

Distribuzioni stabili ed infinitamente divisibili:
teoria ed applicazioni

Luca Carpanese

Indice

1	Distribuzioni infinitamente divisibili	7
1.1	Introduzione	7
1.2	Definizioni di infinita divisibilità	7
1.2.1	Esempi di distribuzioni infinitamente divisibili	8
1.2.2	Formula di Lévy-Khintchine	9
1.3	Distribuzione Compound Poisson	11
1.3.1	Funzione caratteristica di Compound Poisson	11
1.3.2	Convergenza di Compound Poisson	13
2	Distribuzioni stabili	18
2.1	Introduzione	18
2.2	Definizioni di stabilità	19
2.2.1	Proprietà di stabilità	19
2.2.2	Determinazione di c_n	20
2.3	Vettore di parametri	24
2.4	Impiego della funzione caratteristica	25
2.4.1	Dimostrazione del Teorema 2.4.1	25
2.5	Assoluta continuità delle distribuzioni stabili	28
2.6	Distribuzioni stabili con densità esplicita	28
2.7	Momenti di variabili aleatorie stabili	29

2.8	Comportamento delle code nelle distribuzioni stabili	30
2.9	Teorema centrale del limite generalizzato	31
2.10	Grafici	33
3	Procedure di stima dei parametri	35
3.1	Metodo della funzione caratteristica empirica	36
3.1.1	Metodo dei momenti	36
3.1.2	Metodo regressivo	37
3.2	Metodi di stima alternativi	38
3.2.1	Metodo dei quantili	38
3.2.2	Metodo della massima verosimiglianza	40
3.3	Considerazioni sui metodi di stima	41
3.4	Applicazioni in R	41

Introduzione

Negli anni più recenti, la diffusione di dati caratterizzati da *code pesanti* in numerose branche dell'attività umana ha richiesto l'impiego di teorie matematiche adatte alla loro modellizzazione. La distribuzione normale, le cui proprietà sono ampiamente studiate nella letteratura, si rivela spesso inadeguata a tale scopo. A queste carenze sopperiscono particolari distribuzioni di probabilità, dette *distribuzioni stabili*, caratterizzate da quattro parametri, che consentono di descrivere adeguatamente fenomeni che spaziano dall'economia all'astrofisica, dall'idrologia alle telecomunicazioni includendo l'epidemiologia e la finanza, ottenendo modelli impiegabili sia in ottica descrittiva che previsiva.

Nel presente lavoro esamineremo la teoria matematica alla base di questa famiglia di distribuzioni, sottolineando l'appartenenza delle *stabili* alla famiglia delle distribuzioni *infinitamente divisibili*, tentando di offrirne una trattazione unificata ed esaustiva. Saranno forniti infine esempi di applicazioni in ottica inferenziale per mostrare l'elevata potenzialità di esse; il lettore interessato ai numerosi e rilevanti impieghi statistici delle distribuzioni stabili ed infinitamente divisibili troverà in bibliografia alcune monografie specifiche tramite le quali esplorarne le potenzialità. Per una panoramica storica sull'origine di tali leggi di probabilità si rimanda a Mainardi e Rogosin (2008).

Capitolo 1

Distribuzioni infinitamente divisibili

1.1 Introduzione

Introdotta nel 1929 dal probabilista italiano Bruno de Finetti, interessato alle distribuzioni di probabilità sottese ai processi stocastici a tempo continuo con incrementi stazionari ed indipendenti, la famiglia delle distribuzioni infinitamente divisibili può essere definita equivalentemente mediante approcci distinti.

1.2 Definizioni di infinita divisibilità

Definizione 1.2.1. Una variabile aleatoria X è detta *infinitamente divisibile* se, per ogni $n \in \mathbb{N}$, esistono n variabili aleatorie X_1, \dots, X_n tra loro *i.i.d.* tali che

$$X \stackrel{d}{=} X_1 + \dots + X_n \tag{1.1}$$

Osservazione 1.2.1. Nota la distribuzione di X e stabilito n , è possibile determinare la distribuzione di ciascun addendo X_j ricorrendo alla funzione caratteristica $\phi_X(t)$. Dalla (1.1) si ha

$$\mathbb{E}[e^{itX}] = \mathbb{E}[e^{it(X_1 + \dots + X_n)}]$$

Sfruttando l'indipendenza e la identica distribuzione si ottiene

$$\mathbb{E}[e^{itX}] = (\mathbb{E}[e^{itX_1}])^n$$

Quindi la funzione caratteristica $\phi_{X_j}(t)$ della singola X_j è pari a

$$\phi_{X_j}(t) = [\phi_X(t)]^{\frac{1}{n}} \quad (1.2)$$

Da quanto osservato deriva la seguente definizione, equivalente alla 1.2.1

Definizione 1.2.2. Una variabile aleatoria X è infinitamente divisibile se e solo se, per ogni $n \in \mathbb{N}$, $[\phi_X(t)]^{\frac{1}{n}}$ è la funzione caratteristica di una variabile aleatoria.

1.2.1 Esempi di distribuzioni infinitamente divisibili

Mediante la 1.2.1 è possibile fornire alcuni esempi di distribuzioni note che godono della proprietà di infinita divisibilità.

Esempio 1.2.1. Distribuzione degenera

La variabile aleatoria $X = t$ è infinitamente divisibile in quanto, per ogni $n \in \mathbb{N}$, la (1.1) assume la forma

$$t \stackrel{d}{=} \frac{t}{n} + \frac{t}{n} + \cdots + \frac{t}{n}$$

Esempio 1.2.2. Distribuzione normale

Ai fini di dimostrare l'infinita divisibilità della distribuzione normale, ricordiamo che, data una famiglia $\{X_i\}_{i=1}^n$ di variabili aleatorie indipendenti ciascuna distribuita come una $\mathcal{N}(\mu_i, \sigma_i^2)$, allora la variabile aleatoria $X = \sum_{i=1}^n X_i$ è tale che

$$X \sim \mathcal{N}\left(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2\right)$$

Considerando quindi che $X \sim \mathcal{N}(\mu, \sigma^2)$, essa può essere espressa come somma di n variabili aleatorie X_i normali indipendenti ciascuna con valore atteso $\frac{\mu}{n}$ e varianza $\frac{\sigma^2}{n}$.

Esempio 1.2.3. Distribuzione di Poisson

Analogamente a quanto svolto per la distribuzione normale, ricordiamo che, data una famiglia $\{X_i\}_{i=1}^n$ di variabili aleatorie indipendenti distribuite secondo una $Poi(\lambda_i)$, allora la variabile aleatoria $X = \sum_{i=1}^n X_i$ è tale che

$$X \sim Poi\left(\sum_{i=1}^n \lambda_i\right)$$

Nuovamente è possibile affermare che, data $X \sim Poi(\lambda)$, la (1.1) è valida ponendo $X_i \sim Poi\left(\frac{\lambda}{n}\right)$ per ogni $i = 1, \dots, n$ e da ciò segue l'infinita divisibilità.

1.2.2 Formula di Lévy-Khintchine

La Formula di Lévy-Khintchine, dai nomi di coloro che la determinarono durante gli anni trenta del ventesimo secolo, coadiuvati da notevoli contributi di Bruno de Finetti e Andrej Kolmogorov, rappresenta un utile strumento nella caratterizzazione delle variabili aleatorie infinitamente divisibili. Essa consente infatti di definire la famiglia delle distribuzioni infinitamente divisibili mediante la loro funzione caratteristica.

Teorema 1.2.1. Formula di Lévy-Khintchine

Una variabile aleatoria X a valori in \mathbb{R} è infinitamente divisibile se e solo se ha funzione caratteristica data da

$$\mathbb{E}[e^{itX}] = \exp\left\{itb - \frac{1}{2}\sigma^2 t^2 + \mathcal{B}(t)\right\}$$

in cui la funzione $\mathcal{B}(t)$ è definita come

$$\mathcal{B}(t) = \int_{-\infty}^{+\infty} (e^{itx} - itx\mathbb{I}(x)_{[-1,1]} - 1) \nu(dx)$$

per opportuni $b \in \mathbb{R}$, $\sigma^2 \in \mathbb{R}^+$ ed una misura $\nu(dx)$ detta *misura di Lévy* che deve soddisfare

$$\int_{-\infty}^{+\infty} \min(1, |x|^2) \nu(dx) < +\infty$$

Dimostrazione. Si veda Applebaum (2009), pg. 30. □

Osservazione 1.2.2. La tripletta (b, σ^2, ν) identifica univocamente la distribuzione di probabilità e si definisce *tripletta di Lévy*.

Esempi

Esempio 1.2.4. Variabile degenera

Sia $X = b$ una variabile degenera.

$$\mathbb{E}[e^{itX}] = e^{itb}$$

Da questo segue che la tripletta di Lévy è pari a $(b, 0, 0)$.

Esempio 1.2.5. Variabile normale

Data $X \sim \mathcal{N}(b, \sigma^2)$, la sua funzione caratteristica sarà pari a

$$\mathbb{E}[e^{itX}] = e^{itb - \frac{1}{2}\sigma^2 t^2}$$

da cui si deduce che la tripletta di Lévy corrisponde a $(b, \sigma^2, 0)$.

1.3 Distribuzione Compound Poisson

Definiamo una variabile aleatoria data dalla somma di un numero stocastico di addendi *i.i.d.* che vedremo essere infinitamente divisibile.

