

## CAPITOLO 2: STRUMENTI STATISTICI PRELIMINARI

Gli eventi estremi, in campo assicurativo e finanziario, hanno il vantaggio, dal punto di vista matematico, di poter essere per la maggior parte quantificabili in unità monetarie. D'altra parte, però, molti di essi hanno una componente non misurabile che sempre più studiosi stanno cercando di tenere in considerazione. Esiste quindi un rilevante bisogno di "modellizzare" quegli eventi che provocano grandi scosse nel sistema finanziario, quali le catastrofi, soprattutto alla luce della loro crescente importanza, sia in termini di frequenza che in termini di costi, come evidenziato dai dati presentati nelle pagine precedenti.

### 2.1 IL PROBLEMA DELLA ROVINA

Nel presente paragrafo viene introdotta un'analisi matematica della stima di rovina sotto precise ipotesi di modelli "heavy-tailed"; per poter far questo, però, è innanzitutto necessario introdurre il seguente modello di base di teoria del rischio:

$$U(t) = u + ct - S(t), \quad \text{con } S(t) = \sum_{i=1}^{N(t)} X_i \quad t \geq 0,$$

dove  $U(t)$  rappresenta il processo di rischio all'istante  $t$ ,  $u$  il capitale iniziale,  $c$  il tasso di caricamento del premio, mentre il costo sinistri aggregato  $S(t)$  è dato dalla somma casuale delle  $N(t)$  variabili indipendenti ed identicamente distribuite (iid)  $X_i$ , essendo  $N(t)$  la variabile aleatoria "numero di sinistri generato dal portafoglio in esame", nell'unità di tempo considerata<sup>22</sup>. Una possibile rappresentazione di  $U(t)$  è fornita dalla figura 4.

---

<sup>22</sup> EMBRECHTS PAUL, KLÜPPELBERG CLAUDIA, MIKOSCH THOMAS, *Modelling extremal events: for insurance and finance*, Berlin, Springer, 2001.

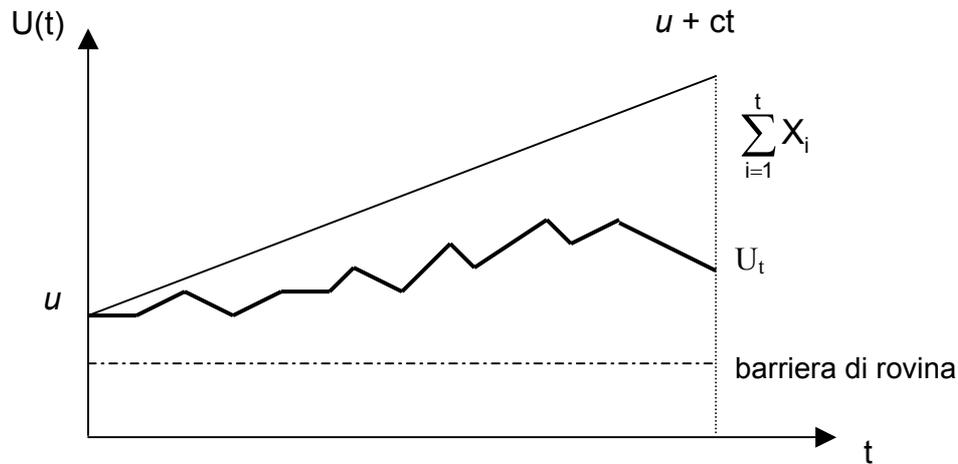


Figura 4- Una possibile realizzazione del processo di rischio  $U(t)$ .

Nel presente modello,  $S(t)$  viene analizzato considerando il portafoglio nella sua globalità, senza prestare alcuna attenzione al singolo rischio (seguendo cioè un approccio collettivo). Per potere quindi applicare il modello, il portafoglio dovrà essere composto da rischi con alto grado di omogeneità, in quanto in genere si richiede che le  $X_i$  siano indipendenti ed identicamente distribuite, ossia si ipotizza:

- la reciproca indipendenza di tutti i sinistri;
- che il costo  $X_i$  di ogni singolo sinistro sia identicamente distribuito secondo la funzione di ripartizione:

$$F(x) = \Pr \{ X_i \leq x \} = \Pr \{ X_{i+1} \leq x \} = \Pr \{ X_{i+2} \leq x \} = \dots$$

Gli eventi estremi, quali le catastrofi, però, sono caratterizzati da una scarsa omogeneità; risulta dunque importante analizzare quale influenza essi abbiano sul comportamento globale del processo di rischio  $U(t)$ ; una soluzione in tal senso ci viene fornita da una analisi dettagliata della probabilità di rovina associata ad  $U(t)$ , possibile grazie all'applicazione del modello di Cramér-Lundberg.

Sotto una condizione di piccoli sinistri, tale stima della probabilità di rovina restituisce valori esponenziali nel capitale iniziale  $u$ . D'altra parte, nella realtà, i sinistri sono in gran parte caratterizzati da distribuzioni *heavy-tailed* come la Pareto, la Loggamma, la Lognormale o la Weibull *heavy-tailed*.

### 2.1.1 Il modello di Cramér-Lundberg

Il modello di Cramér-Lundberg è dato dalle seguenti condizioni:

a) il *processo di "ampiezza dei sinistri"*:

le ampiezze dei sinistri  $(X_k)_{k \in \mathbb{N}}$  sono variabili casuali positive, indipendenti ed identicamente distribuite, aventi tutte la medesima funzione di distribuzione  $F$ , media finita  $\mu = E(X_1)$  e varianza  $\sigma^2 = \text{var}(X_1) \leq \infty$

b) l'*istante di accadimento dei sinistri*:

i sinistri avvengono in istanti di tempo casuali

$$0 < T_1 < T_2 < \dots \quad \text{quasi certamente (q.c.)};$$

c) il *processo di accadimento dei sinistri*:

il numero dei sinistri che avvengono nell'intervallo di tempo  $[0, t]$  è indicato da

$$N(t) = \sup \{ n \geq 1 : T_n \leq t \}, \quad t \geq 0,$$

dove, per convenzione,  $\sup \emptyset = 0$ ;

d) gli *intervalli di tempo* che intercorrono tra un sinistro e quello successivo:

$$Y_1 = T_1, \quad Y_2 = T_2 - T_1, \quad \dots, \quad Y_k = T_k - T_{k-1}, \quad \text{con } k = 2, 3, \dots$$

sono variabili indipendenti ed identicamente distribuite secondo una funzione esponenziale, con media finita  $E(Y_1) = 1/\lambda$ ;

e) le successioni  $(X_k)$  e  $(Y_k)$  sono tra loro indipendenti.

Se, alla condizione d), si sostituisce la seguente:

d') gli intervalli di tempo che intercorrono tra un sinistro e quello successivo,

$Y_k$ , sono indipendenti ed identicamente distribuiti, con media  $E(Y_1) = 1/\lambda$  finita,

il modello, che rappresenta una generalizzazione di quello di Cramér-Lundberg, prende il nome di *renewal model* (o modello di rinnovo).

Dalla presente definizione si può quindi dedurre che  $(N(t))$  è un processo di Poisson omogeneo con intensità  $\lambda$ . Quindi:

$$P(N(t) = k) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}, \quad k = 0, 1, 2, \dots$$

Il *renewal model* consente, invece, un approccio più generale, pertanto  $(N(t))$  viene principalmente assimilato ad un processo di Poisson misto, più adatto a descrivere il fenomeno nelle applicazioni assicurative, come dimostrato dall'esperienza reale.

Il processo *costo sinistri aggregato*  $(S(t))_{t \geq 0}$  di un dato portafoglio è definito come segue:

$$S(t) = \begin{cases} \sum_{i=1}^{N(t)} X_i & \text{se } N(t) > 0 \\ 0 & \text{se } N(t) = 0 \end{cases}$$

Un'altra importante quantità, in questo contesto, è costituita dalla funzione di distribuzione di  $S(t)$ , ossia dalla *distribuzione aggregata dei sinistri*:

$$G_t(x) = P(S(t) \leq x) = \sum_{n=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} F^{n*}(x), \quad x \geq 0, t \geq 0,$$

essendo  $F^{n*}(x) = P(\sum_{i=1}^n X_i \leq x)$  la convoluzione  $n$ -esima di  $F$ , che pertanto risulta essere una funzione di distribuzione di tipo Poisson Composto.

Prima di continuare nell'analisi del processo di rischio  $U(t)$ , è necessario introdurre il concetto di probabilità di rovina, il cui calcolo costituisce l'obiettivo principale del modello in questione.

La probabilità di rovina su un orizzonte temporale finito è definita nel seguente modo:

$$\psi(u, T) = P(U(t) < 0 \text{ per alcuni } t \leq T), \quad 0 < T < \infty, \quad u \geq 0.$$

Su un orizzonte temporale infinito, invece, è indicata come segue:

$$\psi(u) = \psi(u, \infty), \quad u \geq 0.$$

L'istante della rovina è così definito:

$$\tau(T) = \inf \{ t : 0 \leq t \leq T, U(t) < 0 \}, \quad 0 < T \leq \infty,$$

dove, per convenzione,  $\inf \emptyset = \infty$ .

