$ si ha

$$\langle \exp i\theta X \rangle = \exp \left\{ iM\theta + \int_{-\infty}^{\infty} \left(e^{i\theta x} - 1 - \frac{i\theta x}{1+x^2} \right) L(dx) \right\}$$

dove la misura

$$L(dx) = \frac{P}{x^{1+\alpha}} \mathbb{I}_{(0,\infty)}(x) dx + \frac{Q}{|x|^{1+\alpha}} \mathbb{I}_{(-\infty,0)}(x) dx$$

con P e Q numeri non negativi, è chiamata formula non simmetrica della *misura di Lévy*. È importante anche perché l'interpretazione fisica di $L(dx)$ riguarda la frequenza dei 'salti' di lunghezza x in una passeggiata aleatoria (cfr. [4]).

2.2 Probabilità delle code e momenti

Quando $\alpha = 2$, la distribuzione normale ha proprietà delle code ben conosciute. Ad esempio:

Proposizione 1. Sia X un numero aleatorio con distribuzione normale standard. Per ogni $x \geq 0$ vale che:

$$\frac{n(x)}{x} - \frac{n(x)}{x^3} < P(X \geq x) < \frac{n(x)}{x}$$

dove $n(x) := \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.

Asintoticamente, si ottiene quindi che:

$$P(X < -x) = P(X > x) \sim \frac{1}{2\sqrt{\pi}x} e^{-x^2/2}$$

per $x \rightarrow \infty$. Per la dimostrazione, si veda [6].

In questa sezione, invece, sottolineiamo le proprietà delle code per distribuzioni stabili non-Gaussiani ($\alpha < 2$), che sono leggi esponenziali asintotiche a code pesche. Vale la seguente:

Proposizione 2 (Approssimazione delle code). *Sia X una variabile stabile con $0 < \alpha < 2$ e varianza σ^2 . Allora:*

$$\lim_{x \rightarrow \infty} x^\alpha P(X < -x) = x^\alpha P(X > x) = \frac{C_\alpha}{2} \sigma^\alpha,$$

dove

$$C_\alpha = \frac{1}{\int_0^\infty x^{-\alpha} \operatorname{sen} x dx} = \begin{cases} \frac{1-\alpha}{\Gamma(2-\alpha)\cos(\pi\alpha/2)} & \text{se } \alpha \neq 1 \\ 2/\pi & \text{se } \alpha = 1 \end{cases}$$

Di conseguenza, in forza del teorema Tauberiano (cfr. [3]), si trova:

$$P(X > x) \sim \frac{\alpha^\alpha}{\Gamma(1-\alpha)} x^{-\alpha} = \frac{\sigma^\alpha}{\Gamma(1-\alpha)\cos(\pi\alpha/2)} x^{-\alpha} \quad (2.1)$$

$$(2.2)$$

$$= \sigma^\alpha C_\alpha x^{-\alpha} \quad (2.3)$$

per $x \rightarrow \infty$.

Le distribuzioni di Pareto sono una classe di leggi di probabilità con code le cui probabilità sono date da (2.3). Il termine 'legge stabile di Pareto' viene quindi usata per distinguere tra la decrescita rapida delle code della Gaussiana e quelle più lente tipo Pareto con il caso $\alpha < 2$.

Un primo problema delle heavy tails é che non tutti i momenti esistono. Nella maggior parte dei problemi statistici si fa uso solitamente solo due primi due momenti: il primo (cioé la media) $\langle X \rangle$, e la varianza $\langle X^2 \rangle - \langle X \rangle^2$. Purtroppo, nell'approccio delle distribuzioni stabili non sono in generale utili, perché le espressioni integrali per le medie possono divergere; al loro posto conviene introdurre i momenti frazionari assoluti

$$\langle |X|^p \rangle = \int P(|X|^p > x) dx = \int |x|^p f(x) dx$$

con p numero reale. Vale comunque in generale un criterio:

Proposizione 3 (Caratterizzazione dei momenti). *Sia $0 < \alpha < 2$, allora:*

$$\langle |X|^p \rangle < \infty \quad \text{per ogni } 0 < p < \alpha$$

$$\langle |X|^p \rangle = \infty \quad \text{per ogni } p \geq \alpha$$

Quindi, per $0 < \alpha < 2$, $\langle |X|^2 \rangle = \langle X^2 \rangle = +\infty$ e le distribuzioni stabili non possiedono secondi momenti finiti o varianze. Questo fatto é una delle cause principali al mancato uso delle distribuzioni stabili nelle applicazioni pratiche. Per la forma analitica, si faccia riferimento a [3].

Per quanto riguarda le medie, per $1 < \alpha \leq 2$, $\langle |X| \rangle = \mu \in \mathbb{R}$, cioè le distribuzioni possiedono il valore atteso. D'altra parte, se $\alpha \leq 1$, allora $\langle |X| \rangle = +\infty$, quindi le medie sono indefinite. Per la forma analitica, si veda [5].

2.3 Sulla somma di variabili aleatorie

Un altro motivo fondamentale per la popolarità della distribuzione Gaussiana riguarda il Teorema del Limite Centrale. Ora, detto a parole, questo teorema stabilisce che la somma di N variabili iid,

$$S_N = \sum_{n=1}^N X_n,$$

segue al limite la Gaussiana, cioè per $N \rightarrow \infty$, fornendo quindi i primi due momenti di X_n (che non divergono). Queste restrizioni sono così leggere che molte distribuzioni appartengono al dominio di attrazione della Gaussiana.

Tuttavia, un famoso controesempio è la *distribuzione di Cauchy*,

$$p(x) = \frac{a}{\pi} \frac{1}{a^2 + x^2},$$

il cui secondo momento è infinito. Quindi questa distribuzione è solo un esempio di una classe più generale di distribuzioni che sono tipo heavy tails:

$$p(x) \sim \frac{1}{|x|^{1+\alpha}}, \quad 0 < \alpha < 2 \quad (|x| \rightarrow \infty)$$

Ci si chiede quindi se la distribuzione di Cauchy può essere il limite per S_N , e come può essere la distribuzione limite se le variabili iid seguono Cauchy.

Per iniziare dunque, consideriamo una successione di variabili casuali $\{X_n\}_{n=1,2,\dots,N}$ indipendenti ed identicamente distribuite secondo

$$\text{Prob}(x < X_n < x + dx) = p(x)dx \quad (n = 1, \dots, N).$$

Sorgono quindi delle domande:

- È possibile trovare due costanti reali A_N e B_N (> 0) tali che la distribuzione della somma normalizzata,

$$\bar{S}_N = \frac{1}{B_N} \sum_{n=1}^N X_n - A_N,$$

converge a una distribuzione limite quando N tende a infinito, cioè:

$$\text{Prob}(x < \bar{S}_N < x + dx) \rightarrow L(dx) \quad (N \rightarrow \infty)?$$

- Che forma hanno tutte le possibili distribuzioni limite?

- Quando la densità di probabilità $p(x)$ appartiene al dominio di attrazione di una specifica $L(x)$?

La risposta a queste domande richiede intanto la quarta definizione equivalente di *distribuzione stabile*:

Definizione 4. Una variabile aleatoria X ha una distribuzione stabile se, data una successione di variabili iid X_1, X_2, \dots e successioni di numeri positivi $\{B_n\}$ e di reali $\{A_n\}$ tali che:

$$\frac{X_1 + \dots + X_n}{B_n} + A_n \stackrel{d}{=} X. \quad (2.4)$$

Notiamo che, se X è Gaussiana e le X_i hanno varianza finita, allora (2.4) è il Teorema del Limite Centrale ordinario.

Esiste una caratterizzazione anche rispetto alla densità. Infatti vale

Proposizione 4. Una densità di probabilità è stabile se è invariante rispetto alla convoluzione, cioè esistono costanti $a > 0$ e b tali che

$$p(a_1x + b_1) * p(a_2x + b_2) := \int_{-\infty}^{+\infty} p(a_1(z - x) + b_1)p(a_2x + b_2)dx \quad (2.5)$$

$$= p(az + b) \quad (2.6)$$

per tutte le costanti reali $a_1 > 0, b_1, a_2 > 0, b_2$.