Definizione 1.3.1. Variabile aleatoria Compound Poisson

Dati $N \sim \text{Poi}(\lambda)$ ed una sequenza $\{X_j\}_{j=1}^{+\infty}$ di variabili aleatorie *i.i.d.* e indipendenti da N , si definisce *variabile Compound Poisson* la quantità aleatoria

$$X = \sum_{j=1}^N X_j$$

1.3.1 Funzione caratteristica di Compound Poisson

Si consideri la variabile Compound Poisson

$$S_N = \sum_{j=1}^N X_j$$

La sua funzione di ripartizione è pari a

$$P(S_N \leq x) = \sum_{n=0}^{+\infty} P\left(\sum_{j=1}^N X_j \leq x \mid N = n\right) \cdot P(N = n)$$

Dal fatto che $N \sim \text{Poi}(\lambda)$ segue

$$P(S_N \leq x) = \sum_{n=0}^{+\infty} P\left(\sum_{j=1}^n X_j \leq x\right) \cdot e^{-\lambda} \frac{\lambda^n}{n!}$$

Derivando tale quantità si ottiene la funzione di densità

$$f_{S_N}(x) = \sum_{n=0}^{+\infty} f_{S_n}(x) \cdot e^{-\lambda} \frac{\lambda^n}{n!}$$

La funzione caratteristica di X sarà dunque

$$\mathbb{E}[e^{itS_N}] = \int_{-\infty}^{+\infty} e^{itx} \sum_{n=0}^{+\infty} f_{S_n}(x) \cdot e^{-\lambda} \frac{\lambda^n}{n!} dx$$

$$\begin{aligned}
&= \sum_{n=0}^{+\infty} \left(\int_{-\infty}^{+\infty} e^{itx} f_{S_n}(x) \cdot e^{-\lambda} \frac{\lambda^n}{n!} dx \right) \\
&= \sum_{n=0}^{+\infty} \mathbb{E}[e^{itS_N}] \cdot e^{-\lambda} \frac{\lambda^n}{n!}
\end{aligned}$$

Data la natura *i.i.d.* delle variabili in S_n

$$\mathbb{E}[e^{itS_n}] = (\mathbb{E}[e^{itX_1}])^n$$

Ciò implica che

$$\begin{aligned}
\mathbb{E}[e^{itS_N}] &= \sum_{n=0}^{+\infty} [\phi_{X_1}(t)]^n \cdot e^{-\lambda} \frac{\lambda^n}{n!} \\
&= e^{-\lambda} \sum_{n=0}^{+\infty} \frac{\lambda [\phi_{X_1}(t)]^n}{n!}
\end{aligned}$$

ed infine

$$\phi_{S_N}(t) = e^{-\lambda + \lambda \phi_{X_1}(t)} \quad (1.3)$$

dove $\phi_{X_1}(t)$ corrisponde alla funzione caratteristica di X_1 . In particolare, restringendoci al caso in cui X_1 sia assolutamente continua con densità f

$$\phi_{X_1} = \int_{-\infty}^{+\infty} e^{itx} f(x) dx$$

da cui segue

$$\mathbb{E}[e^{itS_N}] = \exp \left\{ -\lambda + \lambda \int_{-\infty}^{+\infty} e^{itx} f(x) dx \right\} \quad (1.4)$$

Tramite semplici manipolazioni la (1.4) diviene

$$\mathbb{E}[e^{itS_N}] = \exp \left\{ \int_{-\infty}^{+\infty} (e^{itx} - 1) \lambda f(x) dx \right\}$$

e ponendo $\lambda f(x) = \nu(x)$, dove $\nu(x)$ rappresenta la densità di Lévy, si ricava

$$\mathbb{E}[e^{itS_N}] = \exp \left\{ \int_{-\infty}^{+\infty} (e^{itx} - 1) \nu(x) dx \right\}$$

Inoltre, aggiungendo e sottraendo la quantità $itx\mathbb{I}(y)_{[-1,1]}$ e svolgendo semplici calcoli si giunge all'equazione

$$\mathbb{E}[e^{itS_N}] = \exp \left\{ \int_{-\infty}^{+\infty} (e^{itx} - 1 - itx\mathbb{I}(x)_{[-1,1]}) \nu(x) dx + it \int_{-1}^1 x \nu(x) dx \right\} \quad (1.5)$$

Osservazione 1.3.1. Dalla (1.3) segue l'infinita divisibilità della Compound Poisson. Infatti, sfruttando la (1.2), è possibile trovare la funzione caratteristica di ciascun addendo, che è data da

$$\sqrt[n]{\phi_{S_N}(t)} = e^{-\frac{\lambda}{n} + \frac{\lambda}{n} \phi_{X_1}(t)}$$

L'infinita divisibilità è garantita inoltre dalla formula di Lévy-Khintchine (1.2.1). In effetti, osservando la (1.5) risulta evidente che la tripletta di Lévy per variabili Compound Poisson sia pari a $(b, 0, \nu)$, dove $b = \int_{-1}^1 x \nu(x) dx$.

Osservazione 1.3.2. Dalla funzione caratteristica di una Compound Poisson si deduce che l'esponente nella formula di Lévy-Khintchine possa essere scomposto nella somma di tre componenti indipendenti di tipo degenere, normale e Compound Poisson.

1.3.2 Convergenza di Compound Poisson

Data una distribuzione infinitamente divisibile con tripletta di Lévy (b, σ^2, ν) , è possibile determinare una successione di variabili Compound Poisson che convergano ad essa in distribuzione, come garantito dal seguente teorema.

Teorema 1.3.1. Sia X una variabile aleatoria infinitamente divisibile. Allora esiste una successione di Compound Poisson

$$Y^{(n)} = \sum_{j=1}^{N^{(n)}} X_j^{(n)}$$

con $N^{(n)} \sim Poi(\lambda_n)$ e $X_j^{(n)}$ i.i.d. e indipendente da N tale che

$$Y^{(n)} \xrightarrow[n \rightarrow +\infty]{d} X$$

Dimostrazione. Dimostriamo il teorema nel caso assolutamente continuo.

Tramite la rappresentazione di Lévy-Khintchine, la funzione caratteristica di una distribuzione infinitamente divisibile si esprime come

$$\phi(t) = \exp \left\{ ibt - \frac{1}{2} \sigma^2 t^2 + k(t) \right\}$$

dove

$$k(t) = \int_{-\infty}^{+\infty} (e^{itx} - 1 - itx \mathbb{I}(x)_{[-1,1]}) \nu(x) dx$$

Consideriamo per semplicità il caso in cui $b = \sigma^2 = 0$; sfruttando la linearità dell'integrale, si può esprimere $k(t)$ come

$$k(t) = \int_{[-1,1]^C} (e^{itx} - 1) \nu(x) dx + \int_{-1}^1 (e^{itx} - 1 - itx) \nu(x) dx$$

da cui segue

$$\phi(t) = \exp \left\{ ibt - \frac{1}{2} \sigma^2 t^2 + \int_{[-1,1]^C} (e^{itx} - 1) \nu(x) dx + \int_{-1}^1 (e^{itx} - 1 - itx) \nu(x) dx \right\}$$

Si consideri allora la relazione

$$\nu(x) \mathbb{I}_{[-1,1]^C} = \frac{\nu(x) \mathbb{I}_{[-1,1]^C}}{\int_{[-1,1]^C} \nu(x) dx} \cdot \int_{[-1,1]^C} \nu(x) dx$$

Ponendo

$$f(x) = \frac{\nu(x)\mathbb{I}_{[-1,1]^C}}{\int_{[-1,1]^C} \nu(x) dx}, \quad \lambda = \int_{[-1,1]^C} \nu(x) dx$$

si verifica che $\lambda f(x)$ rappresenta una misura di Lévy. Inoltre $f(x)$ costituisce una funzione di densità di probabilità in quanto

$$\begin{aligned} \int_{-\infty}^{+\infty} f(x) dx &= \int_{-\infty}^{+\infty} \frac{\nu(x)\mathbb{I}_{[-1,1]^C}}{\lambda} dx \\ &= \frac{1}{\lambda} \int_{[-1,1]^C} \nu(x) dx = \frac{1}{\lambda} \lambda = 1 \end{aligned}$$

Da quanto detto si deduce che il primo addendo in k

$$\int_{-\infty}^{+\infty} (e^{itx} - 1)\mathbb{I}_{(x)_{[-1,1]^C}} \nu(x) dx$$

identifica una Compound Poisson i cui addendi sono variabili aleatorie con supporto in $[-1, 1]^C$.