A questo punto è possibile calcolare il valore atteso del processo di rischio  $U(t)$ :

$$E[U(t)] = E[u + ct - S(t)] = u + ct - E[S(t)]$$

Nel caso più generale, ossia per quanto riguarda il *renewal model*

$$\begin{aligned}
 E[ S(t) ] &= \sum_{k=0}^{\infty} [S(t) | N(t) = k] P[N(t) = k] = \sum_{k=1}^{\infty} E \left( \sum_{i=1}^{N(t)} X_i \mid N(t) = k \right) P(N(t) = k) \\
 &= \sum_{k=1}^{\infty} E \left( \sum_{i=1}^k X_i \right) P(N(t) = k) = \mu \sum_{k=1}^{\infty} k P(N(t) = k) = \mu E(N(t)).
 \end{aligned}$$

Di conseguenza, si deriva che

$$E(U(t)) = u + ct - \mu E(N(t)). \quad (II.1)$$

Nel caso, invece, più specifico del modello di Cramér-Lundberg, essendo  $E(N(t)) = \mu t$ , risulta

$$E(U(t)) = u + ct - \lambda \mu t.$$

Dai valori ottenuti è possibile osservare che il processo di rischio  $U(t)$  dipende dal tasso di incremento di premio  $c$ ; la determinazione di tale valore, a sua volta, costituisce un ulteriore problema in campo assicurativo. La definizione di un adatto tasso di incremento del premio dipende dalla misura di solvibilità che si vuole ottimizzare in un dato periodo di tempo. Alcuni valori di cui si è immediatamente a disposizione sono le probabilità di rovina  $\psi(u, T)$  per  $T \leq \infty$ . Il tasso di premio  $c$  dovrebbe essere scelto in modo che, per determinati  $u$  e  $T$ , risulti una bassa  $\psi(u, T)$ . Un primo passo in questa direzione sarebbe quello di richiedere che  $\psi(u) < 1$ , per tutte le  $u \geq 0$ . D'altra parte, dal momento in cui  $\psi(u) = P(\tau < \infty)$ , questo equivale a dire che  $P(\tau = \infty) > 0$ : la Compagnia ha, cioè, una piccola probabilità positiva di sopravvivenza in un periodo infinitamente lungo. Dalla (II.1) segue immediatamente che, nel *renewal model*<sup>23</sup>, per  $t \rightarrow \infty$

$$E[U(t)] = u + (c - \lambda \mu) t (1 + o(1)) = u + \left( \frac{c}{\lambda \mu} - 1 \right) \lambda \mu t (1 + o(1)).$$

Quindi  $\frac{E(U(t))}{t} \rightarrow c - \lambda \mu$  ed una ovvia condizione per la solvibilità è che  $c - \lambda \mu > 0$ ,

ossia che il premio puro risulti maggiore del valore atteso delle perdite, e ciò implica che  $(U(t))$  ha accumuli positivi per valori ampi di  $t$ .

Questo porta alla condizione del profitto netto di base, per il *renewal model*:

<sup>23</sup> Per quanto riguarda il valore atteso del numero sinistri  $N(t)$ , nel *renewal model*, vale la seguente relazione:

$$E[N(t)] = (\lambda + o(1))t, \quad \text{per } t \rightarrow \infty;$$

$$\rho = \frac{c}{\lambda\mu} - 1 > 0.$$

La costante  $\rho$  è detta *caricamento di sicurezza* e può essere interpretata come un tasso di premio di rischio; infatti, l'incremento del premio nel periodo  $[0,t]$  equivale a  $ct = (1+\rho)\lambda\mu t$ .

Dalla definizione del processo di rischio, la rovina può avvenire solo negli istanti di accadimento  $T_i$ , quindi per  $u \geq 0$ ,

$$\begin{aligned} \psi(u) &= P(u + ct - S(t) < 0, \text{ per alcuni } t \geq 0) \\ &= P(u + cT_n - S(T_n) < 0, \text{ per alcuni } n \geq 1) \\ &= P\left(u + \sum_{k=1}^n (cY_k - X_k) < 0, \text{ per alcuni } n \geq 1\right) \\ &= P\left(\sup_{n \geq 1} \sum_{k=1}^n (X_k - cY_k) > u\right) \end{aligned}$$

Quindi  $\psi(u) < 1$  equivale alla seguente condizione:

$$1 - \psi(u) = P\left(\sup_{n \geq 1} \sum_{k=1}^n (X_k - cY_k) \leq u\right) > 0 \quad u \geq 0.$$

La determinazione della probabilità  $1 - \psi(u)$ , dunque, si riduce all'analisi della funzione di distribuzione dell'estremo superiore di un cammino aleatorio.

### **2.1.2 La stima di Cramér-Lundberg per i piccoli sinistri**

Nelle pagine precedenti, è stato presentato un metodo generale per la stima della probabilità di rovina  $\psi(u)$ , per il *renewal model*. Se si fa riferimento al modello di Cramér-Lundberg, è possibile ottenere una formula per  $\psi(u)$ , che coinvolga esplicitamente la funzione di distribuzione dell'ampiezza dei sinistri  $F$ . Infatti per il modello in questione, sotto la condizione del profitto netto

$\rho = \frac{c}{\lambda\mu} - 1 > 0$ , vale

$$1 - \psi(u) = \frac{\rho}{1 + \rho} \sum_{n=0}^{\infty} (1 + \rho)^{-n} F_1^{n*}(u), \quad (\text{II.2})$$

dove

$$F_1(x) = \frac{1}{\mu} \int_0^x \bar{F}(y) dy, \quad x \geq 0$$

indica la distribuzione integrata della coda e

$$\bar{F}(x) = 1 - F(x), \quad x \geq 0$$

rappresenta la coda della funzione di distribuzione  $F$ .

La formula (II.2) è uno strumento base per la stima della probabilità di rovina e la prova della sua validità viene fornita dal seguente teorema.

**Teorema di Cramér-Lundberg (2.1):** *si consideri il modello di Cramér-Lundberg e la condizione del profitto netto  $\rho > 0$  ;  
si supponga che esista un valore  $v > 0$  tale che*

$$\hat{f}_1(-v) = \int_0^{\infty} e^{vx} dF_1(x) = \frac{C}{\lambda\mu} = 1 + \rho. \quad (II.3)$$

Allora valgono le seguenti condizioni:

- per tutti gli  $u \geq 0$ ,

$$\psi(u) \leq e^{-vu}$$

- se, inoltre,

$$\int_0^{\infty} xe^{vx} \bar{F}(x) dx < \infty,$$

allora

$$\lim_{u \rightarrow \infty} e^{vu} \psi(u) = C < \infty,$$

dove

$$C = \left[ \frac{v}{\rho\mu} \int_0^{\infty} xe^{vx} \bar{F}(x) dx \right]^{-1};$$

- nel caso di una funzione di distribuzione esponenziale  $F(x) = 1 - e^{-x/\mu}$ , la (II.2) diventa:

---

<sup>24</sup> Viene qui applicata la *trasformata di Laplace-Stieltjes*; si consideri infatti una funzione di distribuzione  $H$  concentrata su  $(0, \infty)$ , allora  $\hat{h}(s) = \int_0^{\infty} e^{-sx} dH(x)$ , con  $s \in \Re$ , rappresenta la trasformata di Laplace-Stieltjes della funzione  $H$ .

$$\psi(u) = \frac{1}{1+\rho} \exp\left\{-\frac{\rho}{\mu(1+\rho)}u\right\}, \quad u \geq 0.$$

La fondamentale, cosiddetta, *condizione di Cramér-Lundberg*, espressa dalla (II.3), può anche essere riscritta nel seguente modo:

$$\int_0^{\infty} e^{\nu x} \bar{F}(x) dx = \frac{c}{\lambda}.$$

Inoltre, data la funzione di distribuzione dell'ampiezza sinistri  $F$ , la costante  $\nu > 0$ , che soddisfa la precedente equazione, è chiamata *esponente di Lundberg* o *coefficiente di aggiustamento* del processo di rischio. L'esistenza di  $\nu$  implica che  $\hat{f}_1(s)$  esiste in un intorno non vuoto di 0, il che significa che la coda della distribuzione integrata  $\bar{F}_1$  (e quindi anche la coda  $\bar{F}$ ) è esponenzialmente limitata.

Inoltre, dalla disuguaglianza di Markov, secondo la quale  $P(|X| \geq a) \leq \frac{E|X|}{a}$ , segue che

$$\bar{F}(x) \leq e^{-\nu x} E(e^{\nu X_1}), \quad x > 0.$$

Questa espressione indica che è molto inverosimile che avvengano sinistri di grande entità, in quanto ad essi sono associate probabilità esponenzialmente piccole. Proprio per questa ragione, la (II.3) è detta *condizione dei piccoli sinistri*.

La condizione di Cramér-Lundberg può essere analizzata anche a livello grafico. L'esistenza di  $\nu$ , nella (II.3), dipende dall'ascissa sinistra di convergenza di  $\hat{f}_1$ .

Nella realtà si possono verificare diverse situazioni, ma quella più comune è rappresentata dalla figura 5, che considera il caso dei sinistri di piccola entità, per i quali è calcolabile l'esponente di Lundberg e di conseguenza è applicabile il teorema di Cramér-Lundberg. Funzioni di distribuzione dell'ampiezza dei sinistri, tipiche del caso in questione, sono rappresentate nella tabella 4.