Questa caratterizzazione può sembrare ostica, ma in realtà aiuta a trovare possibili distribuzioni stabili. L'equazione (2.6) risulta infatti particolarmente apprezzabile nello spazio di Fourier, dove la convoluzione $p(z) = f(x) * g(x)$ è il prodotto delle trasformate:

$$\begin{aligned} p(k) &=: \mathcal{F}[p(z)] = \int_{-\infty}^{\infty} \exp(ikz)p(z)dz \\ &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} f(z-x)g(x)dx \right) e^{ikz} dz \\ &= f(k)g(k). \end{aligned}$$

Ad esempio, prendiamo la densità di una Gaussiana e scegliamo $a_1 = a_2 = 1$ e $b_1 = b_2 = 0$. Allora

$$\mathcal{F}[p(x) * p(x)] = \exp\left(-\frac{k^2}{2}\right)\exp\left(-\frac{k^2}{2}\right) = \exp(-k^2)$$

in modo che $a = 1/\sqrt{2}$ e $b = 0$. La formula appena trovata non è specifica di una Gaussiana, ma è soddisfatta anche da esponenziali con altri argomenti. Ad esempio, dal semplice esponenziale

$$\exp\left(-\frac{|k|}{2}\right)\exp\left(-\frac{|k|}{2}\right) = \exp(-|k|),$$

che é la trasformata di Fourier della distribuzione di Cauchy

$$\begin{aligned}\mathcal{F}^{-1}[\exp(-|k|)] &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp(-ikx) \exp(-|k|) dk \\ &= \frac{1}{\pi} \frac{1}{1+x^2}.\end{aligned}$$

Quindi, sia la Normale sia la distribuzione di Cauchy sono, a meno di costanti di normalizzazione, invarianti rispetto alla convoluzione e sono dunque stabili in forza della Proposizione precedente.

Teorema 1. *Una densit  di probabilit  pu  essere solamente una distribuzione limite della somma normalizzata \bar{S}_N si variabili aleatorie indipendenti se   stabile.*

Si noti che non viene richiesto che le X_i abbiano la stessa distribuzione. Le distribuzioni Gaussiana e di Cauchy sono quindi potenziali distribuzioni limite. Ma non sono chiaramente solo loro, ce ne sono molte altre. Il prossimo teorema ne fornisce una caratterizzazione.

Teorema 2. *(Rappresentazione canonica generale) Una densit  di probabilit    stabile se e solo se il logaritmo della sua funzione caratteristica,*

$$L_{\alpha,\beta}(k) = \langle e^{ikx} \rangle = \int_{-\infty}^{\infty} L_{\alpha,\beta}(x) \exp(ikx) dx,$$

si pu  scrivere come

$$\ln L_{\alpha,\beta}(k) = i\gamma k - c|k|^\alpha \left(1 + i\beta \frac{k}{|k|} \omega(k, \alpha)\right),$$

dove $\gamma, c \geq 0$ sono costanti reali e

$$0 < \alpha \leq 2, \quad -1 \leq \beta \leq 1,$$

e la funzione $w(k, \alpha)$   data da

$$w(k, \alpha) = \begin{cases} \tan(\pi\alpha/2) & \alpha \neq 1 \\ (2/\pi)\ln|k| & \alpha = 1 \end{cases}$$

Le costanti γ e c sono costanti di normalizzazione. sostituendo $x - \gamma$ con $c^{1/\alpha}x$ trasla l'origine ma non cambia la funzione $L_{\alpha,\beta}(x)$. Al contrario, i parametri α e β definiscono la forma e le propriet  di $L_{\alpha,\beta}(x)$. Il parametro α caratterizza il comportamento asintotico di $L_{\alpha,\beta}$ determina i momenti:

- $0 < \alpha < 2$: Ogni distribuzione stabile si comporta come

$$L_{\alpha,\beta}(x) \sim \frac{1}{|x|^{1+\alpha}} \quad \text{per } x \rightarrow \pm\infty, \quad (2.7)$$

e riotteniamo quanto trovato nella Proposizione 3, cioè ha momenti finiti assoluti di ordine p

$$\langle |x|^p \rangle = \int_{-\infty}^{\infty} |x|^p L_{\alpha,\beta}(x) dx \quad \text{se } 0 < p < \alpha.$$

- $\alpha = 2$: $L_{\alpha,\beta}(x)$ é indipendente da β ed é Gaussiana.

A causa di queste proprietà, α viene quindi chiamato *esponente caratteristico*. Il secondo esponente caratteristico, β , determina l'asimmetria di $L_{\alpha,\beta}(x)$:

- $\beta = 0$: $L_{\alpha,\beta}(x)$ é una funzione pari in x .
- $\beta = \pm 1$: $L_{\alpha,\beta}(x)$ é fortemente asimmetrica.

La rappresentazione canonica definisce tutte le possibili espressioni per le distribuzioni stabili. Tuttavia, non specifica le condizioni che la densità di probabilità $p(x)$ deve soddisfare affinché la distribuzione della somma normalizzata \bar{S}_N converga a una particolare $L_{\alpha,\beta}(x)$ nel limite per $N \rightarrow \infty$. Se succede effettivamente questo, si dice che $p(x)$ appartiene al dominio di attrazione di $L_{\alpha,\beta}(x)$. Questo problema é stato risolto.

Teorema 3. *La densità di probabilità $p(x)$ appartiene al dominio di attrazione di una densità stabile $L_{\alpha,\beta}(x)$ con esponente caratteristico α ($0 < \alpha < 2$) se e solo se*

$$p(x) \sim \frac{\alpha a^\alpha c}{|x|^{1+\alpha}} \quad \text{per } x \rightarrow \infty,$$

dove $c > 0$, $a > 0$ sono costanti.

Inoltre, se $p(x)$ appartiene al dominio di attrazione di una distribuzione stabile, esistono i suoi momenti assoluti di ordine p per $p < \alpha$:

$$\langle |x|^p \rangle = \int_{-\infty}^{\infty} |x|^p p(x) dx \begin{cases} < \infty & \text{per } 0 \leq p < \alpha (< 2) \\ = \infty & \text{per } p > \alpha (< 2) \end{cases}$$

e la costante di normalizzazione B_N , che caratterizza il comportamento di S_N , é dato da

$$B_N = aN^{1/\alpha}$$

di modo che

$$\lim_{N \rightarrow \infty} \text{Prob} \left(x < \frac{1}{aN^{1/\alpha}} S_N - A_N < x + dx \right) = L_{\alpha,\beta}(x) dx,$$

dove

$$\begin{aligned} 0 < \alpha < 1 : &= A_N = 0, \\ 1 < \alpha < 2 : &= A_N B_N = N \langle x \rangle . \end{aligned}$$

In particolare, si hanno $B_N = \sigma N^{1/2}$ e $A_N B_N = N \langle x \rangle$ per la Normale.

Riassunto, abbiamo visto il comportamento limite della distribuzione limite $L_{\alpha,\beta}(x)$ per $\alpha < 2$. Questo mostra la differenza tra la Gaussiana e le altre distribuzioni stabili. Considerato che tutte le densità di probabilità che vanno a zero abbastanza rapidamente per grande x (almeno $|x|^{-3}$) appartengono al dominio di attrazione della Gaussiana; le distribuzioni stabili con $\alpha < 2$ attirano solamente quelle $p(x)$ che hanno lo stesso comportamento asintotico per x grande. Questa condizione restrittiva è la ragione per la prevalenza della Gaussiana nel mondo naturale.