Si consideri ora il secondo termine in k

$$\int_{-1}^1 (e^{itx} - 1 - itx)\nu(x) dx$$

Data la natura problematica del punto, si escluda l'origine passando a

$$\lim_{n \rightarrow +\infty} \int_{\frac{1}{n} < |x| < 1} (e^{itx} - 1 - itx)\nu(x) dx$$

Quest'ultima quantità rappresenta il limite di una Compound Poisson centrata

$$Y^{(n)} - \mathbb{E}[Y^{(n)}]$$

In particolare, data l'equazione

$$\mathbb{I}_{[\frac{1}{n} < |x| < 1]} \nu(x) = \frac{\mathbb{I}_{[\frac{1}{n} < |x| < 1]} \nu(x)}{\int_{\frac{1}{n} < |x| < 1} \nu(x) dx} \cdot \int_{\frac{1}{n} < |x| < 1} \nu(x) dx$$

si pongano

$$f_{(n)} = \frac{\mathbb{I}_{[\frac{1}{n} < |x| < 1]} \nu(x)}{\int_{\frac{1}{n} < |x| < 1} \nu(x) dx}, \quad \lambda_{(n)} = \int_{\frac{1}{n} < |x| < 1} \nu(x) dx$$

Ciò equivale a disporre di una Compound Poisson $Y^{(n)} = \sum_{j=1}^{N^{(n)}} X_j$ in cui

$$N^{(n)} \sim Poi \left(\int_{\frac{1}{n} < |x| < 1} \nu(x) dx \right), \quad f(x^{(n)}) = \frac{\mathbb{I}_{[\frac{1}{n} < |x| < 1]} \nu(x)}{\int_{\frac{1}{n} < |x| < 1} \nu(x) dx}$$

Per la formula di Wald, la centratura di $Y^{(n)}$ si ottiene come

$$\begin{aligned} Y^{(n)} - \mathbb{E}[N^{(n)}] \mathbb{E}[X_j^{(n)}] &= \\ &= Y^{(n)} - \int_{-\infty}^{+\infty} x \nu(x) \mathbb{I}_{[\frac{1}{n} < |x| < 1]} dx \end{aligned}$$

Si ponga ora $C_n = \int_{-\infty}^{+\infty} x \nu(x) \mathbb{I}_{[\frac{1}{n} < |x| < 1]} dx$. La funzione caratteristica di $Y^{(n)}$ sarà pari a

$$\begin{aligned} \mathbb{E}[e^{it(Y^{(n)} - C_n)}] &= e^{-itC_n} \cdot \mathbb{E}[e^{itY^{(n)}}] \\ &= e^{-it \int_{-\infty}^{+\infty} x \mathbb{I}_{[\frac{1}{n} < |x| < 1]} \nu(x) dx} \cdot e^{\int_{-\infty}^{+\infty} (e^{itx} - 1) \nu(x) dx} \end{aligned}$$

$$= e \int_{-\infty}^{+\infty} (e^{itx} - 1 - itx \mathbb{I}_{[\frac{1}{n} < |x| < 1]}) \nu(x) \, dx$$

e da ciò segue la tesi.

□

Osservazione 1.3.3. Oltre a garantire l'esistenza di una successione $Y^{(n)}$ di Compound Poisson che converge in distribuzione ad una distribuzione infinitamente divisibile, la dimostrazione del teorema consente di determinare $\lambda^{(n)}$ e $f^{(n)}$ esplicitamente.

Capitolo 2

Distribuzioni stabili

2.1 Introduzione

Nella teoria della probabilità e nelle applicazioni statistiche, la distribuzione normale riveste un ruolo fondamentale. La predominanza della *gaussiana* risiede principalmente nel teorema centrale del limite.

Teorema 2.1.1. Teorema centrale del limite

Sia X_1, \dots, X_n una successione di variabili aleatorie *i.i.d.* ciascuna avente valore atteso μ finito e varianza σ^2 finita. Allora, data $S_n = \sum_{i=1}^n X_i$ si ha che

$$\frac{S_n - n\mu}{\sqrt{n}\sigma} \xrightarrow{d} \mathcal{N}(0, 1) \quad n \rightarrow +\infty \quad (2.1)$$

La distribuzione normale, tuttavia, è caratterizzata da alcune criticità operative. In particolare, la sua forma funzionale prevede il decadimento esponenziale delle *code* della distribuzione; ciò la rende inadatta a modellizzare fenomeni con elevata variabilità. Inoltre la funzione di densità risulta simmetrica rispetto al valore atteso μ , ipotesi sovente non soddisfatta dai dati. In ultima istanza, il teorema centrale del limite prevede successioni *i.i.d.* con momento primo e secondo finiti, escludendo alcune distribuzioni di probabilità, *e.g.* Pareto.

Per sopperire a questi limiti venne definita intorno al 1920 una nuova famiglia di distribuzioni assolutamente continue, le *stabili*, principalmente ad opera di Paul

Lévy, probabilista e statistico francese, interessato in particolare alla generalizzazione del teorema del limite centrale a variabili che non ne soddisfino le ipotesi. Contributi successivi sono dovuti alla scuola russa, nella quale si annoverano Aleksandr Khintchine, Egor Zolotarev, Boris Gnedenko e Andrej Kolmogorov.

2.2 Definizioni di stabilità

2.2.1 Proprietà di stabilità

La famiglia delle distribuzioni *stabili* è definita come tale in virtù della proprietà di *stabilità* rispetto alla somma. In particolare, una variabile aleatoria X è definita *stabile* se la somma di n sue copie indipendenti si distribuisce come una opportuna combinazione lineare di X . Tale concetto viene formalizzato nella definizione seguente.

Definizione 2.2.1. Una variabile aleatoria X è detta *stabile* se, per ogni $n \in \mathbb{N}$, date X_1, \dots, X_n variabili *i.i.d.* ciascuna avente la medesima distribuzione di X , esistono $c_n > 0$ e $d_n \in \mathbb{R}$ tali che

$$X_1 + \dots + X_n \stackrel{d}{=} c_n X + d_n \quad (2.2)$$

Proposizione 2.2.1. Ogni distribuzione stabile è infinitamente divisibile, ovvero stabilità implica infinita divisibilità.

Dimostrazione. La (2.2) può essere espressa come

$$X \stackrel{d}{=} \frac{X_1 + \dots + X_n - d_n}{c_n} \quad \forall n \in \mathbb{N}$$

Ponendo

$$Y_i = \frac{X_i - \frac{d_n}{n}}{c_n} \quad \forall i = 1, \dots, n$$

segue che

$$X \stackrel{d}{=} Y_1 + \dots + Y_n$$

I singoli addendi Y_i sono indipendenti ed identicamente distribuiti in quanto trasformazioni affini di variabili aleatorie *i.i.d.* Dalla definizione di infinita divisibilità (1.2.1) segue la tesi.

□

2.2.2 Determinazione di c_n

Introduciamo tre lemmi utili alla dimostrazione del teorema seguente, i primi due dei quali sono tratti da Gut (2005), pg. 570.

Lemma 2.2.1. Sia $g : I \rightarrow \mathbb{R}$, $I \subset \mathbb{R}$ una funzione continua e monotona. Allora l'unica funzione che soddisfa $g(x \cdot y) = g(x) \cdot g(y)$ è $g(x) = x^c$.

Lemma 2.2.2. Siano X_1, \dots, X_n variabili aleatorie con distribuzione simmetrica rispetto all'origine. Data $S_n = \sum_{i=1}^n X_i$, essa è ancora simmetrica rispetto all'origine e vale $\mathbb{P}(S_k > x) \leq 2\mathbb{P}(S_n > x)$ per $k < n$.

Lemma 2.2.3. Siano X_1, \dots, X_n variabili aleatorie stabili *i.i.d.* Allora la somma $X_1 + \dots + X_n$ è ancora stabile.

Dimostrazione. Dalla stabilità di X_1 segue che, per opportuni c_n e d_n

$$X_1 + \dots + X_n \stackrel{d}{=} c_n X_1 + d_n \quad (2.3)$$

in cui le X_i sono *i.i.d.* rispetto a X_1 . Poiché per ipotesi le variabili X_1, \dots, X_n sono *i.i.d.* la (2.3) diviene

$$X_1 + \dots + X_n \stackrel{d}{=} c_n X_1 + d_n \quad (2.4)$$

A questo punto resta da dimostrare che $c_n X_1 + d_n$ sia stabile. Nuovamente, poiché X_1 è stabile, vale la (2.4). Tramite una semplice manipolazione algebrica si otterrà

$$(c_n X_1 + d_n) + \dots + (c_n X_n + d_n) \stackrel{d}{=} c_n (c_n X_1 + d_n) + n d_n \quad (2.5)$$

ovvero la definizione di stabilità per la variabile $c_n X_1 + d_n$.

□

Teorema 2.2.1. La costante c_n nella Definizione (2.2.1) è pari a $n^{\frac{1}{\alpha}}$, $0 < \alpha \leq 2$.

Dimostrazione. La seguente dimostrazione è tratta da Gut (2005), pg. 425.

Caso simmetrico, $d_n = 0$.

Sia $S_n = \sum_{j=1}^n X_j$. Dalla definizione di stabilità, dato $d_n = 0$, si ricava

$$S_n \stackrel{d}{=} c_n \cdot X_1$$

Sia ora $S_{nm} = \sum_{j=1}^{nm} X_j$. È possibile spezzare la somma in

$$\sum_{j=1}^n X_j + \sum_{j=n+1}^{2n} X_j + \cdots + \sum_{j=(m-1)n+1}^{mn} X_j$$

Per costruzione, ognuna di queste sommatorie è *i.i.d.* e stabile, in quanto somma di variabili stabili con medesimo parametro α . Da ciò segue che

$$\sum_{j=1}^n X_j + \cdots + \sum_{j=(m-1)n+1}^{mn} X_j \stackrel{d}{=} c_m \cdot Z_1$$

dove $Z_1 = \sum_{j=1}^n X_j$ e dunque $c_m \cdot S_n \stackrel{d}{=} c_m \cdot c_n \cdot X_1$. Ciò implica che $S_{nm} \stackrel{d}{=} c_m \cdot c_n \cdot X_1$.

Tuttavia è anche vero che $S_{nm} = c_{nm} \cdot X_1$. Per confronto si ottiene $c_n \cdot c_m = c_{nm}$ valido $\forall n, k \in \mathbb{N}$.