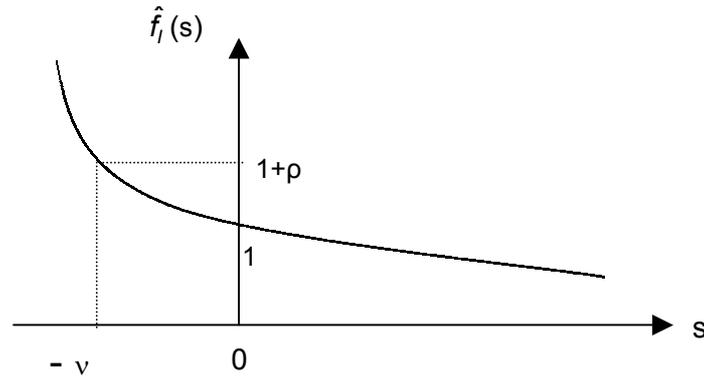


Figura 5- La condizione di Cramér-Lundberg nel caso di piccoli sinistri

Nome	Coda $\bar{F}$ o densità $f$	Parametri
Esponenziale	$\bar{F}(x) = e^{-\lambda x}$	$\lambda > 0$
Gamma	$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$	$\alpha, \beta > 0$
Weibull	$\bar{F}(x) = e^{-cx^\tau}$	$c > 0, \tau \geq 1$
Normale troncata	$f(x) = \sqrt{\frac{2}{\pi}} e^{-x^2/2}$	

Tabella 4- Funzioni di distribuzione dell'ampiezza sinistri per "piccoli sinistri". Tutte le funzioni di distribuzione hanno supporto  $(0, \infty)$ .

### 2.1.3 La teoria della rovina per distribuzioni *heavy-tailed*

Nella sezione precedente è stata presentata la stima della probabilità di rovina di Cramér-Lundberg, su un orizzonte di tempo infinito, nel caso di sinistri di piccola entità. Nella realtà, però, sempre più spesso si ha a che fare con eventi estremi, cioè avvenimenti caratterizzati da funzioni di distribuzione *heavy-tailed*, per i quali non esiste l'esponente di Lundberg e di conseguenza non è possibile applicare la stima della probabilità di rovina vista sopra. Nelle pagine seguenti si cercherà dunque l'equivalente della suddetta teoria, applicabile, però, ai sinistri di grande entità.

Da questo momento si considereranno variabili casuali positive con supporto infinito, ossia  $F(x) < 1$  per tutte le  $x > 0$ .

La tabella 5 indica le possibili funzioni di distribuzione dell'ampiezza dei sinistri peculiari del caso in esame, mentre la figura 6 fornisce una rappresentazione di  $\hat{f}_1(s)$  per i "grandi" sinistri.

Nome	Coda $\bar{F}$ o densità $f$	Parametri
Lognormale	$f(x) = \frac{1}{\sqrt{2\pi}\sigma x} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}$	$\mu \in \mathfrak{R}, \sigma > 0$
Pareto	$\bar{F}(x) = \left(\frac{\kappa}{\kappa + x}\right)^\alpha$	$\alpha, \kappa > 0$
Burr	$\bar{F}(x) = \left(\frac{\kappa}{\kappa + x^\tau}\right)^\alpha$	$\alpha, \kappa, \tau > 0$
Benktander 1	$\bar{F}(x) = (1 + 2(\beta/\alpha)\ln x)e^{-\beta(\ln x)^2 - (\alpha+1)\ln x}$	$\alpha, \beta > 0$
Benktander 2	$\bar{F}(x) = e^{\alpha/\beta} x^{-(1-\beta)} e^{-\alpha x^\beta/\beta}$	$\alpha > 0$ $0 < \beta < 1$
Weibull	$\bar{F}(x) = e^{-cx^\tau}$	$c > 0$ $0 < \tau < 1$
Loggamma	$f(x) = \frac{\alpha^\beta}{\Gamma(\beta)} (\ln x)^{\beta-1} x^{-\alpha-1}$	$\alpha, \beta > 0$
Troncata $\alpha$ -stabile	$\bar{F}(x) = P( X  > x)$ con X variabile casuale $\alpha$ -stabile <sup>25</sup>	$1 < \alpha < 2$

Tabella 5- Funzioni di distribuzioni dell'ampiezza sinistri per "grandi" sinistri.

<sup>25</sup> Una variabile si definisce stabile se soddisfa la seguente relazione:

$$c_1 X_1 + c_2 X_2 \stackrel{d}{=} b(c_1, c_2) X + a(c_1, c_2)$$

per le variabili  $X, X_1, X_2$  indipendenti ed identicamente distribuite, per tutti i numeri  $c_1$  e  $c_2$  non negativi e per appropriati numeri reali  $b(c_1, c_2) > 0$  e  $a(c_1, c_2)$ .

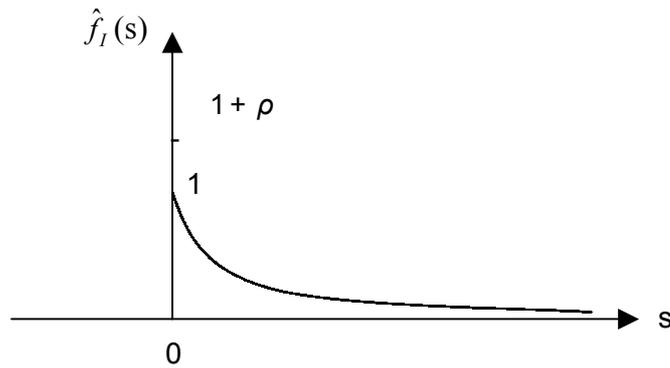


Figura 6- Il caso particolare della condizione di Cramér-Lundberg per i “grandi” sinistri.

Al fine di poter procedere risulta, però, necessario introdurre il concetto di *funzione regularly varying secondo Karamata*:

**Definizione (2.1):**

- una funzione positiva e misurabile  $L$  su  $(0, \infty)$  è *slowly varying ad  $\infty$*  (ossia  $L \in \mathcal{R}_0$ ) se

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

- una funzione positiva e misurabile  $h$  su  $(0, \infty)$  è *regularly varying ad  $\infty$*  con indice  $\alpha \in \mathfrak{R}$  (ossia  $h \in \mathcal{R}_\alpha$ ) se:

$$\lim_{x \rightarrow \infty} \frac{h(xt)}{h(x)} = t^\alpha, \quad t > 0.$$

Tipici esempi di funzioni *slowly varying* sono costanti positive o funzioni convergenti a costanti positive, logaritmi e logaritmi iterati; ad esempio

$$x^\alpha, x^\alpha \ln(1+x), (x \ln(1+x))^\alpha, x^\alpha \ln(\ln(e+x)), \quad \alpha \in \mathfrak{R}$$

sono funzioni *regularly varying ad  $\infty$*  con indice  $\alpha$ .

A questo punto è possibile enunciare il lemma che, insieme alla (II.2), ci consente di trovare un metodo alternativo a quello di Cramér-Lundberg ed applicabile agli eventi estremi:

**Lemma (2.1):** se  $F_1$  ed  $F_2$  sono due funzioni di distribuzione tali che  $\bar{F}_i(x) = x^{-\alpha}L_i(x)$ , per  $\alpha \geq 0$ , e  $L_i \in \mathcal{R}_0$ , con  $i = 1, 2$ , allora la convoluzione  $G = F_1 * F_2$  ha una coda regularly varying tale che:

$$\bar{G}(x) \sim x^{-\alpha}(L_1(x) + L_2(x)), \quad x \rightarrow \infty.$$

Tramite l'introduzione di  $n$  è dunque possibile ottenere il seguente

**Corollario (2.1):** se  $\bar{F}(x) = x^{-\alpha}L(x)$ , per  $\alpha \geq 0$  e  $L \in \mathcal{R}_0$ , allora, per tutte le  $n \geq 1$ ,

$$\bar{F}^{n^*}(x) \sim n\bar{F}(x), \quad x \rightarrow \infty$$

Si supponga ora che  $X_1, X_2, \dots, X_n$  siano variabili iid e che abbiano come funzione di distribuzione quella presentata nel corollario (2.1). Si indichi la loro somma parziale con  $S_n = X_1 + X_2 + \dots + X_n$  ed il loro massimo con  $M_n = \max(X_1, X_2, \dots, X_n)$ . Allora, per ogni  $n \geq 2$ :

$$P(S_n > x) = \bar{F}^{n^*}(x) \sim n\bar{F}(x), \quad x \rightarrow \infty$$

$$P(M_n > x) = \bar{F}^n(x) = \bar{F}(x) \sum_{k=0}^{n-1} F^k(x) \sim n\bar{F}(x), \quad x \rightarrow \infty$$

Di conseguenza, il precedente corollario ci permette di affermare che, se  $\bar{F} \in \mathcal{R}_{-\alpha}$ , essendo  $\alpha \geq 0$ , allora  $P(S_n > x) \sim P(M_n > x)$ , con  $x \rightarrow \infty$ .

Questo implica che, per funzioni di distribuzione con code *regularly varying*, la coda della funzione di distribuzione della somma  $S_n$  è principalmente determinata dalla coda della funzione di distribuzione del massimo  $M_n$ . Di conseguenza, per quanto riguarda i grandi sinistri, si può affermare che, sotto l'assunzione di variazione regolare, la coda del massimo determina la coda della somma.