Capitolo 3

Oltre il Limite Centrale: gli sviluppi asintotici

3.1 Funzioni generatrici per \bar{S}_N

Fino ad ora abbiamo studiato la distribuzione della media S_N senza aver accennato alle approssimazioni. Questo problema che stiamo introducendo non appartiene solo ad esercizi teorici, ma si presta anche al fine operativo di semplificare i conti. Succede infatti che, quando si studiano fenomeni aleatori in cui molti fattori sono presenti, questi fattori possono avere un'influenza di tipo additiva: cioè si sommano, contribuendo ciascuno a una piccola perturbazione. Queste approssimazioni sono più frequenti di quanto uno si aspetta. La Legge dei grandi numeri dice, ad esempio, che se esiste l'aspettazione finita $\langle X \rangle$, allora per $n \rightarrow \infty$

$$\frac{S_n}{n} \rightarrow \langle X \rangle \quad \text{quasi sicuramente.}$$

In questo senso, si può vedere il valore $n \langle X \rangle$ come *l'approssimazione del primo ordine* della somma S_n . Il Limite Centrale, come visto precedentemente,

$$\bar{S}_n = \frac{(S_n - n \langle X \rangle)}{\sqrt{n\sigma^2}} \sim N(0, 1) \quad \text{per } n \rightarrow \infty$$

dice che \bar{S}_n si distribuisce come una normale standard. Quindi, si può chiamare la quantità

$$n \langle X \rangle + \bar{S}_n \sqrt{n\sigma^2}$$

l'approssimazione di secondo ordine di S_n . In altre parole, per deviazioni della forma $x = n \langle X \rangle + v\sqrt{n\sigma^2}$,

$$P(S_n \geq x) \sim P(\bar{S}_n \geq (x - n \langle X \rangle)/\sqrt{n\sigma^2}) = 1 - \phi(v) \quad \text{per } n \rightarrow \infty$$

uniformemente in $v \in \mathbb{R}$ e ϕ é la funzione di ripartizione della normale standard.

Quindi in questo capitolo vogliamo, seguendo la filosofia appena esposta, trovare questi 'approssimatori' di S_n e svilupparli asintoticamente per correggere errori di approssimazione dovuti alla convergenza in legge (quindi debole): parleremo dunque di momenti e di cumulanti.

Definiamo velocemente quindi i momenti e i cumulanti. I momenti possono essere ottenuti dalle derivate a 0 della funzione caratteristica $L_{\alpha,\beta}(k)$:

$$m_n = \widehat{D}^n L_{\alpha,\beta}(k) |_{k=0}$$

dove \widehat{D} é l'operatore differenziale complesso $\widehat{D} = -i\partial_k$. Inoltre si definiscono i cumulanti:

$$c_n = \widehat{D}^n \ln L_{\alpha,\beta}(k) |_{k=0}$$

Queste definizioni non risultano tuttavia applicabili: si preferisce l'approccio delle funzioni generatrici, che approssimano queste trasformate.

Sia X una variabile casuale che ammette funzione generatrice dei cumulanti $K_X(t)$ e siano X_1, \dots, X_n copie indipendenti di X . Sia $S_n = \sum X_i$ la variabile aleatoria somma e sia $\bar{S}_n = (S_n - A_n B_n)/B_n$ la variabile somma standardizzata. Come detto prima, se si tratta di variabili normali, $c_1 = A_1 B_1 = \langle X \rangle = \mu$ e $c_2 = B_2^2 = \sigma^2$. Per la standardizzazione, si ha $c_1(\bar{S}_n) = 0$ e $c_2(\bar{S}_n) = 1$. Inoltre i cumulanti di ordine superiore al secondo sono invarianti rispetto a traslazioni. Ricorsivamente, si ottiene per $r = 3, 4, \dots$

$$c_r(\bar{S}_n) = c_r \left(\frac{S_n}{\sqrt{n\sigma^2}} \right) = \frac{nc_r}{n^{r/2}\sigma^r} = n^{1-r/2}\rho_r$$

dove c_r e ρ_r sono rispettivamente i cumulanti e i cumulanti standardizzati di X . Si osservi che, essendo $c_2(\bar{S}_n) = 1$, si ha per $r = 3, 4, \dots$

$$\rho_r(\bar{S}_n)c_r(\bar{S}_n) = O(n^{-\frac{r}{2}+1})$$

All'aumentare di n quindi, i cumulanti di \bar{S}_n si avvicinano a quelli della distribuzione $N(0, 1)$, come ci aspettavamo. Per t fissato, vale lo sviluppo:

$$C_{\bar{S}_n}(t) = c_1(\bar{S}_n)t + c_2(\bar{S}_n)\frac{t^2}{2!} + c_3(\bar{S}_n)\frac{t^3}{3!} + c_4(\bar{S}_n)\frac{t^4}{4!} + O(c_5(\bar{S}_n))$$

da cui si ottiene lo sviluppo asintotico

$$C_{\bar{S}_n}(t) = \frac{1}{2}t^2 + \frac{\rho_3}{6\sqrt{n}}t^3 + \frac{\rho_4}{24n}t^4 + O(n^{-3/2}).$$

Si consideri poi la funzione generatrice dei momenti della variabile aleatoria \bar{S}_n . Per t fissato al divergere di n essa ammette lo sviluppo asintotico

$$M_{\bar{S}_n}(t) = e^{t^2/2} \left(1 + \frac{\rho_3}{6\sqrt{n}} t^3 + \frac{\rho_4}{24n} t^4 + \frac{\rho_3^2}{72n} t^6 + O(n^{-3/2}) \right). \quad (3.1)$$

Si ha infatti

$$M_{\bar{S}_n}(t) = \exp(C_{\bar{S}_n}(t)) = e^{t^2/2} \exp\left(\frac{\rho_3}{6\sqrt{n}} t^3 + \frac{\rho_4}{24n} t^4 + O(n^{-3/2})\right)$$

e, poiché $e^x = 1 + x + x^2/2 + O(x^3)$, si vede che dal termine $x^2/2$ sorge un termine non trascurabile rispetto al resto di ordine $O(n^{-3/2})$, ossia $(\rho_3 t^3 / (6\sqrt{n}))^2 / 2$. Al divergere per n fissato, la (3.1) fornisce la solita giustificazione del Teorema del Limite Centrale. Essa evidenzia anche, rispetto alla funzione generatrice dei momenti della Normale $N(0, 1)$, termini correttivi di ordine $O(n^{-1/2})$ e $O(n^{-1})$.

L'inversione richiesta é praticabile utilizzando i polinomi di Hermite, introdotti nella sezione successiva, e conduce alle espansioni di Edgeworth.

3.2 Polinomi di Hermite

Si consideri per $r = 0, 1, \dots$ la successione di funzioni $H_r(y)$ definita dalla relazione

$$H_r(y)\phi(y) = (-1)^r \frac{d^r}{dy^r} \phi(y)$$

ove $\phi(y)$ é la densità di una distribuzione normale $N(0, 1)$: si intende che la derivata di ordine zero di una funzione sia la funzione stessa. Per $r = 0$ si ha

$$H_0(y) = 1.$$

Inoltre, derivando ambo i membri, si ottiene la relazione ricorsiva

$$(-1)^{r+1} \frac{d^{r+1}}{dy^{r+1}} \phi(y) = H_{r+1}(y)\phi(y).$$

Pertanto le funzioni H_r soddisfano

$$H_{r+1}(y) = yH_r(y)\phi(y).$$

Essendo $H_0(y) = 1$ e $H'_0 = 0$ si mostra per induzione che $H_r(y)$ é un polinomio in y di grado r . Ricorsivamente, si ottengono i polinomi

$$\begin{aligned} H_1(y) &= y \\ H_2(y) &= y^2 - 1 \\ H_3(y) &= y^3 - 3y \\ H_4(y) &= y^4 - 6y^2 + 3 \\ &\dots \end{aligned}$$

Abbiamo introdotto questi polinomi per una notevole proprietà conseguente alla definizione: per $r \geq 1$ la funzione $H_r \phi(y)$ ha primitiva $-H_{r-1}(y)\phi(y)$. Infatti

$$-H_{r-1}\phi(y) = (-1)^r \int \left(\frac{d^r}{dy^r} \phi(y) \right) = \int H_r(y)\phi(y)dy. \quad (3.2)$$

La relazione (3.2) si può dimostrare che i polinomi di Hermite formano una successione di polinomi ortogonali rispetto al prodotto scalare con funzione di ponderazione $\phi(y)$. Precisamente per $r, s = 0, 1, \dots$ si ha

$$\int_{-\infty}^{\infty} H_r(y)H_s(y)\phi(y)dy = 0.$$

Il risultato chiave per l'inversione della funzione generatrice dei momenti della variabile aleatoria somma \bar{S}_n è l'identità

$$\int_{-\infty}^{\infty} e^{ty} H_r(y)\phi(y)dy = t^r e^{\frac{1}{2}t^2}, \quad (3.3)$$

valida per ogni r naturale.