Si consideri ora la variabile aleatoria

$$S_{(n^k)} = \sum_{j=1}^{(n^k)} X_j \stackrel{d}{=} c_{n^k} X_1$$

e si calcoli $P(X_1 > x) = P(c_{(n^k)} X_1 > c_{n^k} x) = P(S_{(n^k)} > c_{(n^k)} x)$. La somma di variabili aleatorie simmetriche è ancora simmetrica, per cui, dal Lemma 2.2.2

$$P(S_{(n^k)} > c_{(n^k)}x) \leq 2P(S_{(n+1)^k} > c_{(n^k)}x)$$

Sfruttando la proprietà di stabilità e il fatto che $c_{n^k} = (c_n)^k$ segue che

$$P(X_1 > x) \leq 2P\left(X_1 > \left(\frac{c_n}{c_{n+1}}\right)^k x\right)$$

La precedente disequazione è valida per ogni n, k e variabile aleatoria stabile X_1 . In particolare quindi deve valere per $k \rightarrow +\infty$. Fissando n e andando al limite per $k \rightarrow +\infty$ si prospettano due possibilità:

1. $\frac{c_n}{c_{n+1}} > 1$. Per $k \rightarrow \infty$ si ottiene $P(X_1 > x) \leq 2P(X_1 > +\infty)$ ovvero $P(X_1 > x) \leq 0$ per ogni x . Questo implica che, per $k \rightarrow +\infty$, X_1 non sia una variabile aleatoria, il che conduce ad un assurdo.
2. $\frac{c_n}{c_{n+1}} \leq 1$ In questo caso, per $k \rightarrow +\infty$, $P(X_1 > x) \leq 2P(X_1 > 0)$. Ma per simmetria rispetto all'origine questo implica $P(X_1 > x) \leq 1$.

Deve dunque essere $\frac{c_n}{c_{n+1}} \leq 1$ per ogni $n \in \mathbb{N}$ e ciò implica che la successione c_n sia monotona crescente. In virtù di ciò e del Lemma 2.2.2, deriva che c_n debba necessariamente assumere la forma $c_n = n^{\frac{1}{\alpha}}$.

Abbiamo così dimostrato che la definizione di stabilità può essere espressa, date le ipotesi di simmetria e $d_n = 0$, come

$$X_1 + \cdots + X_n = n^{\frac{1}{\alpha}} X$$

Questa equazione implica

$$\mathbb{E}[e^{it(X_1 + \cdots + X_n)}] = \mathbb{E}[e^{itn^{\frac{1}{\alpha}} X}]$$

ovvero

$$\phi(t)^n = \phi(tn^{\frac{1}{\alpha}})$$

L'unica $\phi(t)$ a soddisfare questa equazione ed il fatto che $|\phi(t)| \leq 1$ è

$$\phi(t) = e^{-c|t|^\alpha}$$

Come noto, $\phi(t) = e^{-c|t|^\alpha}$ corrisponde a una distribuzione simmetrica rispetto all'origine, essendo la funzione caratteristica reale¹. Si consideri ora la relazione tra $\phi(t)$ e momenti di una distribuzione espressa da

$$\mathbb{E}[X^r] = \frac{\frac{d^r}{dt^r} \phi(t) \Big|_{t=0}}{i^r} \quad (2.6)$$

Nel caso $\alpha > 2$, considerando $\phi(t) = e^{-c|t|^\alpha}$, dalla (2.6) si ottiene $\mathbb{E}[X] = 0 = \mathbb{V}[X]$, ovvero una variabile degenere $X = 0$. Tuttavia, data $X = 0$, la sua funzione caratteristica $\phi(t)$ sarebbe anche pari a

$$\phi(t) = \mathbb{E}[e^{itX}] = \mathbb{E}[1] = 1 \quad \forall t \in \mathbb{R}$$

il che condurrebbe ad un assurdo in quanto avremmo due funzioni caratteristiche distinte per la medesima distribuzione di probabilità. Da ciò segue che, per $\alpha > 2$, $\phi(t) = e^{-c|t|^\alpha}$ non sia una funzione caratteristica e dunque α debba necessariamente appartenere all'intervallo $(0, 2]$, nel caso simmetrico $d_n = 0$.

Caso non simmetrico, $d_n \neq 0$.

Si consideri nuovamente una variabile aleatoria X stabile. Siano inoltre date due successioni X_1, \dots, X_n e Y_1, \dots, Y_n *i.i.d.* con la medesima distribuzione di X tali che

$$\sum_{i=1}^n X_i \stackrel{d}{=} c_n \cdot X_1 + d_n \quad (2.7)$$

¹ *Dimostrazione.* $\phi_X(t) \in \mathbb{R} \iff \phi_X(t) = \overline{\phi_X(t)} \iff \mathbb{E}[e^{itX}] = \mathbb{E}[e^{-itX}] \iff X \stackrel{d}{=} -X$.

$$\sum_{i=1}^n Y_i \stackrel{d}{=} c_n \cdot Y_1 + d_n \quad (2.8)$$

Sia ora la quantità $Q_n = \sum_{i=1}^n X_i - Y_i$. Sottraendo le equazioni 2.7 e 2.8 si ottiene

$$Q_n \stackrel{d}{=} c_n(X_1 - Y_1)$$

da cui segue che $(X_1 - Y_1)$ è una variabile aleatoria stabile con $d_n = 0$. X_1 e Y_1 sono indipendenti ed identicamente distribuite, perciò

$$X_1 - Y_1 \stackrel{d}{=} Y_1 - X_1$$

il che verifica la condizione necessaria e sufficiente per la simmetria rispetto all'origine di una variabile casuale. Da ciò segue l'estensione al caso asimmetrico dei risultati ottenuti per quello simmetrico, ovvero $0 < \alpha \leq 2$. \square

2.3 Vettore di parametri

A differenza delle distribuzioni di probabilità comunemente impiegate, con funzione di densità ad uno o due parametri, la famiglia delle stabili è definita da un vettore parametrico pari ad $(\alpha, \beta, \mu, \sigma)$. In particolare

1. $0 < \alpha \leq 2$ *Indice di stabilità*
2. $-1 \leq \beta \leq 1$ *Parametro di asimmetria*
3. $\mu \in \mathbb{R}$ *Indice di posizione*
4. $\sigma > 0$ *Parametro di scala*

Inoltre, data una variabile aleatoria X stabile, si può indicare la distribuzione di probabilità ad essa sottesa mediante la notazione

$$X \sim \mathcal{S}_\alpha(\beta, \mu, \sigma)$$

Nel corso della trattazione saranno ottenuti analiticamente gli intervalli di variazione dei parametri α e β .

2.4 Impiego della funzione caratteristica

La famiglia delle distribuzioni stabili è caratterizzata da variabili assolutamente continue, ovvero ammette densità di probabilità. Quest'ultima tuttavia può essere espressa in forma chiusa unicamente in rari casi, rendendo necessario l'impiego della funzione caratteristica $\phi(t)$ che, come noto, è posta in corrispondenza biunivoca con la funzione di densità di probabilità della variabile in esame.

Teorema 2.4.1. Sia X una variabile aleatoria stabile. Allora la sua funzione caratteristica assume la forma

$$\phi(t) = \exp \left\{ i\mu t - \sigma^\alpha |t|^\alpha \left(1 - i\beta \operatorname{sgn}(t) \tan \frac{\pi\alpha}{2} \right) \right\} \quad \alpha \neq 1, \quad t \in \mathbb{R}$$

$$\phi(t) = \exp \left\{ i\mu t - \sigma |t| \left(1 + i\beta \frac{2}{\pi} \operatorname{sgn}(t) \log |t| \right) \right\} \quad \alpha = 1, \quad t \in \mathbb{R}$$

in cui il vettore di parametri $(\alpha, \beta, \mu, \sigma)$ è dato nella Sezione 2.3.

Osservazione 2.4.1. Il Teorema 2.4.1 costituisce una caratterizzazione delle distribuzioni stabili equivalente alla Definizione 2.2.1.

2.4.1 Dimostrazione del Teorema 2.4.1

Si considerino la Definizione (2.2.1) ed il Teorema 2.2.1. Da essi segue l'equazione

$$\mathbb{E}[e^{it(X_1 + \dots + X_n)}] = \mathbb{E}[e^{it(n^{\frac{1}{\alpha}} X + d_n)}]$$

equivalente a

$$[\phi(t)]^n = e^{itd_n} \cdot \mathbb{E}[e^{itn^{\frac{1}{\alpha}} X}] \quad (2.9)$$

La condizione da soddisfare per la (2.9) corrisponde dunque a

$$[\phi(t)]^n = e^{itd_n} \cdot \phi(tn^{\frac{1}{\alpha}})$$

Data l'appartenenza delle distribuzioni stabili alla famiglia delle infinitamente divisibili, è possibile esprimere $\phi(t)$ mediante la formula di Lévy-Khintchine determinata nel Capitolo 1. In formule

$$\phi(t) = \exp \left\{ i\mu t - \frac{1}{2}\gamma^2 t^2 + \int_{\mathbb{R}} (e^{itx} - 1 - itx\mathbb{I}_{[-1,1]}(x)) \nu(dx) \right\}$$

con tripletta di Lévy (μ, γ^2, ν) . Come strategia per la dimostrazione, Gnedenko e Kolmogorov (1954, pg. 164) impongono che la $\phi(t)$ data dalla formula di Lévy-Khintchine soddisfi l'equazione (2.9). Riportiamo di seguito i risultati ottenuti, tralasciando per esigenze espositive i lunghi procedimenti da cui scaturiscono ed il caso $\alpha = 1$, data la reperibilità del tutto sui testi specialistici di cui in bibliografia.