Riprendendo il modello di Cramér-Lundberg, vale la seguente relazione:

$$\psi(u) = \frac{\rho}{1 + \rho} \sum_{n=0}^{\infty} (1 + \rho)^{-n} \bar{F}_1^{n^*}(u), \quad u \geq 0,$$

dove  $F_I(x) = \mu^{-1} \int_0^x \bar{F}(y) dy$  rappresenta la coda di distribuzione integrata. Sotto la condizione  $\bar{F}_1 \in \mathcal{R}_{-\alpha}$  per alcune  $\alpha \geq 0$ , vale la seguente stima asintotica:

$$\frac{\psi(u)}{\bar{F}_1(u)} = \frac{\rho}{1+\rho} \sum_{n=0}^{\infty} (1+\rho)^{-n} \frac{\bar{F}_1^{n^*}(u)}{\bar{F}_1(u)} \rightarrow \frac{\rho}{1+\rho} \sum_{n=0}^{\infty} (1+\rho)^{-n} n = \rho^{-1}, \quad u \rightarrow \infty.$$

Si può dunque concludere che, per la distribuzione dell'ampiezza sinistri con code *regularly varying*, la probabilità di rovina  $\psi(u)$ , per un ampio capitale iniziale  $u$ , è essenzialmente determinata dalla coda  $\bar{F}(y)$  della distribuzione dell'ampiezza sinistri per grandi valori di  $y$ , quindi:

$$\psi(u) \sim \frac{1}{\rho\mu} \int_0^{\infty} \bar{F}(y) dy, \quad u \rightarrow \infty.$$

Il risultato così ottenuto si mostra valido, in particolare, per le funzioni di distribuzione di Pareto, di Burr e la Loggamma.

Nel presente paragrafo si è dunque riusciti ad ottenere, a partire dal teorema di Cramér-Lundberg, una stima della probabilità di rovina anche per i sinistri di grande entità; questo è stato possibile grazie all'introduzione delle distribuzioni con code regolarmente variabili. Il medesimo obiettivo si può raggiungere tramite l'utilizzo di un'altra classe di distribuzioni: quella subesponenziale.

▫ **Il caso delle distribuzioni subesponenziali**

Nella precedente sezione è stato possibile ottenere una stima asintotica della probabilità di rovina, grazie alla cruciale proprietà secondo cui

$$\bar{F}_1^{n^*}(x) \sim n\bar{F}_1(x) \quad \text{per } x \rightarrow \infty \text{ ed } n \geq 2.$$

Questo ci porta direttamente alle funzioni di distribuzione subesponenziali, che ben rappresentano la distribuzione del danno nel caso di lunghe code; esse, infatti, giocano un ruolo fondamentale nell'analisi di variabili aleatorie che rappresentano grossi rischi, come, appunto, quelli catastrofali.

**Definizione (2.2):** una funzione di distribuzione  $F$  con supporto  $(0, \infty)$  è subesponenziale se, per ogni  $n \geq 2$ ,

$$\lim_{x \rightarrow \infty} \frac{\bar{F}^{n^*}(x)}{\bar{F}(x)} = n. \quad (II.4)$$

La classe delle funzioni subesponenziali viene indicata con  $S$ .

La (II.4) può anche essere riscritta nel seguente modo:

$$\lim_{x \rightarrow \infty} \frac{P(X_1 + X_2 + \dots + X_n > x)}{P(\max(X_1, X_2, \dots, X_n) > x)} = 1, \quad \text{per ogni } n \geq 2.$$

Quindi, le code della distribuzione della somma e di quella del massimo dei primi  $n$  sinistri sono, asintoticamente, dello stesso ordine e ciò indica la forte influenza del più ampio sinistro sul totale sinistri aggregato.

Dalle proprietà delle distribuzioni in considerazione, è possibile evincere che, per  $F \in S$ , la coda  $\bar{F}(x)$  decresce a 0 più lentamente di qualsiasi esponenziale del tipo  $e^{-\varepsilon x}$ ,  $\varepsilon > 0$ . Ciò rappresenta una prima caratteristica delle distribuzioni subesponenziali ed, in genere, di quelle *heavy-tailed*. Al contrario, le code delle funzioni di distribuzione utilizzate per il caso di piccole perdite, per le quali è possibile calcolare l'esponente di Lundberg, decrescono a 0 più velocemente di qualsiasi esponenziale del tipo  $e^{-\varepsilon x}$ ,  $\varepsilon > 0$ . La figura 7 rappresenta graficamente quanto appena detto.

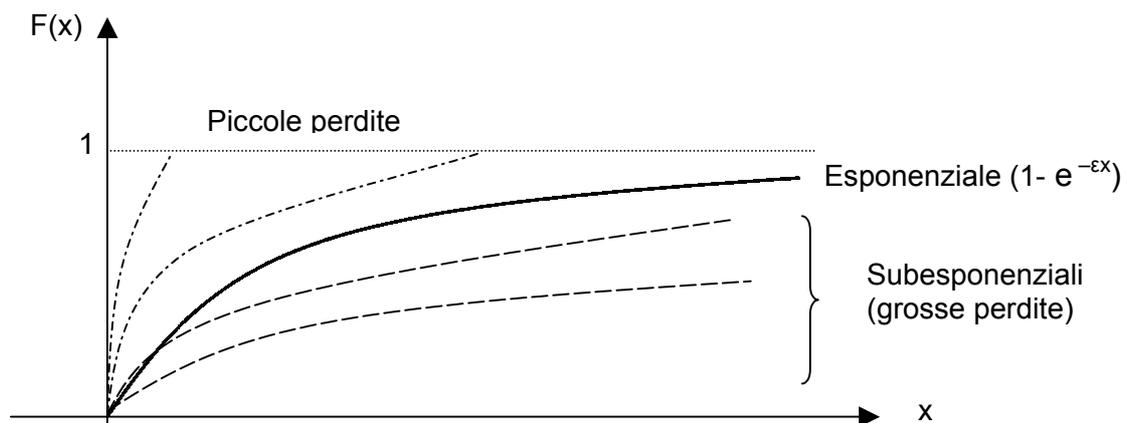


Figura 7- Funzioni di ripartizione per distribuzioni subesponenziali, esponenziale e per quelle rappresentanti i piccoli sinistri.

Ritornando al problema della stima della probabilità di rovina, la (II.4) ci permette di verificare immediatamente l'importantissima proprietà sopra citata:

$$P(S_n > x) \sim P(M_n > x), \quad \text{per ogni } n \geq 2 \text{ e con } x \rightarrow \infty$$

Allo scopo di verificare la subesponenzialità di una funzione non è necessario provare che la precedente relazione sia rispettata per ogni  $n$ , ma condizione sufficiente affinché la funzione  $F$  sia subesponenziale è la seguente:

$$\lim_{x \rightarrow \infty} \frac{\overline{F}^{2^*}(x)}{\overline{F}(x)} \leq 2.$$

Si ricordi che, per una funzione di distribuzione  $F$  con media finita  $\mu$ ,  $F_1(x) = \mu^{-1} \int_0^x \overline{F}(y) dy$ ; il seguente teorema rappresenta una immediata e fondamentale conseguenza di quanto sinora detto.

**Teorema di Cramér-Lundberg per i grandi sinistri (2.2):**

*si consideri il modello di Cramér-Lundberg, la condizione di profitto netto  $\rho > 0$  e la funzione di distribuzione  $F_1 \in S$ ; allora:*

$\psi(u) \sim \rho^{-1} \overline{F}_1(u), \quad u \rightarrow \infty.$	(II.5)
---	--------

Di conseguenza, la presente formula ci consente di ottenere la stima della rovina  $\psi(u)$ , per distribuzioni dell'ampiezza sinistri aventi distribuzioni integrate della coda subesponenziali.

In aggiunta alle funzioni di distribuzione con code *regularly varying*, le seguenti permettono di ottenere la stima rappresentata nella (II.5): Lognormale, Benktander 1 e 2 e Weibull (con  $0 < \tau < 1$ ).

Il teorema appena enunciato può essere ulteriormente incrementato, da un punto di vista matematico, come mostra il seguente risultato.

**Teorema (2.3):** *si supponga che  $0 < \alpha < 1$  e che le funzioni di distribuzione  $F$  e  $G$  siano relazionate nel seguente modo:*

$$G(x) = (1 - \alpha) \sum_{n=0}^{\infty} \alpha^n F^{n^*}(x);$$

allora le seguenti affermazioni sono equivalenti:

$$F \in \mathcal{S};$$

$$G \in \mathcal{S};$$

$$\lim_{x \rightarrow \infty} \frac{\overline{G}(x)}{\overline{F}(x)} = \frac{\alpha}{1-\alpha}.$$

A questo punto è così possibile scrivere la modifica del teorema di Cramér-Lundberg per i grandi sinistri.

**Teorema (2.4):** *si consideri il modello di Cramér-Lundberg e la condizione del profitto netto  $\rho > 0$ . Allora le seguenti affermazioni sono equivalenti:*

$$F_1 \in \mathcal{S};$$

$$1-\psi(u) \in \mathcal{S};$$

$$\lim_{x \rightarrow \infty} \frac{\psi(u)}{F_1(u)} = \rho^{-1}.$$

Di conseguenza, la stima (II.5) è possibile solo sotto la condizione  $F_1 \in \mathcal{S}$ . Nel caso della teoria di Cramér-Lundberg, dunque,  $\mathcal{S}$  è la classe naturale in cui poter determinare la probabilità di rovina, nel momento in cui la condizione di Cramér-Lundberg (II.3) è violata.

▫ **L'ammontare totale dei sinistri nel caso delle funzioni subesponenziali**

Nelle pagine precedenti, è stata evidenziata l'importanza della classe delle funzioni subesponenziali per la stima della probabilità di rovina, nel caso di "grandi" sinistri. Da un punto di vista matematico, è importante che, nel modello di Cramér-Lundberg,  $1-\psi(u)$  possa essere espresso come somma geometrica composta; a tal proposito, si veda la (II.2).