Inoltre, l'insieme delle funzioni $\{H_r(y)\phi(y)\}$ forma una base per le densità standardizzate $f(y)$ di media 0 e varianza 1 quando $f(y)^2$ è integrabile. Vale infatti:

$$f(y) = \sum_r a_r H_r(y)\phi(y)$$

per delle costanti $\{a_r\}$ che rappresentano i coefficienti nella base dei polinomi di Hermite con la funzione di peso. Ogni coordinata si trova, come in Fourier, considerando il prodotto interno tra l'elemento $H_r(y)$ e $f(y)$:

$$a_r = \frac{1}{r!} \int H_r(y)f(y)dy.$$

3.3 Espansioni asintotiche di Edgeworth

Ora che abbiamo introdotto lo strumento dei polinomi di Hermite, possiamo studiare l'approssimazione della densità della variabile aleatoria somma standardizzata che si ottiene per inversione diretta della approssimazione (3.1) per la funzione generatrice dei momenti di S_n .

La relazione (3.3) consente di rappresentare lo sviluppo asintotico (3.1) di $M_{S_n}(t)$ nella forma integrale

$$M_{S_n}(t) = \int_{-\infty}^{\infty} e^{ty} \left(1 + \frac{\rho_3}{6\sqrt{n}} H_3(y) + \frac{\rho_4}{24n} H_4(y) + \frac{\rho_3^2}{72n} H_6(y) + O(n^{-3/2}) \right) \phi(y)dy.$$

Il confronto con la definizione

$$M_{S_n}(t) = \int_{-\infty}^{\infty} e^{ty} p_{S_n}(y) dy$$

ove $p_{S_n}(y)$ é la densitá di S_n , suggerisce per l'approssimazione

$$p_{S_n}(y) = p_{S_n}^E(y) + O(n^{-3/2}), \quad (3.4)$$

con

$$p_{S_n}^E(y) = \phi(y) \left(1 + \frac{\rho_3}{6\sqrt{n}} H_3(y) + \frac{\rho_4}{24n} H_4(y) + \frac{\rho_3^2}{72n} H_6(y) \right). \quad (3.5)$$

La (3.5) si chiama *formula o espansione di Edgeworth*. Essa é in sintonia con il risultato del primo ordine stabilito dalla versione locale del Teorema del Limite Centrale. La validitá dello sviluppo (3.4) viene dimostrata con l'assunzione che X ammetta finiti tutti i momenti fino al quinto, con anche l'ipotesi che la densitá $p_{S_n}(y)$ sia continua e limitata (o equivalentemente che la funzione caratteristica di S_n sia assolutamente integrabile). Si dimostra inoltre che l'errore assoluto nell'approssimazione puó essere maggiorato uniformemente in y da un termine $O(n^{-3/2})$.

Le espansioni di Edgeworth servono anche a calcolare anche una serie molto importante che collega funzioni caratteristiche e cumulanti.

3.4 Serie di Gram-Charlier di Tipo A

Abbiamo visto prima, per funzioni di densitá standardizzate f con media 0 e varianza unitaria, uno sviluppo in serie con base $\{H_r(y)\phi(y)\}$ e coefficienti $a_r = 1/r! \int H_r(y)f(y)dy$. Risulta interessante quando questi a_r vengono espressi nei momenti di f . Si trova che gli $\{a_r\}$ sono funzioni lineari dei primi r momenti di f , dove il momento r -esimo é definito come

$$m_r = \int y^r f(y) dy.$$

Sostituendo l'espressione lineare di a_r in termini di m_1, \dots, m_r , abbiamo la serie di Gram-Charlier di Tipo A per la densitá f della forma

$$f(y) = \phi(y) \left(1 + \frac{m_3}{6\sqrt{n}} H_3(y) + \frac{(m_4 - 3)}{24n} H_4(y) + \dots \right).$$

La serie di Gram-Charlier per densitá non standard $f(x)$ di media μ e varianza σ^2 si ottiene con la trasformazione $x = \sigma y + \mu$

$$f(y) = \frac{1}{\sigma} \phi(y) \left(1 + \frac{m_3}{6\sqrt{n}} H_3(y) + \frac{(m_4 - 3)}{24n} H_4(y) + \dots \right).$$

I momenti m_r sono ora i momenti standardizzati; cioè, data $X \sim N(0, 1)$, allora sono definiti come

$$m_r = \left\langle \frac{X - \mu}{\sigma} \right\rangle.$$

Sostituiamo ora i momenti standardizzati della serie di Gram-Charlier con i cumulanti associati ρ_i . I coefficienti a_1, \dots, a_r nella formula

$$f(y) = \sum_r a_r H_r(y) \phi(y)$$

si possono infatti riscrivere mediante i ρ_1, \dots, ρ_r . Per i primi quattro coefficienti si ha:

$$a_0 = 1$$

$$a_1 = 0$$

$$a_2 = 0$$

$$a_3 = \rho_3/6$$

$$a_4 = \rho_4/24$$

Quando i cumulanti sono conosciuti, si riottiene la serie di Edgeworth dalla serie di Gram-Charlier.

Capitolo 4

Correzioni asintotiche razionali

Riassumendo fino ad ora, abbiamo visto che, data la somma di $S_n = X_1 + \dots + X_n$ di variabili aleatorie normali standard di media nulla, allora il Teorema del Limite Centrale standard ci dice che

$$P\left(S_n/\sqrt{n} \leq x\right) = P(N(0, \sigma^2) < x) + o(1)$$

per $n \rightarrow \infty$. Le espansioni di Edgeworth forniscono una stima dell'errore, ma solo a condizione che esistano i momenti fino a un certo ordine. In particolare, se $\langle |X_k|^m \rangle < \infty$ per $m \geq 3$, l'espansione di Edgeworth di ordine $m - 2$ ha forma

$$P\left(S_n/\sqrt{n} \leq x\right) = P(N(0, \sigma^2) < x) + \sum_{j=1}^{m-2} n^{-j/2} G_j(x) + o(n^{(m-2)/2})$$

per $n \rightarrow \infty$, dove G_j é appropriatamente definita in termini di polinomi di Hermite e dei momenti. Come conseguenze, lo sviluppo tradizionale di Edgeworth non può andare oltre la $(m - 2)$ -esima correzione nel caso in cui il $m - 2 + 3 = m + 1$ esimo momento non é definito.

Il risultato principale di questo ultimo capitolo può essere espresso così: supponiamo che una successione di variabili aleatorie stabili $\{X_k\}$ abbia funzione di densità positiva p_X della forma:

$$p_X(x) = (1 + |x|)^{-(1+\alpha)} L(|x|),$$

dove L é una funzione a variazione lenta (cioé $L(xt)/L(x) \rightarrow 1$ per $x \rightarrow \infty$ per ogni $t > 0$). Sia $\phi(x)$ la densità normale standard. Allora la densità di S_n/\sqrt{n} si comporta asintoticamente come

$$p_{S_n/\sqrt{n}}(x) \sim \phi(x) \left(1 + \sum_{3 \leq j < \alpha} \frac{G_j(x)}{n^{j/2-1}}\right) + \begin{cases} \frac{L(n^{1/2}x)}{n^{\alpha/2-1}} G_\alpha(x) & x \text{ dispari} \\ \frac{\zeta_L(n^{1/2}x)}{n^{\alpha/2-1}} G_\alpha(x) & x \text{ pari} \end{cases}$$

per $n \rightarrow \infty$ con j pari, dove ζ_L é una funzione a varianza lenta e $G_\alpha(x)$ é scritta in termini di α . Vediamo che quindi viene aggiunto un nuovo termine correttivo oltre le espansioni asintotiche di Edgeworth. Questo risultato é stato trovato recentemente da Lam-Blanchet-Burch-Bazant, e noi vogliamo scriverlo attraverso uno sviluppo migliore, in modo che sia una formula facilmente implementabile come nell'Introduzione.