Risultati. Caso $\alpha = 2$.

Si tratta semplicemente del caso gaussiano. In particolare, la misura di Lévy ν è nulla e la funzione caratteristica assume la nota struttura

$$\phi(t) = e^{i\mu t - \frac{1}{2}\gamma^2 t^2}$$

Caso $0 < \alpha < 2$, con $\alpha = 1$ escluso.

Date le restrizioni circa il parametro α , si osserva che $\gamma = 0$ con conseguente elisione della componente gaussiana nella formula di Lévy-Khintchine. Inoltre la misura di Lévy ν assume la forma

$$\nu(x) = \begin{cases} \frac{c_1}{x^{\alpha+1}} & x > 0 \\ \frac{c_2}{|x|^{\alpha+1}} & x < 0 \end{cases}$$

con $c_1 \geq 0$, $c_2 \geq 0$ e tali che $c_1 + c_2 > 0$, ovvero che almeno uno tra c_1 e c_2 sia strettamente positivo.

La misura di Lévy, come evidente dal Paragrafo 1.3.1, risulta legata al contributo Compound Poisson della successione approssimante. In particolare, se $c_1 = c_2$, $\nu(x)$ è simmetrica, *i.e.* data l'approssimazione delle distribuzioni stabili mediante distribuzioni di Poisson composte, i salti S_j di queste ultime sono simmetrici ed analoga considerazione può essere estesa alla loro sommatoria o successione di sommatorie. Naturalmente, se $c_1 = 0$, si avrà massima asimmetria dei salti verso sinistra, e viceversa, data la forma di ν . Sostituendo quest'ultima nella formula di Lévy-Khintchine e tenendo conto dei risultati che soddisfano l'uguaglianza (2.9), Gnedenko e Kolmogorov (1954) giungono all'espressione

$$\phi(t) = \exp \left\{ i \frac{d_n}{n - n^{\frac{1}{\alpha}}} t - |t|^\alpha (c_1 + c_2) \cos \frac{\pi\alpha}{2} \cdot \int_0^{+\infty} (1 - e^{-x}) \frac{dx}{x^{1+\alpha}} \left(1 - i \frac{c_1 - c_2}{c_1 + c_2} \operatorname{sgn}(t) \tan \frac{\pi\alpha}{2} \right) \right\}$$

Si noti che, ponendo

$$\begin{aligned} \sigma^\alpha &= (c_1 + c_2) \cos \frac{\pi\alpha}{2} \cdot \int_0^{+\infty} (1 - e^{-x}) \frac{dx}{x^{1+\alpha}} \\ \beta &= \frac{c_1 - c_2}{c_1 + c_2} \\ \mu &= \frac{d_n}{n - n^{\frac{1}{\alpha}}} \end{aligned}$$

si ottiene la forma funzionale di cui al Teorema (2.4.1), ovvero la classica definizione di distribuzione stabile mediante funzione caratteristica.

Osservazione 2.4.2. Si consideri $\beta = \frac{c_1 - c_2}{c_1 + c_2}$. Appare ora evidente il fatto che esso governi l'asimmetria della distribuzione, dato il legame dei coefficienti c_1 , c_2 con il segno dei salti. Inoltre $|\beta| \leq 1$ e dunque $-1 \leq \beta \leq 1$.

Osservazione 2.4.3. Per $\alpha \neq 1$, $d_n = \mu(n - n^{\frac{1}{\alpha}})$.

2.5 Assoluta continuità delle distribuzioni stabili

Nonostante la funzione di densità sia esprimibile in forma chiusa unicamente per un suo ristretto sottoinsieme, è noto che la famiglia delle distribuzioni stabili sia di natura assolutamente continua. Ciò può essere formalmente verificato mediante il seguente teorema, riportato da Gut (2005).

Teorema 2.5.1. Sia X una variabile aleatoria con $\phi(t) = \mathbb{E}[e^{itX}]$. Allora, se

$$\int_{-\infty}^{+\infty} |\phi(t)| dt < +\infty$$

X ha distribuzione assolutamente continua con funzione di densità f data da

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{-itx} \phi(t) dt$$

Dimostrazione. Si veda Gut (2005), pg. 161. □

Osservazione 2.5.1. Come esempio, riportiamo di seguito la forma della funzione di densità $f(x)$ per una variabile $X \sim \mathcal{S}_\alpha(\beta, \mu, \sigma)$, $\alpha < 1$.

$$f(x) = \frac{1}{\pi x} \sum_{k=1}^{+\infty} \left\{ \frac{\Gamma(k\alpha + 1)}{k!} (-x^{-\alpha})^k \sin \left[\frac{k\pi}{2} \left(\lambda \frac{x}{|x|} - \alpha \right) \right] \right\}$$

in cui λ è una costante dipendente dal vettore parametrico e Γ indica la funzione Gamma. A causa della sommatoria infinita, si tratta di una forma generalmente non esplicita.

2.6 Distribuzioni stabili con densità esplicita

Come evidenziato nell'Osservazione 2.5.1, spesso la funzione di densità non risulta esprimibile in forma chiusa. Vi sono tuttavia alcuni esempi in cui essa può essere esplicitamente determinata.

Esempio 2.6.1. Distribuzione normale

La distribuzione normale si ottiene come caso particolare di stabile, ponendo $\alpha = 2$. In tal caso infatti $\phi(t) = e^{i\mu t - \sigma^2 t^2}$, ovvero la funzione caratteristica di una $\mathcal{N}(\mu, 2\sigma^2)$.

Esempio 2.6.2. Distribuzione di Cauchy standard

Un ulteriore esempio in cui la densità sia espressa in forma chiusa è quello in cui $(\alpha, \beta, \mu, \sigma) = (1, 0, 0, 1)$. Nella suddetta eventualità, $\phi(t) = e^{-|t|}$, cioè la funzione caratteristica di una Cauchy standard, con densità $f(x) = (\pi + \pi x^2)^{-1}$, $x \in \mathbb{R}$.

2.7 Momenti di variabili aleatorie stabili

Proposizione 2.7.1. Sia $X \sim \mathcal{S}_\alpha(\beta, \mu, \sigma)$, con $0 < \alpha < 2$. Allora

$$\mathbb{E}|X|^p < +\infty \quad 0 < p < \alpha$$

$$\mathbb{E}|X|^p = +\infty \quad p \geq \alpha$$

Dimostrazione. La dimostrazione segue dal Lemma 4.4.4 in Gut (2005) il quale afferma che, dati $X \sim \mathcal{S}_\alpha(\beta, \mu, \sigma)$ con funzione caratteristica $\phi(t)$ e $p \in (0, 2)$

$$\mathbb{E}|X|^p = \mathcal{B}(p) \int_{-\infty}^{+\infty} \frac{1 - \Re(\phi(t))}{|t|^{p+1}} dt$$

$$\text{con } \mathcal{B}(p) = \frac{1}{\pi} \Gamma(p+1) \sin\left(\frac{p\pi}{2}\right).$$

□

Osservazione 2.7.1. La Proposizione (2.7.1) implica che per $\alpha < 2$ la varianza di X , $\mathbb{V}(X)$, sia infinita. Inoltre, per $\alpha \leq 1$, $\mathbb{E}[X] = +\infty$, complicando ulteriormente l'uso dei momenti.

2.8 Comportamento delle code nelle distribuzioni stabili

Tra le caratteristiche che rendono le distribuzioni stabili rilevanti nella modellizzazione, si annovera la velocità di decadimento delle loro *code*. In particolare esse decrescono con legge di potenza quando $\alpha < 2$, mentre mostrano decadimento esponenziale per $\alpha = 2$. In particolare, per $\alpha = 2$

$$P(|X| > x) \simeq \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi x}} \quad |x| \rightarrow +\infty$$

Circa il caso $\alpha < 2$ introduciamo la seguente proposizione.

Proposizione 2.8.1. Sia X una variabile aleatoria distribuita secondo una $S_\alpha(\beta, \mu, \sigma)$, $\alpha < 2$. Allora

$$\lim_{x \rightarrow +\infty} x^\alpha \cdot P(X > x) = C_\alpha \frac{1 + \beta}{2} \sigma^\alpha \quad (2.10)$$

$$\lim_{x \rightarrow +\infty} x^\alpha \cdot P(X < -x) = C_\alpha \frac{1 - \beta}{2} \sigma^\alpha \quad (2.11)$$

dove

$$C_\alpha = \left(\int_0^{+\infty} x^{-\alpha} \sin(x) dx \right)^{-1}$$

Dimostrazione. La dimostrazione è data in Samorodnitsky e Taqqu (1994), pg. 17. □

Osservazione 2.8.1. Le equazioni (2.10) e (2.11) possono essere espresse come

$$\begin{aligned} P(X > x) &\simeq \frac{k_1}{x^\alpha} & x \rightarrow +\infty \\ P(X < -x) &\simeq \frac{k_2}{x^\alpha} & x \rightarrow +\infty \end{aligned}$$

Questo implica che, per $\alpha < 2$, le code decadano con legge di potenza proporzionale a $\frac{1}{x^\alpha}$. Il parametro α , oltre ad intervenire nella proprietà di stabilità, descrive dunque la velocità di decadimento delle code della distribuzione. In particolare, al crescere di α , esse divengono progressivamente *più leggere*, ovvero la densità di probabilità tende a 0 più rapidamente. La citata utilità delle distribuzioni stabili nella descrizione di fenomeni con elevata variabilità risiede nella possibilità di variare la densità di probabilità nelle code, adottando un opportuno α .