Precedentemente, si è osservato che, all'interno del modello di Cramér-Lundberg, per tutte le  $t \geq 0$ ,

$$G_t(x) = P(S(t) \leq x) = \sum_{n=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^n}{n!} F^{n*}(x), \quad x \geq 0, \quad (\text{II.6})$$

dove  $S(t)$  è l'ammontare sinistri aggregato, fino all'istante  $t$ . Il processo *numero dei sinistri*  $(N(t))_{t \geq 0}$ , contenuto nella precedente equazione, è un processo di Poisson omogeneo con intensità  $\lambda$ , quindi

$$P(N(t) = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}, \quad n \geq 0.$$

Gli stessi calcoli che conducono alla (II.6), portano, per un generale processo *numero dei sinistri* (ancora supposto indipendente dal processo di ampiezza dei sinistri  $(X_k)$ ), alla formula

$$G_t(x) = \sum_{n=0}^{\infty} p_t(n) F^{n^*}(x), \quad x \geq 0,$$

dove

$$p_t(n) = P(N(t) = n), \quad n \in \mathbb{N}_0$$

definisce una misura di probabilità su  $\mathbb{N}_0$ .

Risulta ora possibile enunciare il teorema inerente l'ammontare sinistri aggregato, per il caso subesponenziale.

**Teorema (2.5):** *si consideri  $G_t(x) = \sum_{n=0}^{\infty} p_t(n) F^{n^*}(x)$ , con  $x \geq 0$  ed  $F \in \mathcal{S}$ ; si fissi*

*un  $t > 0$  e si supponga che  $(p_t(n))$  soddisfi la seguente relazione:*

$$\sum_{n=0}^{\infty} (1 + \varepsilon)^n p_t(n) < \infty$$

*per alcune  $\varepsilon > 0$ . Allora  $G_t \in \mathcal{S}$  e*

$$\overline{G}_t(x) \sim E[N(t)] \overline{F}(x), \quad \text{per } x \rightarrow \infty$$

## 2.2 IL COMPORTAMENTO DELLE SOMME

Nel presente paragrafo viene fornita una teoria asintotica generale per le somme di variabili casuali indipendenti ed identicamente distribuite, rivolgendo, in particolare, l'attenzione verso il costo sinistri aggregato  $S(t)$ , introdotto nelle pagine precedenti per la definizione del processo di rischio  $U(t)$ . Prima di iniziare, però, la suddetta analisi, è necessario presentare alcuni risultati classici come la

legge dei grandi numeri e il teorema del limite centrale, fondamentali per la teoria asintotica.

### **2.2.1 La legge dei grandi numeri**

Nella presente sezione,  $X, X_1, X_2, \dots$  rappresenterà una sequenza di variabili indipendenti ed identicamente distribuite non degeneri, su uno spazio di probabilità  $[\Omega, \mathcal{F}, P]$ <sup>26</sup> con comune funzione di distribuzione  $F$ <sup>27</sup>. Se si desidera avere una idea delle oscillazioni delle  $X_n$  si può richiedere la convergenza della sequenza  $(X_n)$ . Sfortunatamente, per quasi tutti gli  $\omega \in \Omega$  tale sequenza non converge. D'altra parte, si possono ottenere informazioni su come le  $X_n$  si comportano "in media". Questo porta a considerare le somme parziali

$$S_0 = 0, \dots, S_n = X_1 + X_2 + \dots + X_n, \quad n \geq 1,$$

e la media aritmetica (o campionaria)

$$\bar{X}_n = n^{-1}S_n, \quad n \geq 1.$$

A livello intuitivo, una media aritmetica dovrebbe possedere una certa "stabilità" in  $n$ . Perciò ci si aspetta che, per valori ampi di  $n$ , il valore individuale  $X_i$  abbia meno influenza sul valore di  $\bar{X}_n$ , quindi che una sequenza  $(\bar{X}_n)$  converga ad un particolare valore, per  $n \rightarrow \infty$ . Tale effetto è noto sotto il nome di *legge dei grandi numeri*.

Supponiamo, per il momento, che  $\sigma^2 = \text{var}(X)$  sia finita e che  $\mu = E(X)$ . Dalla disuguaglianza di Chebyshev<sup>28</sup> si può concludere che, per  $\varepsilon > 0$ ,

$$P(|\bar{X}_n - \mu| > \varepsilon) \leq \varepsilon^{-2} \text{var}(\bar{X}_n) = (n\varepsilon^2)^{-1}\sigma^2 \rightarrow 0, \quad n \rightarrow \infty.$$

Quindi, per  $n \rightarrow \infty$ ,

<sup>26</sup> Si definisce spazio di probabilità una terna  $[\Omega, \mathcal{F}, P]$ , costituita da uno spazio misurabile  $(\Omega, \mathcal{F})$  – in cui  $\Omega$  rappresenta lo spazio campionario ed  $\mathcal{F}$  una famiglia di sottoinsiemi di uno stesso insieme  $\Omega$ - e da una misura di probabilità  $P$ .

<sup>27</sup> EMBRECHTS P., KLÜPPELBERG C., MIKOSCH T., *Modelling extremal events...*, op. cit., pagg. 60 e sgg.

<sup>28</sup> Secondo la disuguaglianza di Chebyshev-Bienaymé:

$$P(|X - E(X)| \geq a) \leq \frac{\text{VAR}(X)}{a^2}.$$

$$\bar{X}_n \xrightarrow{P} \mu.$$

Questa relazione viene chiamata *legge dei grandi numeri in senso debole* (o semplicemente *legge dei grandi numeri*) per la sequenza  $(\bar{X}_n)$ . Si è visto come tale risultato sia valido nel caso in cui la varianza  $\sigma^2$  sia finita; questa condizione può, però, essere sostanzialmente indebolita, come viene mostrato nel seguente

**Teorema (2.6):** *la legge dei grandi numeri in senso debole*

$$\bar{X}_n \xrightarrow{P} 0$$

vale se e solo se le due seguenti condizioni sono soddisfatte:

$$nP(|X| > n) \rightarrow 0;$$

$$E(X)I_{\{|X| \leq n\}} \rightarrow 0.$$

Dove:

$$I_A(\omega) = \begin{cases} 1 & \text{se } \omega \in A \\ 0 & \text{altrove} \end{cases}$$

rappresenta la funzione indicatrice dell'evento  $A$ .

Le assunzioni del teorema appena presentato sono facilmente verificate per la sequenza centrata  $(\bar{X}_n - \mu)$ . Di conseguenza, si può concludere che la legge dei grandi numeri in senso debole  $\bar{X}_n \xrightarrow{P} \mu$  vale se il valore atteso di  $X$  è finito.

Ora si può andare ad analizzare quali condizioni sono necessarie per assicurare che la successione  $\bar{X}_n$  non solo converga in probabilità, ma anche con probabilità 1, cioè quasi certamente (q.c.). Il risultato così ottenuto rappresenta la *legge dei grandi numeri in senso forte* per la sequenza  $(\bar{X}_n)$ . L'esistenza del primo momento è una condizione necessaria perché sia valida la legge in questione.

Dato che  $\bar{X}_n \xrightarrow{\text{q.c.}} a$  per alcune costanti finite  $a$ , si ha che

$$n^{-1} \bar{X}_n = n^{-1}(S_n - S_{n-1}) \xrightarrow{\text{q.c.}} a - a = 0.$$

Di conseguenza, per  $\varepsilon > 0$ ,

$$P\left(n^{-1}|X_n| > \varepsilon \text{ i.s.}\right) = 0.$$

Questo risultato, insieme al lemma di Borel-Cantelli<sup>29</sup>, implica che, per  $\varepsilon > 0$

$$\sum_{n=1}^{\infty} P\left(|n^{-1}X_n| > \varepsilon\right) = \sum_{n=1}^{\infty} P\left(|X| > \varepsilon n\right) < \infty,$$

e significa che  $E|X| < \infty$ .

Questa condizione risulta inoltre sufficiente perché sia verificata la legge dei grandi numeri in senso forte, come viene mostrato nel seguente teorema.

**Teorema di Kolmogorov (2.7):** *la legge dei grandi numeri in senso forte*

$$\bar{X}_n \xrightarrow{\text{q.c.}} a$$

*risulta valida per la sequenza  $(\bar{X}_n)$  per alcune costanti reali  $a$ , se e solo se  $E|X| < \infty$ . Inoltre, se  $(\bar{X}_n)$  rispetta la legge dei grandi numeri in senso forte, allora  $a = \mu$ .*

La legge dei grandi numeri intesa in senso forte fornisce, quasi certamente, una approssimazione del primo ordine della variabile casuale  $\bar{X}_n$ , tramite la quantità deterministica  $\mu$ :

$$\bar{X}_n = \mu + o(1) \quad \text{q.c.}$$

Viene ora presentato un teorema in cui si analizzano le fluttuazioni quasi certe delle somme  $S_n$ , nel caso si dovessero utilizzare altre sequenze di normalizzazione, quali le potenze di  $n$ .