Incominciamo dunque. Sia data una successione di variabili aleatorie $\{X_k\}$ indipendenti e stabili, con media nulla e varianza unitaria. Scriviamo

$$\widehat{p}(z) = \langle e^{izX_k} \rangle$$

per la funzione caratteristica di X_k . Inoltre, al variabile aleatoria somma S_n associamo

$$\widehat{p}_n(z) = \langle \frac{zS_n}{n^{1/2}} \rangle = p^n\left(\frac{z}{n^{1/2}}\right).$$

Come detto, si vuole estendere lo sviluppo di Edgeworth dato dall'approssimazione ottenuta mediante la trasformata di Fourier di $S_n/n^{1/2}$. In particolare, si ha

$$n \log \widehat{p}(z/n^{1/2}) = -\frac{z^2}{2} - c_3 \frac{iz^3}{3!n^{1/2}} + \dots + c_m \frac{i^m z^m}{m!n^{m/2-1}} + o\left(\frac{z^m}{n^{m/2-1}}\right),$$

e quindi

$$\begin{aligned} \langle \exp(zS_n/n^{1/2}) \rangle &= \exp\left[n \log\left(z/n^{1/2}\right)\right] \\ &= \exp(-z^2/2) \left[\sum_{j=3}^m \frac{c_j (iz)^j}{n^{j/2-1}} + o\left(\frac{z^m}{n^{m/2-1}}\right) \right] \\ &= \exp(-z^2/2) \left(1 + \frac{c_3 (iz)^3}{n^{1/2}} + \dots \right) \end{aligned}$$

Per poter invertire l'espressione sopra, bisogna utilizzare la trasformata inversa di Fourier (anche detta antitrasformata) di $(iz)^k \exp(-z^2/2)$ che é

$$\mathcal{F}^{-1}[(iz)^k \exp(-z^2/2)] = \frac{d^k}{dx^k} \eta(x)$$

dove $\eta(x) = \exp(-x^2/\sqrt{2\pi})$. Ricorsivamente si ha

$$\frac{d^k}{dx^k} \eta(x) = \eta(x) (-1)^k H_k(x)$$

dove H_k é il k -esimo polinomio di Hermite. In particolare richiamiamo il fatto che $H_1(x) = x$, $H_2(x) = x^2 - 1$ e $H_3(x) = x^3 - 3x$. Quindi, facendo la trasformata inversa in entrambi i membri, si ottiene che la densit  $p_{S_n/n^{1/2}}$ di $S_n/n^{1/2}$ soddisfa

$$p_{S_n/n^{1/2}}(x) = \eta(x) + \eta(x) \sum_{k=3}^m n^{-k/2+1} P_k + o(n^{-m/2+1}),$$

dove P_k é un polinomio di grado k che dipende solo dai primi k cumulanti di X_j . In particolare, si ha che

$$P_3 = c_3 H_3/3!, \quad P_4 = c_3^2 H_3/72 + (c_4 - 3)H_4/4!$$

Ora conviene specificare la densit  p in una forma conosciuta che per  sia 'generale'. Poich  abbiamo visto nelle sezioni precedenti che la densit  di Cauchy appartiene al dominio di attrazione di distribuzioni stabili, possiamo sceglierla come rappresentante di distribuzioni simmetriche stabili (per le non-simmetriche, si veda [3] e [9]). Quindi, formalmente, abbiamo che nel caso pi  generale la densit  $p(\cdot)$ di X_k sia

$$p(x) = \frac{L(|x|)}{1 + |x|^{1+\alpha}}, \quad \text{con } x \in \mathbb{R}, \quad (4.1)$$

dove L é una funzione a variazione lenta, cio  vale

$$\frac{L(vx)}{x} \rightarrow 1 \quad \text{per } t \rightarrow \infty$$

e che inoltre soddisfi

$$L'(x) = o\left(\frac{L(x)}{x}\right).$$

Introduciamo poi anche la funzione

$$\zeta_L(x) = \int_1^x \frac{L(u)}{u} du$$

da cui segue

$$L(x) = o(\zeta_L(x)).$$

Vale allora il seguente:

Teorema 4 (Correzioni di Lam-Blanchet). *Sia $\{X_k\}$ una successione di variabili aleatorie stabili con varianza unitaria e funzione di densit  simmetrica definita da (4.1). Sia $\alpha > 2$. Allora*

$$p_{S_n/\sqrt{n}}(x) \sim \phi(x) \left(1 + \sum_{3 \leq j < \alpha} \frac{G_j(x)}{n^{j/2-1}} \right) + \begin{cases} \frac{L(n^{1/2}x)}{n^{\alpha/2-1}} G_\alpha(x) & x \text{ dispari} \\ \frac{\zeta_L(n^{1/2}x)}{n^{\alpha/2-1}} G_\alpha(x) & x \text{ pari} \end{cases} \quad (4.2)$$

per $n \rightarrow \infty$, dove $G_j(x)$ sono i coefficienti di Edgeworth e G_α é definita come

$$G_\alpha(x) = \begin{cases} -\frac{\sqrt{2}}{\Gamma(\alpha+1)} \frac{d^\alpha}{dx^\alpha} D(x/\sqrt{2}) & \text{per } \alpha \text{ dispari} \\ -\sqrt{\pi/2} \frac{e^{-x^2/4}}{\Gamma(\alpha+1) \text{sen}(\alpha\pi)} [D_\alpha(x) + D_\alpha(-x)] & \text{per } \alpha \text{ non-intero} \\ \sqrt{2/\pi} \frac{e^{-x^2/2}}{\Gamma(\alpha+1)} H_\alpha(x) & \text{per } \alpha \text{ pari} \end{cases} \quad (4.3)$$

dove $H_k(z)$ sono i polinomi di Hermite di ordine k , $D(z)$ sono gli integrali di Dawson

$$D(z) = e^{-z^2} \int_0^z e^{-t^2} dt \quad (4.4)$$

e $D_\alpha(z)$ é la funzione parabolica cilindrica classica

$$D_\alpha(z) = \sqrt{\frac{2}{\pi}} e^{-z^2/4} \int_0^\infty e^{-t^2/2} t^\alpha \cos(zt - \alpha\pi/2) dt. \quad (4.5)$$

La dimostrazione si trova in [9]. Quello però che qui si vuol fare é approssimare il termine correttivo G_α (4.3), in modo che la formula asintotica del Teorema possa essere facilmente implementabile e non scritto in termini di $D(z)$ e $D_\alpha(z)$.