Osservazione 2.8.2. Si noti come le equazioni (2.10) e (2.11) siano equivalenti nel caso di distribuzioni simmetriche rispetto all' origine. In tal caso, infatti, $\beta = 0$, eguagliando i limiti.

2.9 Teorema centrale del limite generalizzato

Si consideri la successione X_1, \dots, X_n e sia $S_n = \sum_{i=1}^n X_i$. Sia ora

$$\frac{S_n - a_n}{b_n} \quad a_n \in \mathbb{R}, \quad b_n > 0 \quad (2.12)$$

Dal teorema centrale del limite in (2.1) è noto che, nel caso in cui gli addendi in S_n abbiano valore atteso μ e varianza σ^2 finiti, ponendo $a_n = n\mu$ e $b_n = \sigma\sqrt{n}$ si ottiene la convergenza in distribuzione di (2.12) ad una normale standard. Le ipotesi di cui sopra tuttavia non consentono di estendere il risultato a variabili con μ e σ^2 infiniti. Presenteremo una generalizzazione del teorema di Lindeberg-Lévy che consenta di sopperire a questo limite.

Definizione 2.9.1. Una funzione $L : \mathbb{R}^+ \rightarrow \mathbb{R}$ è detta *a variazione lenta* se

$$\lim_{x \rightarrow +\infty} \frac{L(tx)}{L(x)} = 1 \quad \forall t > 0$$

Osservazione 2.9.1. La precedente definizione afferma che, per $x \rightarrow +\infty$, $L(tx) \simeq L(x)$. Ciò si verifica, ad esempio, nel caso in cui si considerino funzioni con asintoto orizzontale.

Teorema 2.9.1. Teorema centrale del limite generalizzato

Siano X_1, \dots, X_n variabili aleatorie *i.i.d.* e sia $S_n = \sum_{i=1}^n X_i$. Allora esistono $a_n \in \mathbb{R}$ e $b_n > 0$ ed una variabile aleatoria G stabile con $\alpha \in (0, 2)$ tali che

$$\frac{S_n - a_n}{b_n} \xrightarrow{d} G$$

se e solo ciascun addendo in S_n soddisfa le proprietà:

1. Per una funzione a variazione lenta $L(x)$ si ha

$$\mathbb{P}(|X_j| > x) \simeq \frac{2 - \alpha}{\alpha} \frac{L(x)}{x^\alpha} \quad x \rightarrow +\infty, \quad \alpha \in (0, 2) \quad (2.13)$$

2. Per un $p > 0$ si ha

$$\lim_{x \rightarrow +\infty} \frac{\mathbb{P}(X_j > x)}{\mathbb{P}(|X_j| > x)} = p \quad , \quad \lim_{x \rightarrow +\infty} \frac{\mathbb{P}(X_j < -x)}{\mathbb{P}(|X_j| > x)} = 1 - p$$

Dimostrazione. Si veda Gut (2005), pg. 431. □

Osservazione 2.9.2. L'ipotesi (2.13) implica che gli addendi in S_n abbiano valore atteso infinito per $\alpha \in (0, 1)$ e varianza infinita per $\alpha \in (0, 2)$.

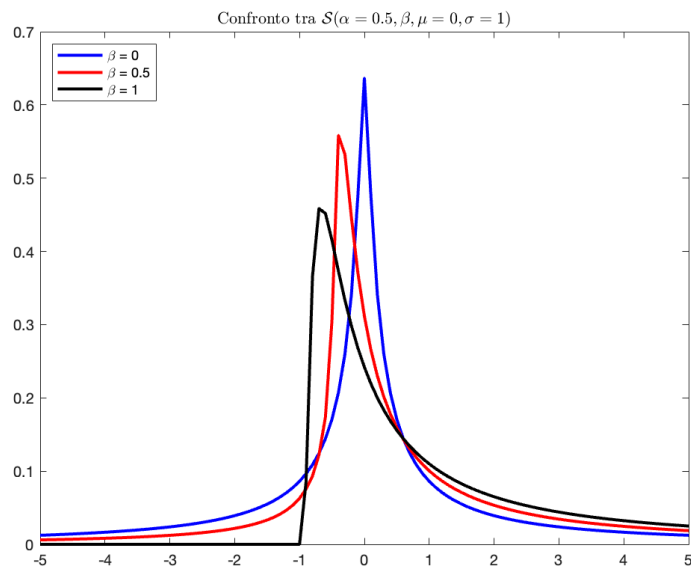
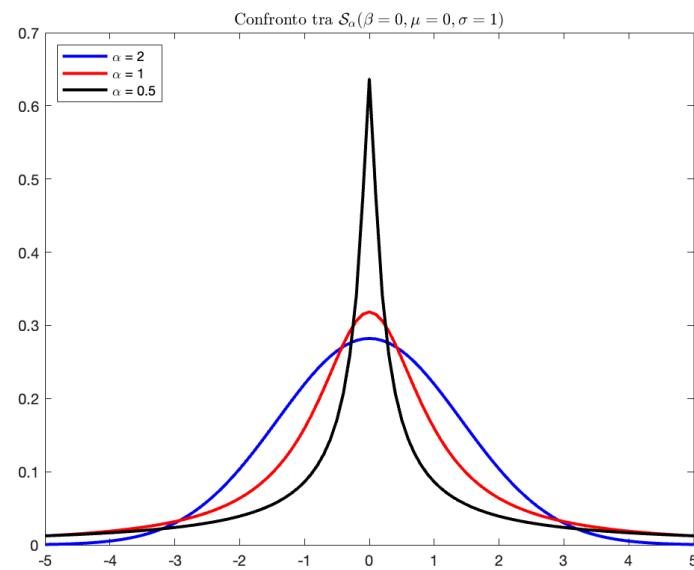
Osservazione 2.9.3. Nel caso in cui $L(x) = 1$ in (2.13), vale che

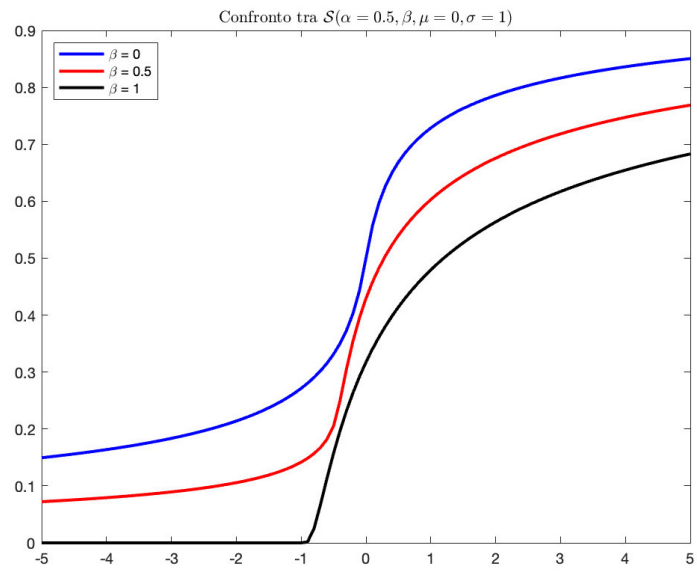
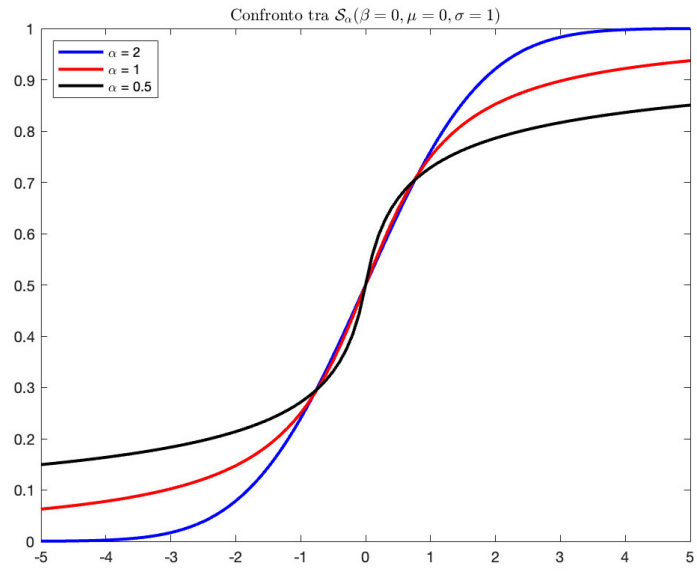
$$\frac{S_n - a_n}{n^{\frac{1}{\alpha}}} \xrightarrow{d} G$$

Nella suddetta eventualità il teorema enunciato è analogo al teorema centrale del limite ordinario, in cui $\alpha = 2$.

2.10 Grafici

Riportiamo di seguito quattro grafici che descrivono l'andamento delle funzioni di densità di probabilità e di ripartizione al variare dei parametri α e β .





Capitolo 3

Procedure di stima dei parametri

In letteratura si riscontra una pluralità di tecniche impiegate nella stima del vettore parametrico delle distribuzioni stabili. In questa sezione, seguendo l'argomentazione di Kateregga *et al.* (2017), forniremo una sintetica panoramica delle metodologie più comunemente adottate nelle procedure di stima: metodo della funzione caratteristica empirica, dei quantili e della massima verosimiglianza.