**Teorema di Marcinkiewicz-Zygmund (2.8):**

*si supponga che  $p \in (0,2)$ . La legge dei grandi numeri in senso forte*

$$n^{-1/p}(S_n - an) \xrightarrow{\text{q.c.}} 0$$

<sup>29</sup> Il lemma di Borel-Cantelli afferma che

$$\sum_{n=1}^{\infty} P(A_n) < \infty \text{ implica che } P(A_n, \text{infinitamente spesso}) = 0.$$

Il suo parziale converso, per  $A_n$  indipendenti, dice che

$$\sum_{n=1}^{\infty} P(A_n) = \infty \text{ implica che } P(A_n, \text{infinitamente spesso}) = 1$$

vale per alcune costanti reali  $a$  se e solo se  $E|X|^p < \infty$ . Se  $(X_n)$  obbedisce a tale legge, allora si può scegliere

$$a = \begin{cases} 0 & \text{se } p < 1 \\ \mu & \text{se } p \in [1,2) \end{cases}$$

Inoltre, se  $E|X|^p = \infty$  per alcune  $p \in (0,2)$ , allora per ogni  $a$  reale

$$\limsup_{n \rightarrow \infty} n^{-1/p} |S_n - an| = \infty, \quad \text{q.c.}$$

Il presente teorema fornisce una completa caratterizzazione della legge dei grandi numeri in senso forte, utilizzando come sequenza di normalizzazione le potenze di  $n$  e ci consente, inoltre, di ottenere una rifinita approssimazione, quasi certa, del primo ordine di  $\bar{X}_n$ :

$$\bar{X}_n = \mu + o(n^{(1/p)-1}) \quad \text{q.c.}$$

che è valida se  $E|X|^p < \infty$ , per alcuni  $p \in [1,2)$ .

Si può, infine, fornire un'ultima rifinitura della legge in questione, nel caso in cui esista il momento secondo di  $X$ ; a tale scopo contribuisce il seguente teorema, detto anche *legge del logaritmo iterato*.

**Teorema di Hartman-Wintner (2.9):** se  $\sigma^2 = \text{var}(X) < \infty$ , allora

$$\limsup_{n \rightarrow \infty} (2n \ln \ln n)^{-1/2} (S_n - \mu n) = -\liminf_{n \rightarrow \infty} (2n \ln \ln n)^{-1/2} (S_n - \mu n) = \sigma \quad \text{q.c.}$$

Se  $\sigma^2 = \infty$ , allora per ogni sequenza reale  $(a_n)$

$$\limsup_{n \rightarrow \infty} (2n \ln \ln n)^{-1/2} |S_n - a_n| = \infty \quad \text{q.c.}$$

Dunque, la legge del logaritmo iterato restituisce l'approssimazione quasi certa del primo ordine

$$\bar{X}_n = \mu + O[(\ln \ln n)^{-1/2}] \quad \text{q.c.}$$

che rappresenta la migliore approssimazione possibile di  $\bar{X}_n$ , tramite il suo valore atteso  $\mu$ .

### 2.2.2 Il teorema del limite centrale

Un problema classico nella teoria della probabilità è rappresentato dalla ricerca delle possibili leggi limite per la somma  $S_n$ , propriamente normalizzata e centrata. Questo equivale a studiare quali siano le distribuzioni che soddisfano l'identità in legge

$$c_1 X_1 + c_2 X_2 \stackrel{d}{=} b(c_1, c_2) X + a(c_1, c_2) \quad (\text{II.7})$$

per tutti i numeri  $c_1, c_2$  non negativi e appropriati numeri reali  $b(c_1, c_2) > 0$  e  $a(c_1, c_2)$ .

Le possibili leggi limite per somme di variabili indipendenti ed identicamente distribuite sono semplicemente le funzioni che soddisfano la precedente equazione per tutti i numeri  $c_1, c_2$  non negativi e vengono definite *distribuzioni* (o *funzioni di distribuzione*) *stabili*. Si consideri, dunque, la somma  $S_n$ ; dalla (II.7) si ha che, per alcune costanti reali  $a_n$  e  $b_n > 0$  e per  $X = X_1$ ,

$$S_n = X_1 + X_2 + \dots + X_n \stackrel{d}{=} b_n X + a_n, \quad n \geq 1$$

che si può anche riscrivere nel seguente modo

$$b_n^{-1} (S_n - a_n) \stackrel{d}{=} X.$$

Di conseguenza, si può affermare che, se una distribuzione è stabile, allora è una distribuzione limite per somme di variabili indipendenti ed identicamente distribuite. Inoltre, la classe delle distribuzioni stabili coincide con la classe di tutte le possibili leggi limite per somme di variabili indipendenti ed identicamente distribuite.

A causa dell'importanza della classe delle distribuzioni stabili, è necessario descriverle analiticamente, tramite la loro funzione caratteristica.

**Teorema (2.10):** *Una distribuzione stabile ha funzione caratteristica*

$$\varphi_X(t) = E[\exp\{iXt\}] = \exp\{i\gamma t - c|t|^\alpha [1 - i\beta \text{sign}(t) z(t, \alpha)]\}, \quad t \in \mathfrak{R},$$

dove  $\gamma$  è una costante reale,  $c > 0$ ,  $\alpha \in (0, 2]$ ,  $\beta \in (-1, 1]$ , e

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

La quantità  $\gamma$  rappresenta un parametro di locazione e d'ora in poi la si considererà nulla. Il valore più importante è costituito da  $\alpha$ , che determina le proprietà di base della classe di distribuzione (quali, ad esempio, i momenti, le code e il comportamento asintotico delle somme). Il parametro in questione viene anche definito *esponente caratteristico* e la corrispondente *distribuzione* è definita  $\alpha$ -*stabile*. Nel caso sia  $\alpha = 2$ , si ottiene una distribuzione normale, mentre, se si pone  $\alpha = 1$  si ha la distribuzione di Cauchy. Infine, i parametri  $c$  e  $\beta$  determinano la natura della distribuzione:  $c$  è una costante di scala, mentre  $\beta$  descrive l'asimmetria della distribuzione.

Una volta definita la classe delle funzioni  $\alpha$ -stabili (di nostro interesse in quanto coincide con la classe di tutte le possibili leggi limite per somme di variabili indipendenti ed identicamente distribuite), si prosegue ora indagando su quali siano le condizioni che, data una distribuzione  $\alpha$ -stabile  $G_\alpha$ , implicano che la somma  $S_n$ , normalizzata e centrata, converga debolmente a  $G_\alpha$ . Questo significa decidere come scegliere le costanti  $a_n \in \mathfrak{R}$  e  $b_n > 0$  tali che

$$b_n^{-1}(S_n - a_n) \xrightarrow{d} G_\alpha. \quad (\text{II.8})$$

Prima di tutto, però, risulta necessario introdurre la nozione di *domain of attraction*:

**Definizione (2.3):** *si dice che una variabile casuale  $X$  (o la sua funzione di distribuzione  $F$ ) appartiene al domain of attraction di una distribuzione  $\alpha$ -stabile  $G_\alpha$  se esistono costanti  $a_n \in \mathfrak{R}$  e  $b_n > 0$ , tali che valga*

$$b_n^{-1}(S_n - a_n) \xrightarrow{d} G_\alpha.$$

*Si può scrivere  $X \in DA(G_\alpha)$  (o  $F \in DA(G_\alpha)$ ) e dire che  $(X_n)$  soddisfa il teorema del limite centrale con limite  $G_\alpha$ .*

Si introducano ora le quantità:

$$K(x) = x^{-2} \int_{|y| \leq x} y^2 dF(y), \quad x > 0$$

$$Q(x) = G(x) + K(x) = P(|X| > x) + K(x), \quad x > 0.$$

Risulta, a questo punto, possibile dire quale valore devono assumere i parametri  $a_n$  e  $b_n$ , perché la (II.8) valga. A questo scopo, vengono enunciate le due seguenti proposizioni:

**Proposizione (2.1):** *le costanti di normalizzazione  $b_n$ , nel teorema del limite centrale, per  $F \in DA(\alpha)$ , possono essere scelte come unica soluzione dell'equazione*

$$Q(b_n) = n^{-1}, \quad n \geq 1.$$

*In particolare, se  $\sigma^2 = \text{var}(X) < \infty$  e  $E(X) = 0$ , allora*

$$b_n \sim n^{1/2} \sigma, \quad n \rightarrow \infty.$$

*Se  $\alpha < 2$ , in alternativa, si può scegliere ( $b_n$ ) in modo tale che*

$$b_n = \inf \{ y : G(y) < n^{-1} \}, \quad n \geq 1.$$

*Risulta importante notare che quest'ultima equazione implica la seguente :*

$$G(b_n) \sim n^{-1}, \quad n \rightarrow \infty,$$

*e che, per  $n \geq 1$*

$$b_n = n^{1/\alpha} L(n),$$

*per una appropriata funzione slowly varying  $L$ .*

**Proposizione (2.2):** *le costanti di centramento  $a_n$  nel teorema del limite centrale (II.8) possono essere scelte nella seguente forma*

$$a_n = n \int_{|y| \leq b_n} y dF(y),$$

*dove  $b_n$  è dato dalla precedente proposizione. In particolare, si può scegliere*

*$a_n = \tilde{a}n$ , dove:*

$$\tilde{a} = \begin{cases} \mu & \text{se } \alpha \in (1, 2] \\ 0 & \text{se } \alpha \in (0, 1) \\ 0 & \text{se } \alpha = 1 \text{ ed } F \text{ è simmetrica} \end{cases}$$

Infine, una volta stabiliti quali valori possono assumere i parametri, è possibile enunciare il seguente teorema.