- Caso α dispari. Per questo primo caso, dobbiamo sviluppare il termine

$$-\frac{\sqrt{2}}{\Gamma(\alpha+1)} \frac{d^\alpha}{dx^\alpha} D(x/\sqrt{2})$$

Ora, quindi si consideri l'approssimazione troncata in serie, ovvero

$$D(z) = f_n(z) + r_n(z)$$

con

$$f_n(z) = \frac{1}{2z} \left(1 + \frac{1}{2z^2} + \frac{1 \cdot 3}{(2z^2)^2} + \dots + \frac{1 \cdot 3 \cdot \dots \cdot (2n-1)}{(2z^2)^n} \right),$$

$$r_n(z) = ie^{-z^2} \left((-1)^{n+1} \frac{1 \cdot 3 \cdot \dots \cdot (2n+1)}{2^{n+1}} \int_{iz}^\infty \frac{e^{-\zeta^2/2}}{\zeta^{2n+2}} d\zeta - \frac{\sqrt{\pi}}{2} \right).$$

Si trova quindi il comportamento asintotico, per $x \rightarrow \infty$:

$$-\frac{\sqrt{2}}{\Gamma(\alpha+1)} \frac{d^\alpha}{dx^\alpha} D(x/\sqrt{2}) \sim \frac{\sqrt{2}}{\Gamma(\alpha+1)} (-1)^{\alpha+1} \frac{\alpha!}{\sqrt{2} x^{1+\alpha}} \quad (4.6)$$

- Caso α non intero. In questo secondo e ultimo caso, vogliamo una somma asintotica che sviluppi la funzione parabolica cilindrica classica D_α in (4.5). Utilizzando i risultati in [11], si ha che

$$D_\alpha(z) = e^{-z^2/4} \left(\frac{\alpha(\alpha+1)}{2z^2} + \frac{\alpha(\alpha-1)(\alpha-2)(\alpha-3)}{2 \cdot 4z^4} - \dots \right) \quad (4.7)$$

da cui si trova il comportamento al limite

$$D_\alpha(z) \sim \begin{cases} z^\alpha e^{-z^2/4} & \text{per } z \rightarrow +\infty \\ \frac{\sqrt{2\pi}}{\Gamma(\alpha)} e^{-z^2/4} |z|^{-\alpha-1} & \text{per } z \rightarrow -\infty \end{cases} \quad (4.8)$$

Quindi, in definitiva, abbiamo trovato che, per α dispari e non intero,

$$G_\alpha(x) \sim x^{-(1+\alpha)}, \quad \text{per } x \rightarrow \infty,$$

in quanto il caso non intero segue dalla proprietà di riflessione della funzione gamma e in quanto la parte negativa di D_α domina in $G_\alpha(x)$. Abbiamo dunque ritrovato la misura simmetrica di Levy! Cioè il termine correttivo che abbiamo sviluppato è proprio la frequenza dei salti di lunghezza x in una passeggiata aleatoria (*Weierstrass Random Walk*, [4]).

Quindi, andando ad inserire la correzione, per α non pari si ha:

$$G_\alpha(x) \frac{L(n^{1/2}x)}{n^{\alpha/2-1}} \sim \frac{L(n^{1/2})}{x^{1+\alpha} n^{\alpha/2-1}} \quad (4.9)$$

per $x \rightarrow \infty$.

Ma si può andare oltre! Il termine generale in tutto questo discorso è la funzione L a variazione lenta. Per questa classe di funzioni vale una importante caratterizzazione:

Proposizione 5 (Rappresentazione integrale). *Sia L una funzione a variazione lenta, allora per $x_0 < x$ si ha*

$$L(x) = c \cdot \exp \left\{ \int_{x_0}^x \frac{\varepsilon(u)}{u} du \right\}$$

dove c è una costante e $\varepsilon(x)$ è una funzione misurabile che tende a zero per x crescente.

In forza delle Proposizione, la funzione

$$L(x) = \ln x$$

è una funzione a variazione lenta (basta prendere $c = 1$, $x_0 = e$ e $\varepsilon(x) = (\ln x)^{-1}$). Ma non è solo una funzione appartenente alla classe: la possiamo scegliere come rappresentante, in forza delle proprietà del logaritmo.

Succede infatti che, con questa scelta,

$$L(n^{1/2}) = \frac{1}{2} \ln n,$$

e concludendo si ottiene da (4.9):

$$G_\alpha(x) \frac{L(n^{1/2}x)}{n^{\alpha/2-1}} \sim \frac{1}{2} \frac{\ln n}{x^{1+\alpha} n^{\alpha/2-1}} \quad (4.10)$$

Vediamo allora il valore dello sviluppo asintotico della correzione trovato ora per la prima volta, al variare di α .

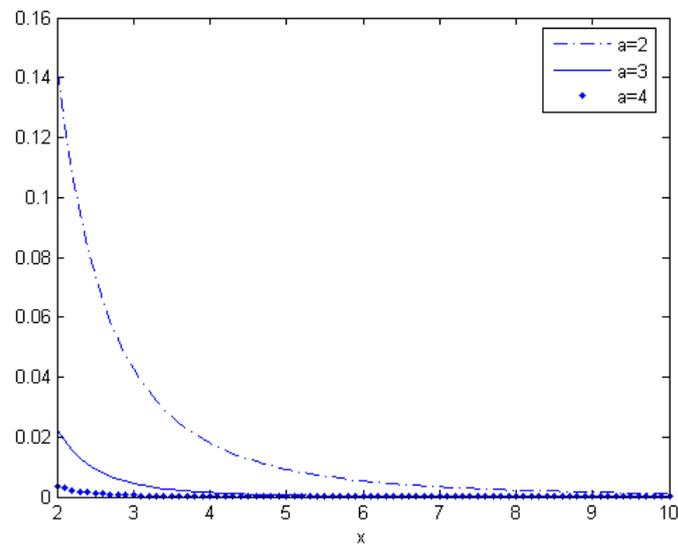


Figura 4.1: Correzione per $n=10$

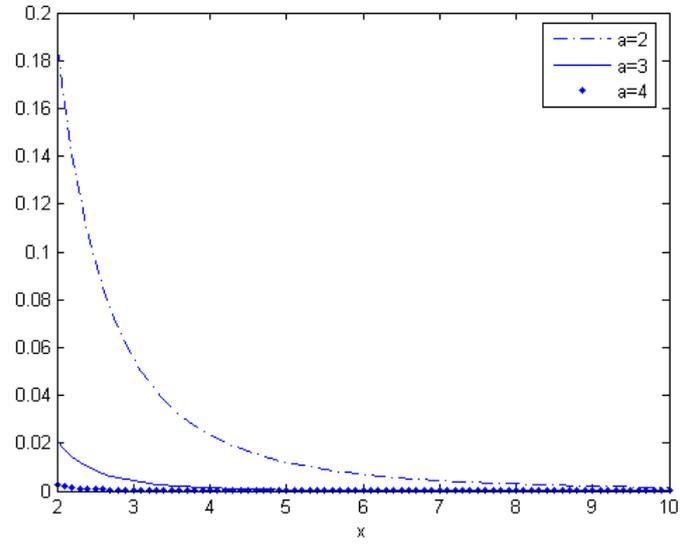


Figura 4.2: Correzione per n=20

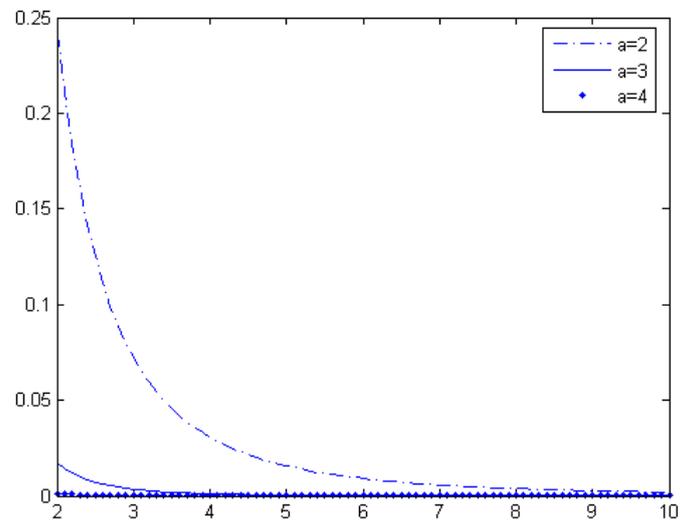


Figura 4.3: Correzione per n=50

4.1 Espansioni classiche per $F_n(x)$

Il risultato trovato riguarda la densità della distribuzione della somma S_n/\sqrt{n} . Ora vogliamo arrivare ad una espansione per la funzione di ripartizione.