Prima di affrontare la trattazione dei suddetti, risulta conveniente delineare quali siano le restrizioni alle quali i metodi sono soggetti affinché possano operare efficientemente, mediante la seguente tabella:

Metodo di stima	Restrizioni
Quantili	$\alpha \geq 0.1$
Funzione caratteristica empirica	$\alpha \geq 0.1$
Massima verosimiglianza	$\alpha \geq 0.4$

Tabella 3.1: Metodi di stima e relative restrizioni

3.1 Metodo della funzione caratteristica empirica

Proposto inizialmente da Press (1972), è stato sottoposto a revisioni successive da parte di Paulson *et al.* (1975), Feuerverger e McDunnough (1981), Koutrouvelis (1980). Di seguito una panoramica del metodo in due declinazioni.

3.1.1 Metodo dei momenti

Sia $\{S_1, \dots, S_n\}$ un insieme di osservazioni *i.i.d.* provenienti da una distribuzione stabile. Possiamo stimare la funzione caratteristica $\phi(t)$ di quest'ultima mediante la quantità

$$\hat{\phi}(t) = \frac{1}{n} \sum_{j=1}^n \exp\{itS_j\}$$

Grazie alla formula di Eulero, è possibile esprimere $\phi(t)$ di una distribuzione stabile come

$$\phi(t) = e^{-|\sigma t|^\alpha} (\cos \eta + i \sin \eta) \quad (3.1)$$

dove

$$\eta = \sigma t - |\sigma t|^\alpha \beta \operatorname{sign}(t) \omega(t, \alpha)$$

$$\omega(t, \alpha) = \begin{cases} \tan \frac{\pi \alpha}{2} & \alpha \neq 1 \\ \frac{2 \ln |t|}{\pi} & \alpha = 1 \end{cases}$$

Ne deriva che $|\phi(t)| = \exp\{-|\sigma t|^\alpha\}$. Stimare la funzione caratteristica mediante $\hat{\phi}$ consente una stima dei parametri del modello. Sia infatti

$$\ln |\hat{\phi}(t_k)| = \sigma^\alpha |t_k|^\alpha; \quad k = 1, 2; \quad t_k > 0; \quad \alpha \neq 1$$

Ciò consente la stima di $\hat{\alpha}$ e $\hat{\sigma}$. In particolare:

$$\hat{\alpha} = \frac{\ln \frac{\ln |\hat{\phi}(t_1)|}{\ln |\hat{\phi}(t_2)|}}{\ln \left| \frac{t_1}{t_2} \right|}$$

$$\ln \hat{\sigma} = \frac{\ln |t_1| \ln(-\ln |\hat{\phi}(t_2)|) - \ln |t_2| \ln(-\ln |\hat{\phi}(t_1)|)}{\ln \left| \frac{t_1}{t_2} \right|}$$

Inoltre la parte reale e immaginaria di $\phi(t) = e^{-|\sigma t|^\alpha}(\cos \eta + i \sin \eta)$ forniscono stime di β e μ , indicate con $\hat{\beta}$ e $\hat{\mu}$. In particolare

$$\gamma(t) := \arctan \frac{Im[\phi(t)]}{Re[\phi(t)]} = \mu t - |\sigma t|^\alpha \beta \operatorname{sign}(t) \omega(t, \alpha)$$

Si considerino allora l'insieme di numeri positivi $\{t_k\}$, con $k = 3, 4$. Le stime $\hat{\beta}$ e $\hat{\mu}$ equivalgono a

$$\begin{aligned} \hat{\mu} &= \frac{t_4^{\hat{\alpha}} \gamma(t_3) - t_3^{\hat{\alpha}} \gamma(t_4)}{t_3 t_4^{\hat{\alpha}} - t_4 t_3^{\hat{\alpha}}} \\ \hat{\beta} &= \frac{t_4^{\hat{\alpha}} \gamma(t_3) - t_3^{\hat{\alpha}} \gamma(t_4)}{\hat{\nu}^{\hat{\alpha}} \tan \frac{\pi \hat{\alpha}}{2} (t_4 t_3^{\hat{\alpha}} - t_3 t_4^{\hat{\alpha}})} \end{aligned}$$

Le stime ottenute sono consistenti, in quanto basate su stimatori consistenti di $\phi(t)$, $Re[\phi(t)]$, $Im[\phi(t)]$. Tuttavia, come sottolinea Weron (1995), la convergenza ai valori reali dei parametri dipende dalla scelta di t_1, \dots, t_4 , generalmente complicata.

3.1.2 Metodo regressivo

Accenniamo ora alla seconda procedura di stima, quella di Koutrouvelis (1980), basata su una tecnica regressiva. Dalla (3.1) si deduce che

$$\ln(-\ln(|\phi(t)|^2)) = \ln(2\sigma^\alpha) + \alpha \ln(t)$$

dunque ponendo

$$y_k = \ln(-\ln(|\phi(t)|^2)); \quad m = \ln(2\sigma^\alpha); \quad x_k = \ln(t_k)$$

otteniamo la relazione lineare

$$y_k = m + \alpha x_k + \epsilon_k; \quad k = 1, \dots, M$$

in cui e_k rappresenta l'errore e che rende possibile stimare i parametri α e σ . Koutrouvelis suggerisce $t_k = \frac{\pi k}{25}$. Le stime ottenute, $\hat{\alpha}$ e $\hat{\sigma}$, sono impiegate nella determinazione delle due successive, riguardo β e μ , tramite la relazione

$$\gamma_n(t_l) + \pi k_l(u_l) = \hat{\sigma}_l t - |\hat{\sigma}_l t|^{\hat{\alpha}} \beta \operatorname{sign}(t) \omega(t, \hat{\alpha}) + \psi_l; \quad l = 1, \dots, Q$$

in cui ψ_l è un termine di errore. Stimati i quattro parametri, pari a $\hat{\alpha}, \hat{\beta}, \hat{\sigma}, \hat{\mu}$, la procedura viene iterata considerando questi ultimi come parametri iniziali, fino al soddisfacimento di un criterio di convergenza. Kogon e Williams (1998) hanno perfezionato questo algoritmo, determinando il vettore di partenza $\hat{\Theta}_0 = (\hat{\alpha}_0, \hat{\beta}_0, \hat{\sigma}_0, \hat{\mu}_0)$ tramite il metodo dei quantili di McCulloch, ottenendo vantaggi in termini computazionali.

3.2 Metodi di stima alternativi

3.2.1 Metodo dei quantili

Inizialmente proposto da Fama e Roll (1971) per distribuzioni stabili simmetriche con $\alpha \geq 1$, venne sottoposto a revisione da McCulloch (1986) il quale lo estese a distribuzioni asimmetriche e $\alpha \geq 0.6$.

Sia x_p il quantile a livello p di una popolazione generata da una distribuzione $S_\alpha(\beta, \mu, \sigma)$ e si consideri il rapporto

$$\frac{x_p - x_q}{x_{p'} - x_{q'}}$$

L'idea di McCulloch consiste nello stimare i parametri α e β mediante due funzioni di questa quantità, le quali verranno indicate con $\psi_1(\alpha, \beta)$ e $\psi_2(\alpha, \beta)$. In particolare ψ_1 confronta la dimensione delle code della distribuzione con quella del *centro* e si definisce come

$$\psi_1 = \frac{x_{0.95} - x_{0.05}}{x_{0.75} - x_{0.25}}$$

Naturalmente, valori maggiori di ψ_1 implicano code *più pesanti* e dunque un indice

di stabilità α minore. La funzione ψ_2 misura la differenza tra la parte destra e quella sinistra della distribuzione ed è pari a

$$\psi_2 = \frac{(x_{0.95} - x_{0.5}) - (x_{0.5} - x_{0.05})}{x_{0.95} - x_{0.05}}$$

Risulta evidente che ψ_2 elevata evidenzia una forte densità di probabilità nella porzione destra della distribuzione. Al contrario, valori bassi del rapporto denotano una distribuzione particolarmente sbilanciata nella parte sinistra. Questo è dovuto alla monotonia della funzione di ripartizione mediante la quale vengono computati i quantili della distribuzione, ed il fatto che la differenza al numeratore sia contenuta sottolinea l'esiguità dell'area sottesa alla curva tra $x_{0.5}$ e $x_{0.95}$. Appare quindi evidente come ψ_2 possa rappresentare una forte indicazione nella stima di β .

Le suddette considerazioni valgono nel caso di una *popolazione*, dunque necessitano di essere estese a campioni di essa, ai fini della stima dei parametri di interesse. La strategia consiste nel rimpiazzare x_p col proprio equivalente campionario \hat{x}_p e definire delle stime $\hat{\psi}_1$ e $\hat{\psi}_2$ come soluzione del sistema

$$\begin{cases} \psi_1(\hat{\alpha}, \hat{\beta}) = \hat{\psi}_1 \\ \psi_2(\hat{\alpha}, \hat{\beta}) = \hat{\psi}_2 \end{cases} \quad (3.2)$$

in cui le equazioni $\hat{\psi}_1$ e $\hat{\psi}_2$ sono definite come

$$\hat{\psi}_1 = \frac{\hat{x}_{0.95} - \hat{x}_{0.05}}{\hat{x}_{0.75} - \hat{x}_{0.25}}$$

$$\hat{\psi}_2 = \frac{\hat{x}_{0.95} + \hat{x}_{0.05} - 2\hat{x}_{0.5}}{\hat{x}_{0.95} - \hat{x}_{0.05}}$$

Ci si interroga dunque su quali siano gli stimatori $\hat{\alpha}$ e $\hat{\beta}$ tali che le funzioni campionarie sopra definite siano equivalenti a quelle computate sulla popolazione. McCulloch fornisce delle tabelle che restituiscono i valori di $\hat{\alpha}$ e $\hat{\beta}$ dati i valori di $\hat{\psi}_1$ e $\hat{\psi}_2$, rendendo dunque possibile la stima dei primi due parametri caratteristici delle distribuzioni stabili: α e β .