**Teorema generale del limite centrale (2.11):** *si supponga che  $F \in DA(\alpha)$  per alcune  $\alpha \in (0, 2]$ .*

a) *Se  $E[X^2] < \infty$ , allora*

$$(\sigma n^{1/2})^{-1} (S_n - \mu n) \xrightarrow{d} \Phi$$

*per la distribuzione normale standard  $\Phi$ , avente media zero e varianza unitaria.*

b) *Se  $E[X^2] = \infty$  e  $\alpha = 2$ , o se  $\alpha < 2$ , allora*

$$[n^{1/\alpha} L(n)]^{-1} (S_n - a_n) \xrightarrow{d} G_\alpha$$

*per una distribuzione  $\alpha$ -stabile  $G_\alpha$ , una appropriata funzione slowly varying  $L$  e appropriate costanti di centramento, come definite precedentemente. In particolare,*

$$[n^{1/\alpha} L(n)]^{-1} (S_n - \tilde{a}n) \xrightarrow{d} G_\alpha.$$

### 2.2.3 Somme aleatorie

Nella matematica assicurativa, le somme casuali assumono un ruolo fondamentale; basti, infatti, pensare al solo fatto che l'ammontare totale dei sinistri di un portafoglio assicurativo è solitamente modellato da somme casuali, come si è avuto modo di vedere nelle pagine precedenti. Infatti

$$S(t) = S_{N(t)} = \begin{cases} 0 & \text{se } N(t) = 0 \\ X_1 + X_2 + \dots + X_{N(t)} & \text{se } N(t) \geq 1 \end{cases}$$

dove  $(N(t))_{t \geq 0}$  è un processo stocastico su  $[0, \infty)$ , tale che le variabili casuali  $N(t)$  siano valori interi non negativi. Solitamente, si assume che  $(N(t))$  sia generato da una sequenza di variabili non negative  $(T_n)_{n \geq 1}$ , tale che

$$0 \leq T_1 \leq T_2 \leq \dots \quad \text{q.c.}$$

e

$$N(t) = \sup \{ n \geq 1 : T_n \leq t \}, \quad \text{per } t \geq 0.$$

Questo è anche definito un processo *counting*, ossia di conteggio. Le variabili casuali  $X_n$  possono essere interpretate come sinistri individuali che accadono in istanti casuali  $T_n$ ,  $N(t)$  conta il numero totale di sinistri individuali e  $S(t)$  rappresenta l'ammontare totale sinistri nel portafoglio, fino all'istante  $t$ .

Nel modello di Cramèr-Lundberg definito nel paragrafo precedente, si assumeva che  $(N(t))$  fosse un processo omogeneo di Poisson, con parametro di intensità  $\lambda$  e con

$$T_n = Y_1 + Y_2 + \dots + Y_n, \quad n \geq 1,$$

dove  $(Y_n)$ , ossia gli intervalli di tempo che intercorrono tra un sinistro e quello successivo, sono variabili casuali indipendenti ed identicamente distribuite, con valore atteso pari a  $1/\lambda$ . Qualsiasi *counting process* generato da un processo di somme indipendenti ed identicamente distribuite  $(T_n)$  viene definito un processo *renewal counting*. Altrimenti, un *processo di Poisson omogeneo* è definito dalle seguenti tre proprietà:

- inizia da zero:  $N(0) = 0$ ;
- ha incrementi indipendenti e stazionari;
- per ogni  $t > 0$ ,  $N(t)$  è una variabile casuale Poisson, con parametro  $\lambda t$ :

$$P ( N(t) = n ) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}, \quad n = 0, 1, 2, \dots$$

Se  $(N(t))$  e  $(X_n)$  sono indipendenti, allora il processo  $(S(t))_{t \geq 0}$  è detto *processo di Poisson composto*.

Le oscillazioni delle somme casuali  $S(t)$ , per valori ampi di  $t$ , possono ancora essere descritte tramite teoremi sui limiti. Nelle prossime pagine, vengono forniti alcuni concetti base, che permettono di mostrare lo stretto legame tra il comportamento di  $(S_n)$  e quello di  $(S(t))$ .

Siano, d'ora in poi,  $(Z_n)_{n \geq 0}$  una sequenza generale di variabili casuali ed  $(N(t))_{t \geq 0}$  un processo di variabili casuali intere e non negative  $N(t)$ . Per poter applicare la legge dei grandi numeri, intesa in senso forte, anche alla somma di variabili casuali, è prima necessario enunciare il seguente lemma.

**Lemma (2.2):** *si supponga che  $Z_n \xrightarrow{q.c.} Z$  per  $n \rightarrow \infty$  e che  $N(t) \xrightarrow{q.c.} \infty$  ( $N(t) \xrightarrow{P} \infty$ ) per  $t \rightarrow \infty$ . Allora*

$$Z_{N(t)} \xrightarrow{\text{q.c.}} Z \quad (Z_{N(t)} \xrightarrow{P} Z), \quad \text{per } t \rightarrow \infty.$$

Grazie al lemma appena presentato, è ora possibile enunciare la variante della legge dei grandi numeri intesa in senso forte di Marcinkiewicz-Zygmund, per quanto riguarda le somme casuali.

**Teorema (2.12):** si supponga che  $E|X|^p < \infty$  per alcune  $p \in (0,2)$  e che

$N(t) \xrightarrow{\text{q.c.}} \infty$ . Allora

$$(N(t))^{-1/p} [S(t) - aN(t)] \xrightarrow{\text{q.c.}} 0,$$

dove

$$a = \begin{cases} 0 & \text{se } p < 1 \\ \mu = E(X) & \text{se } p \in [1,2) \end{cases}$$

Una volta analizzata la legge dei grandi numeri, per il caso in esame, si può ora analizzare il problema del teorema del limite centrale. Prima, però, è necessario fare alcune valutazioni inerenti la convergenza debole, tramite il seguente lemma.

**Lemma (2.3):** si supponga che  $(Z_n)$  e  $(N(t))$  siano indipendenti e che

$N(t) \xrightarrow{P} \infty$ , per  $t \rightarrow \infty$ . Se  $Z_n \xrightarrow{d} Z$  per  $n \rightarrow \infty$ , allora  $Z_{N(t)} \xrightarrow{d} Z$  per  $t \rightarrow \infty$ .

Grazie a queste considerazioni, è ora possibile enunciare il teorema generale del limite centrale, per somme casuali.

**Teorema (2.13):** si supponga che  $(X_n)$  e  $(N(t))$  siano indipendenti e che

$N(t) \xrightarrow{P} \infty$ . Si assuma che  $F \in DA(\alpha)$ , per alcune  $\alpha \in (0,2]$ . Allora il teorema generale del limite centrale<sup>30</sup> rimane valido se  $n$  è ovunque sostituita da  $N(t)$ , quindi se esistono costanti di centramento  $a_n$  ed una funzione slowly varying  $L$  tali che

$$[(N(t))^{1/\alpha} L(N(t))]^{-1} (S(t) - a_{N(t)}) \xrightarrow{d} G_\alpha, \quad \text{per } t \rightarrow \infty,$$

<sup>30</sup> Il teorema in questione è il (2.11), precedentemente enunciato, a pagina 57.

per una funzione  $\alpha$ -stabile  $G_\alpha$ .

La condizione secondo la quale i processi  $(Z_n)$  ed  $(N(t))$  sono indipendenti può essere notevolmente allentata, come mostra il seguente teorema.

**Teorema di Anscombe (2.14):** *si supponga che esista una funzione a valori interi  $b(t) \uparrow \infty$ , tale che*

$$\frac{N(t)}{t} \xrightarrow{P} 1, \quad \text{per } t \rightarrow \infty, \quad (II.9)$$

e che la seguente, cosiddetta condizione di Anscombe, valga:

$\forall \varepsilon > 0, \forall \eta > 0$ , esistono  $\delta > 0$  ed  $n_0$ , tali che

$$P \left( \max_{m: |m-n| < n\delta} |Z_m - Z_n| > \varepsilon \right) < \eta, \quad \text{con } n > n_0 \quad (II.10)$$

Se  $Z_n \xrightarrow{d} Z$  per  $n \rightarrow \infty$ , allora  $Z_{N(t)} \xrightarrow{d} Z$  per  $t \rightarrow \infty$ .

La condizione (II.9) assicura che l'indice casuale  $N(t)$  possa essere sostituito dalla funzione deterministica  $b(t)$ ; se, ad esempio, si sostituisce  $Z_{N(t)}$  a  $Z_{b(t)}$ , allora la (II.10) garantisce che l'errore  $|Z_{N(t)} - Z_{b(t)}|$  sia trascurabile. In particolare, si può notare che la condizione (II.9) risulta valida per ampie classi di processi *renewal counting*, incluso il processo di Poisson omogeneo, mentre la (II.10) vale per le somme  $S_n$  propriamente centrate e normalizzate. Questo è il contenuto del seguente *teorema del limite centrale di Anscombe per somme casuali*.