Sia allora X_1, X_2, \dots una successione di variabili aleatorie iid di media nulla e con momento $(2 + \delta)$ esimo finito, con $\delta > 0$. Poniamo allora

$$B_n = \sum_{k=1}^n \langle X_k^2 \rangle, \quad F_n(x) = P(S_n < xB_n\sqrt{n}),$$

$$L_n = \frac{\sum |X_k|^{2+\delta}}{B_n^{-(1+\delta/2)}}, \quad \phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp(-t^2/2) dt.$$

Il teorema di Lyapunov stabilisce che, se $L_n \rightarrow 0$ per $n \rightarrow \infty$, allora

$$F_n(x) \rightarrow \phi(x)$$

per ogni x . Osserviamo che, in particolare, se $\langle X^2 \rangle < \infty$, allora si riottiene il teorema di Levy

$$P(S_n < \sigma x \sqrt{n}) \rightarrow \phi(x).$$

Viene naturale però farsi domande sulla velocità di convergenza: la risposta sta nelle *formule di Barry-Essen*, che stabiliscono che

$$\sup_x |P\{S_n/\sigma\sqrt{n} < x\} - \phi(x)| \leq \frac{A \cdot \langle X_1^3 \rangle}{\sigma^3 \sqrt{n}}$$

con A costante positiva. Questa formula fornisce quindi l'ordine di convergenza della funzione di ripartizione della somma standardizzata verso la normale. Notiamo che è la prima volta nella Tesi che si pone l'accento su questo ordine di convergenza.

C'è la possibilità di andare ad accelerare la convergenza con un termine correttivo dipendente dai momenti della prima variabile della successione? Sì! Vale infatti il seguente risultato, trovato nel 1966:

Teorema 5 (di Osipov-Petrov). *Sia X_1, \dots, X_n una successione di variabili aleatorie iid, con $\langle X_1 \rangle = 0$, $\langle X_1^2 \rangle = \sigma^2$ e $\langle X_1^r \rangle < \infty$. Allora*

$$|F_n(x) - \phi(x)| \leq \frac{C}{1 + |x|^r} \left(\frac{\langle |X_1|^3 \rangle}{\sigma^3 \sqrt{n}} + \frac{\langle |X_1|^r \rangle}{\sigma^r n^{(r-2)/2}} \right)$$

per ogni x e con C costante positiva.

Appendice A

Approssimanti di Padé

Ora che abbiamo scritto la correzione degli approssimanti di Edgeworth, vogliamo riscriverla in un modo semplificato e piú trattabile. Succede infatti che

$$G_\alpha(x) \frac{L(n^{1/2}x)}{n^{\alpha/2-1}} \sim \frac{1}{2} \frac{\ln n}{x^{1+\alpha} n^{\alpha/2-1}}$$

dipende da x , dal numero di variabili aleatorie sommate n e dall'esponente caratteristico α , che ci dice se le variabili sono stabili o meno.

Il problema ora é quindi riscrivere meglio il termine

$$\frac{1}{2} \frac{\ln n}{n^{\alpha/2-1}}$$

che, essendo $0 < \alpha < 2$ non necessariamente intero, presenta un'esponente frazionario. Alla luce di questa considerazione, ha senso introdurre un metodo di approssimazione razionale implementabile anche in ambiente Matlab: parliamo quindi degli approssimanti di Padé (`pade(f, x = x0, <[m, n]>)`).

L'approssimazione di Padé é una alternativa semplice ed efficace all'approssimazione polinomiale per funzioni analitiche.

Sia $f(z)$ una funzione analitica attorno a $z = 0$, possiamo espanderla in una serie di Taylor

$$f(z) = \sum_{n=0}^{\infty} a_n z^n = a_0 + a_1 z + a_2 z^2 + \dots \quad (\text{A.1})$$

che converge sul disco $|z| < R$. Supponiamo di voler conoscere il valore della funzione f in un punto z_1 fuori dal disco di convergenza ma dove f é analitica. La serie di Taylor é inutile in quanto diverge, essendo fuori dal disco. Ma dalla teoria delle funzioni analitiche, sappiamo che $f(z_1)$ é completamente determinata dai coefficienti a_0, a_1 , ecc. In altre parole, data un'espansione in serie di potenze, vogliamo trovare uno schema di approssimazione compatibile con la continuazione analitica al di fuori del raggio di

convergenza della serie. Gli approssimanti di Padé soddisfano questa richiesta. Definiamo approssimante di Padé $[M, N]$ di $f(z)$ il rapporto di due polinomi in z , $P_M(z)$ e $Q_N(z)$ di grado rispettivamente M e N , che ha le stesse prime $M + N$ derivate di $f(z)$ in $z = 0$. Scriviamo

$$f(z)_{[M,N]} = \frac{P_M(z)}{Q_N(z)} = f(z) + O(z^{M+N+1}) \quad (\text{A.2})$$

Da questo punto di vista, l'insieme degli approssimanti di Padé é una generalizzazione dell'espansione in serie di Taylor; infatti, gli approssimanti $[M, 0]$ coincidono con le somme parziali di (A.1).

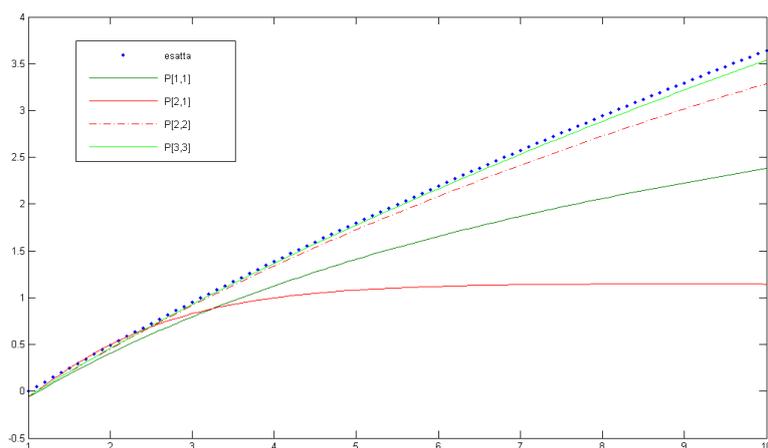
Gli approssimanti di Padé sono costruiti con i coefficienti della serie di Taylor. Per poterla estendere al di fuori del suo disco di convergenza, gli approssimanti costituiscono un metodo di somma per serie divergenti (cfr. [13]).

Equivalentemente la (A.2) si può riscrivere

$$Q_N(z)f(z) - P_M(z) = O(z^{M+N+1})$$

da cui segue che i coefficienti dei polinomi $P_M(z)$ e $Q_N(z)$ sono determinati da un'insieme di equazioni lineari a partire dai primi $M + N + 1$ coefficienti della serie di Taylor, a meno di una costante. Segue che l'approssimante di Padé $[M, N]$ é unico. Gli approssimanti diagonali di Padé sono quelli della forma $[M, M]$.

Abbiamo mostrato diversi approssimanti dei coefficienti della correzione (4.10) in n :



A.1 Proprietá omografiche

La domanda é: quali sono le proprietá invarianti per trasformazioni che sono comuni a $f(z)$ e ai suoi approssimanti di Padé? Per rispondere, consideriamo le trasformazioni omografiche. Una matrice reale unimodulare del secondo ordine definisce una applicazione omografica

$$z \longrightarrow \frac{az + b}{cz + d}, \quad ad - bc = 1.$$

del semipiano superiore complesso in se stesso. Geometricamente, questa applicazione é uno spostamento del piano di Lobacevskij, rappresentato dal semipiano superiore. Per la sua trattazione algebrica, si veda [14]).

Consideriamo ora la trasformazione degli approssimanti diagonale $f_{[M,M]}(z)$ di $f(z)$:

$$R(z) = \frac{af_{[M,M]} + b}{cf_{[M,M]} + d} = \frac{aP_{[M]}(z) + bQ_{[M]}(z)}{cP_{[M]}(z) + dQ_{[M]}(z)}.$$

É ovvio che:

- $R(z)$ é il rapporto di due polinomi di grado M ;
- ha le prime $2M$ derivate della funzione trasformata $f(z) \rightarrow (af(z) + b)/(cf(z) + d)$.