Riguardo il parametro di scala σ , dato che la differenza tra quantili di una distribuzione è strettamente legata ad esso, risulta conveniente definire

$$\hat{\sigma} = \frac{\hat{x}_{0.75} - \hat{x}_{0.25}}{\psi_3(\hat{\alpha}, \hat{\beta})}$$

con $\psi_3(\hat{\alpha}, \hat{\beta})$ fornita nella *Table V* in McCulloch (1986).

Si pone dunque la sfida della determinazione dell' ultimo parametro, di posizione, μ . Tralasciando per semplicità considerazioni inerenti alla discontinuità delle funzioni in gioco, introduciamo un nuovo parametro ϕ definito come

$$\phi = \begin{cases} \mu + \beta \gamma \tan \frac{\pi \alpha}{2} & \alpha \neq 1 \\ \mu & \alpha = 1 \end{cases}$$

stimato da $\hat{\phi} = \hat{x}_{0.5} + \hat{\sigma} \psi_5$, ricavando ψ_5 dalla *Table VIII* di McCulloch. Da ciò deriviamo lo stimatore

$$\hat{\mu} = \hat{\phi} + \hat{\beta} \hat{\sigma} \tan \frac{\pi \hat{\alpha}}{2}$$

3.2.2 Metodo della massima verosimiglianza

Il metodo della massima verosimiglianza rappresenta la tecnica di stima principale in analisi finanziarie ed econometriche, qualora si operi con distribuzioni stabili. La funzione di densità di probabilità ad esse sottesa, tuttavia, risulta raramente espressa in forma esplicita; si sopperisce a tale problema mediante approssimazione numerica.

Sia quindi $\{S_1, \dots, S_n\}$ un insieme di variabili *i.i.d.* distribuite secondo una $S_\alpha(\beta, \mu, \sigma)$. La stima di massima verosimiglianza del vettore parametrico $\Theta = (\alpha, \beta, \mu, \sigma)$ è ottenuta massimizzando la funzione di log-verosimiglianza

$$l(\Theta) = \ln \left(\prod_{i=1}^n \hat{f}(S_i; \Theta) \right)$$

\hat{f} rappresenta una stima numerica della funzione di densità sottesa e Mittnik *et al.* (1999) mostrano che l'algoritmo più efficiente nella sua determinazione sia il *Fast Fourier Transform* nel caso n assuma valore elevato, mentre suggeriscono il *metodo di integrazione diretta* proposto da Nolan (1997) per n contenuto.

3.3 Considerazioni sui metodi di stima

Kateregga *et al.* (2017) sottolineano come i metodi elencati si dimostrino fallaci qualora i parametri delle distribuzioni in esame si avvicinino ai loro valori estremi, mentre appaiono soddisfacenti sotto le restrizioni di cui alla Tabella 3.1. In particolare, gli autori evidenziano il fatto che il metodo della funzione caratteristica empirica performi globalmente meglio rispetto ai restanti due in quanto:

1. Fornisce stimatori robusti
2. Garantisce una precisione maggiore in ampi intervalli di variazione di α e β
3. Mostra un migliore tasso di convergenza

Viene dunque suggerito dagli stessi l'impiego degli altri metodi per la determinazione di una stima iniziale dei parametri tramite la quale inizializzare gli algoritmi modellati sul metodo della funzione caratteristica empirica e sfruttare questi ultimi per le valutazioni definitive.

3.4 Applicazioni in R

In questa sezione mostriamo brevemente la procedura di generazione di variabili casuali distribuite secondo $S_\alpha(\beta, \mu, \sigma)$ e di stima dei parametri basata sul del metodo della funzione empirica caratteristica proposto da Koutrouvelis (1980), tramite il software R.

Esempio 3.4.1. Dopo aver installato i pacchetti `Rcpp` e `libstableR`, generiamo un campione di $n = 1000$ unità **camp** *i.i.d.* secondo una stabile con vettore parame-

trico $\text{par} = (\alpha = 1.5, \beta = 0.9, \sigma = 1, \mu = 0)$ ed otteniamo la stima dei parametri come segue:

```
> library(libstableR)

> library(Rcpp)

> par = c(1.5, 0.9, 1, 0)

> camp = stable_rnd(N=1000, pars=par)

> parametri_iniziali = stable_fit_init(camp)

> stima_finale = stable_fit_koutrouvelis(camp, pars_init = parametri_iniziali)

> stima_finale
[1] 1.40949984 0.85300342 0.96715455 -0.07673328
```

In particolare, il vettore `parametri_iniziali` contiene una stima preliminare del vettore parametrico, ottenuta mediante il metodo dei quantili di McCulloch. In seguito, `stima_finale` servendosi della stima preliminare adotta il metodo della funzione caratteristica empirica per trarne un risultato evidentemente compatibile con i valori *reali* dei parametri.

Esempio 3.4.2. Servendoci dei medesimi pacchetti di cui sopra, generiamo un campione di $n = 10000$ unità `camp` provenienti da una $\mathcal{N}(2, 1)$.

```
> camp = rnorm(10000, 2, 1)

> parametri_iniziali = stable_fit_init(camp)
```

```
> stima_finale = stable_fit_koutrouvelis(camp, pars_init = parametri_iniziali)
```

```
> stima_finale
```

```
[1] 1.9949427 0.0000000 0.7021753 2.0118236
```

Notando dall' `help(libstableR)` che la parametrizzazione default impiegata corrisponde alla 0 in Nolan (1998), ricaviamo che la varianza stimata è pari a $2\hat{\sigma}^2$, ovvero a 0.98, confermando nuovamente la bontà delle stime.

Bibliografia

- [1] D. Applebaum. *Lévy processes and stochastic calculus*. Cambridge university press, 2009.
- [2] B. De Finetti. *Sulle funzioni a incremento aleatorio*. Bardi, 1929.
- [3] E. F. Fama and R. Roll. Parameter estimates for symmetric stable distributions. *Journal of the American Statistical Association*, 66(334):331–338, 1971.
- [4] W. Feller. An introduction to probability theory and its applications. *Wiley*, 1957.
- [5] A. Feuerverger and P. McDunnough. On the efficiency of empirical characteristic function procedures. *Journal of the Royal Statistical Society. Series B (Methodological)*, 43(1):20–27, 1981.
- [6] B. Gnedenko and A. Kolmogorov. Limit distributions for sums of independent random variables. *Am. J. Math*, 105, 1954.
- [7] A. Gut. *Probability: a graduate course*, volume 200. Springer, 2005.
- [8] K. Itô. On stochastic processes (i) infinitely divisible laws of probability. In *Japanese journal of mathematics: transactions and abstracts*, volume 18, pages 261–301. The Mathematical Society of Japan, 1941.
- [9] M. Kateregga, S. Mataramvura, and D. Taylor. Parameter estimation for stable distributions with application to commodity futures log-returns. *Cogent Economics & Finance*, 5(1):1318813, 2017.

- [10] S. M. Kogon and D. B. Williams. Characteristic function based estimation of stable distribution parameters. *A practical guide to heavy tails: statistical techniques and applications*, pages 311–338, 1998.
- [11] I. A. Koutrouvelis. Regression-type estimation of the parameters of stable laws. *Journal of the American statistical association*, 75(372):918–928, 1980.
- [12] P. Lévy. Théorie des erreurs. La loi de gauss et les lois exceptionnelles. *Bulletin de la Société mathématique de France*, 52:49–85, 1924.
- [13] F. Mainardi and S. Rogosin. The origin of infinitely divisible distributions: from de Finetti’s problem to Levy-Khintchine formula. 2008.
- [14] J. H. McCulloch. Simple consistent estimators of stable distribution parameters. *Communications in Statistics-Simulation and Computation*, 15(4):1109–1136, 1986.
- [15] J. H. McCulloch. 13 financial applications of stable distributions. *Handbook of statistics*, 14:393–425, 1996.
- [16] J. P. Nolan. Numerical calculation of stable densities and distribution functions. *Communications in statistics. Stochastic models*, 13(4):759–774, 1997.
- [17] J. P. Nolan. Parameterizations and modes of stable distributions. *Statistics & probability letters*, 38(2):187–195, 1998.
- [18] A. S. Paulson, E. W. Holcomb, and R. A. Leitch. The estimation of the parameters of the stable laws. *Biometrika*, 62(1):163–170, 1975.
- [19] S. J. Press. Estimation in univariate and multivariate stable distributions. *Journal of the American Statistical Association*, 67(340):842–846, 1972.
- [20] G. Samorodnitsky and M. S. Taqqu. Stable non-gaussian random processes: stochastic models with infinite variance. *Chapman & Hall*, 1994.

- [21] V. Zolotarev. Asymptotic behavior of the distributions of processes with independent increments. *Theory of Probability & Its Applications*, 10(1):28–44, 1965.
- [22] V. M. Zolotarev. *One-dimensional stable distributions*, volume 65. American Mathematical Soc., 1986.