Allora, poiché l'approssimante di Padé é unico, gli approssimanti diagonali della funzione trasformata sono la trasformata degli approssimanti diagonali della funzione. Che gli approssimanti di Padé mantengano le proprietá delle trasformazioni omografiche della funzione, é di fondamentale importanza e spiega molte proprietá:

- i coefficienti degli approssimanti si possono trovare in forma ricorsiva;
- i poli di $f(z)$ non devono essere considerati punti speciali per gli approssimanti di Padé (al contrario di Taylor), poiché si puó pensare di regolarizzarli attraverso opportune trasformazioni omografiche. In particolare, gli approssimanti di Padé sommano esattamente le serie geometriche.

Putroppo però, non sempre conviene scegliere gradi alti dei polinomi $P_{[M]}(z)$ e $Q_{[M]}(z)$ per avere una stima dell'errore minima. Succede infatti che, oltre ad aumentare gli zeri del numeratore P , aumentano anche gli zeri del denominatore Q , che sono poi i poli. Viceversa, non conviene nemmeno aumentare M e lasciare N basso: i polinomi, ad alti gradi, oscillano.

Al di lá del discorso numerico, dove vengono ricercati zeri e poli sull'asse reale del piano complesso \mathbb{C} , é interessante indagare la distribuzione degli zeri su \mathbb{C} perché sappiamo da sopra che gli approssimanti sono invarianti per omografie, che sono automorfismi del semipiano superiore in sé. Veniamo quindi al conclusivo

Esempio 2 (Zeri degli approssimanti di Padé per l'esponenziale). Sia $f(z) = e^z$. Andiamo a vedere il luogo degli zeri per $0 < M, N < 15$. Si ottengono con metodi numerici le seguenti figure

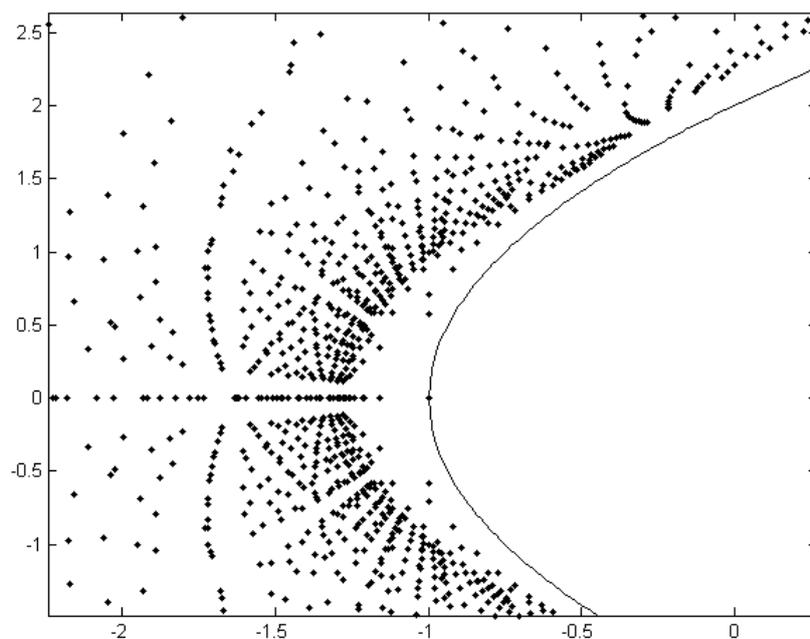
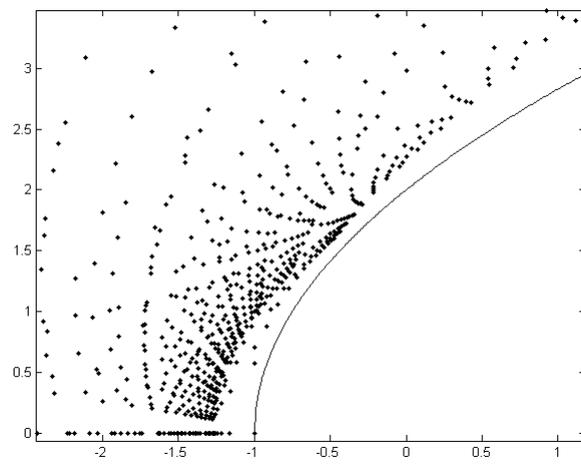
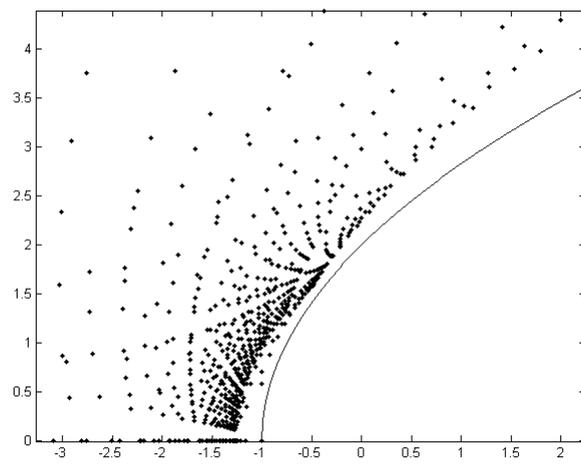
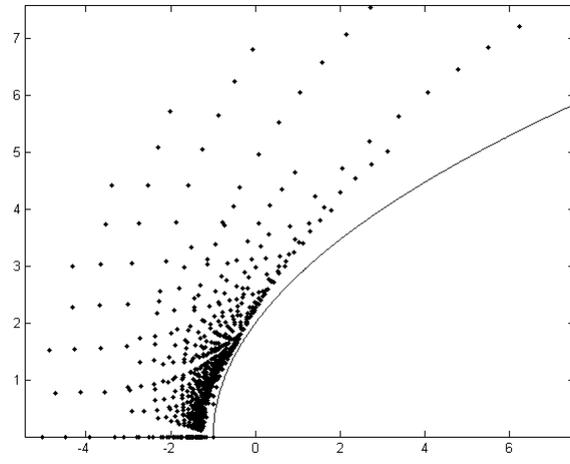


Figura A.1: Zeri su \mathbb{C} . Per il teorema di Saff-Varga, non ci sono zeri oltre la regione parabolica $Im(z)^2 \leq 4(Re(z) + 1)$, $Re(z) > -1$.



Bibliografia

- [1] C. Donnelly e P. Embrechts, The devil is in the tails: actuarial mathematics and the subprime mortgage crisis, *Astin Bulletin*, 2010
- [2] Vincenzo Grecchi, *Appunti di Matematica per le applicazioni economiche e finanziarie*
- [3] G. Samorodnitsky e M.S. Taqqu, *Stable non-Gaussian Random Processes*, Cambridge Uni Press, 1997
- [4] J.P. Bouchaud e M. Potters, *Theory of Financial Risks*, Cambridge Uni Press, 2000
- [5] J.P. Nolan, *Stable Distributions - Models for Heavy Tailed Data*, Birkhauser, 2013
- [6] M. Campanino e F. Biagini, *Elementi di Probabilità e Statistica*, Springer, 2005
- [7] W. Paul e J. Baschnagel, *Stochastic Processes: from physics to finance*, Springer Berlin, 1999
- [8] L. Pace e A. Salvan, *Teoria della Statistica*, Cedam, Padova, 1996
- [9] H. Lam, J. Blanchet, D. Burch e M. Bazant, *Corrections to the Central Limit Theorem for Heavy-tailed Probability Densities*, arXiv.org, 2011
- [10] F. Blomquist, *Interval inclusions for Dawson's Integral*, *Interval Computations* No(3), 1992
- [11] P.F. Daley, *The parabolic cylinder function in elastodynamic problems*, *PCF in elastodynamic problems*, 2001
- [12] L.V. Osipov e V.V. Petrov, *On an Estimate of the Remainder Term in the Central Limit Theorem*, *Theory of Probability and its Application*, 1966
- [13] J.L. Basdevant, *The Padé Approximation and its physical applications*, CERN, Geneve, 1971
- [14] V.I. Arnold, *Metodi geometrici della teoria delle equazioni differenziali ordinarie*, Editori Riuniti, Roma, 2010