**Teorema (2.15):** *si supponga che*

$$\frac{N(t)}{t} \xrightarrow{P} \lambda, \quad \text{per } t \rightarrow \infty,$$

per alcuni valori positivi di  $\lambda$  e che  $F \in DA(\alpha)$  per alcuni  $\alpha \in (0, 2]$  con

$$(n^{1/\alpha} L(n))^{-1} (S_n - \tilde{a}n) \xrightarrow{d} G_\alpha,$$

per una distribuzione  $\alpha$ -stabile  $G_\alpha$ , una funzione slowly varying  $L$  e

$$\tilde{a} = \begin{cases} 0 & \text{se } \alpha \leq 1 \\ \mu & \text{se } \alpha \in (1, 2] \end{cases}$$

Allora:

$$[(N(t))^{1/\alpha} L(N(t))]^{-1} [S(t) - \tilde{a}N(t)] \xrightarrow{d} G_\alpha,$$

$$[(\lambda t)^{1/\alpha} L(t)]^{-1} [S(t) - \tilde{a}N(t)] \xrightarrow{d} G_\alpha.$$

In particolare, se  $\sigma^2 < \infty$ , allora

$$(\lambda \sigma^2 t^{-1/2} (S(t) - \mu N(t))) \xrightarrow{d} \Phi$$

dove  $\Phi$  rappresenta la distribuzione normale standard

▫ **I processi renewal counting**

Nella presente sezione, si andranno ad esaminare la legge dei grandi numeri ed il teorema del limite centrale, per quanto riguarda il caso più particolare del processo *renewal counting*  $((N(t)))_{t \geq 0}$ , dove

$$N(t) = \sup \{ n \geq 1 : T_n \leq t \}, \quad \text{per } t \geq 0$$

e

$$T_n = Y_1 + Y_2 + \dots + Y_n, \quad n \geq 1,$$

per variabili casuali indipendenti ed identicamente distribuite non negative  $Y, Y_1, Y_2, \dots$ . Il processo di Poisson omogeneo è un processo *renewal counting* dove  $Y$  è esponenziale, con valore atteso  $1/\lambda$ .

Nella presente sezione si cercherà di determinare quale sia l'ordine di grandezza di  $N(t)$ , nel caso in cui  $t \rightarrow \infty$ . Si osservi, innanzitutto, che

$$\{T_n \leq t\} = \{N(t) \geq n\}.$$

La legge dei grandi numeri intesa in senso forte, nella versione di Kolmogorov, implica che  $T_n \xrightarrow{\text{q.c.}} \infty$  e che perciò  $N(t) \xrightarrow{\text{q.c.}} \infty$ . D'altra parte, grazie alla *legge dei grandi numeri intesa in senso forte di Marcinkiewicz-Zygmund* (detta anche *legge dei logaritmi iterati per processi renewal counting*), si possono ottenere delle informazioni molto più precise.

**Teorema (2.16):** *si supponga che  $E(Y) = 1/\lambda < \infty$  (se  $E(Y) = \infty$  si ponga  $\lambda = 0$ ). Allora*

$$t^{-1} N(t) \xrightarrow{\text{q.c.}} \lambda.$$

Se  $E(Y^p) < \infty$  per alcune  $p \in (1,2)$  allora

$$t^{-1/p}(N(t) - \lambda t) \xrightarrow{q.c.} 0.$$

Se  $\sigma_Y^2 = \text{var}(Y) < \infty$  allora

$$\limsup_{t \rightarrow \infty} (2t \ln \ln t)^{-1/2} (N(t) - \lambda t) = -\liminf_{t \rightarrow \infty} (2t \ln \ln t)^{-1/2} (N(t) - \lambda t) = \sigma_Y \lambda^{3/2} \quad \text{q.c.}$$

Questo teorema suggerisce che  $E(N(t)) \sim \lambda t$  e che  $\text{var}(N(t)) = \sigma_Y^2 \lambda^3 t$ , dal momento in cui  $E(Y) = 1/\lambda$  e  $\sigma_Y^2 = 1/\lambda^2$ .

Una volta analizzata la legge dei grandi numeri per il caso in analisi, è possibile passare a considerare il teorema del limite centrale. Prima, però, per avere una idea della crescita del processo *renewal counting*, si ritiene necessario presentare i momenti del processo in esame.

**Proposizione (2.3):** *valgano le seguenti relazioni:*

- $E(N(t)) = (\lambda + o(1))t$  se  $t \rightarrow \infty$ ;
- si supponga che  $\sigma_Y^2 = \text{var}(Y) < \infty$ . Allora

$$E(N(t)) = \lambda t + O(1), \quad t \rightarrow \infty;$$

$$\text{var}(N(t)) = \sigma_Y^2 \lambda^3 t + o(t), \quad t \rightarrow \infty.$$

Risulta, a questo punto, possibile enunciare *il teorema del limite centrale per il processo renewal counting*.

**Teorema (2.17):** *si supponga che  $\sigma_Y^2 < \infty$ . Allora*

$$(\sigma_Y^2 \lambda^3 t)^{-1/2} (N(t) - \lambda t) \xrightarrow{d} \Phi,$$

dove  $\Phi$  rappresenta la distribuzione normale standard.

Ricordando che, dall'ultima proposizione, si era ricavato che  $E(N(t)) \sim \lambda t$  e che  $\text{var}(N(t)) = \sigma_Y^2 \lambda^3 t$ , questo teorema risulta di conseguenza simile a quello del limite centrale classico, per le somme indipendenti ed identicamente distribuite.

▫ **Somme aleatorie condotte da processi *renewal counting***

Nella presente sezione, viene trattato uno dei più importanti modelli della matematica assicurativa. In essa si assume che le somme casuali  $S(t) = S_{N(t)}$  sono condotte da un processo *renewal counting*  $N(t)$  (che è stato definito in precedenza). Il processo  $(S(t))$  rappresenta l'ammontare totale sinistri di un portafoglio assicurativo. I modelli *renewal* e di Cramér-Lundberg sono inclusi in questo campo come casi particolari in cui  $(N(t))$  ed  $(X_n)$  sono indipendenti. In generale, però, non si richiede che tale ipotesi sia verificata. Nella presente sezione si pone l'attenzione sulle proprietà asintotiche del processo  $(S(t))$ .

Si richiami dalla sezione precedente che  $N(t) \xrightarrow{q.c.} \infty$ . Di conseguenza, è possibile applicare la legge dei grandi numeri intesa in senso forte di Marcinkiewicz-Zygmund, per le somme casuali<sup>31</sup> :

$$(N(t))^{-1/p} [S(t) - aN(t)] \xrightarrow{q.c.} 0, \quad (II.11)$$

se  $E|X|^p < \infty$ , per alcune  $p \in (0,2)$  e

$$a = \begin{cases} 0 & \text{se } p < 1 \\ \mu = E(X) & \text{se } p \in (1,2] \end{cases} \quad (II.12)$$

Si va ora ad analizzare se è possibile sostituire  $N(t)$ , nella (II.11), con una funzione deterministica, come ad esempio  $\lambda t$ . Solitamente, questo non è consentito; d'altra parte, però, dalla legge dei logaritmi iterati applicata al processo *renewal counting*,  $N(t)/t \xrightarrow{q.c.} \lambda$ , se  $E(Y) < \infty$ . Dunque, si potrebbe sostituire il processo di normalizzazione  $(N(t))^{1/p}$  con  $(\lambda t)^{1/p}$ . Il processo di centramento causa alcuni problemi. Per procedere si supponga che  $E|X|^p < \infty$  per alcune  $p \in [1,2)$ . Si può quindi scrivere

$$t^{-1/p}(S(t) - \lambda \mu t) = t^{-1/p}(S(t) - \mu N(t)) + \mu t^{-1/p}(N(t) - \lambda t).$$

---

<sup>31</sup> Il teorema in questione è il (2.12), enunciato a pagina 59.

Grazie alla (II.11), il primo addendo dell'equazione converge a zero quasi certamente, una volta posto che il primo momento di  $X$  sia finito. D'altra parte, però,

$$\mu t^{-1/p}(N(t) - \lambda t) \xrightarrow{\text{q.c.}} 0,$$

in generale, non vale. Se, però,  $E|Y|^p < \infty$ , si può concludere, dal teorema della legge dei logaritmi iterati applicata al processo *renewal counting*, che la (II.11) è soddisfatta. In sintesi, si ottiene così una modifica del teorema appena richiamato:

**Teorema (2.18):** *si supponga che  $E|X|^p < \infty$  per alcuni  $p \in (0,2)$ .*

- Se  $E(Y) < \infty$ , allora

$$t^{-1/p}(S(t) - aN(t)) \xrightarrow{\text{q.c.}} 0,$$

dove  $a$  è definita dalla (II.12).

- Se  $p \geq 1$  e  $E|Y|^p < \infty$ , allora

$$t^{-1/p}(S(t) - \lambda \mu t) \xrightarrow{\text{q.c.}} 0.$$

Analizzata la legge dei grandi numeri, per il caso in questione, è possibile passare all'analisi del teorema del limite centrale. Per quanto riguarda la convergenza debole, si può procedere in modo analogo a quello appena presentato. Dal momento in cui  $N(t) \rightarrow \infty$  quasi certamente, per un processo *renewal counting*, il *teorema del limite centrale per somme casuali* viene applicato sotto miti condizioni, ossia *nella variante di Anscombe*.

**Teorema (2.19):** *si assuma che  $F \in DA(\alpha)$  per alcune  $\alpha \in (0,2]$  e che*

$$[n^{1/\alpha}L(n)]^{-1}[S_n - \tilde{a}n] \xrightarrow{d} G_\alpha,$$

*per alcune distribuzioni  $\alpha$ -stabili  $G_\alpha$  e una funzione slowly varying  $L$ . In questo caso*

$$\tilde{a} = \begin{cases} 0 & \text{se } \alpha \leq 1 \\ \mu & \text{se } \alpha \in (1,2] \end{cases}$$

- Se  $E(Y) < \infty$ , allora

$$[(\lambda t)^{1/\alpha} L(t)]^{-1} [S(t) - \tilde{a}N(t)] \xrightarrow{d} G_\alpha.$$

In particolare, se  $\sigma^2 = \text{var}(X) < \infty$ , allora

$$(\lambda \sigma^2 t)^{-1/2} [S(t) - \mu N(t)] \xrightarrow{d} \Phi,$$

dove  $\Phi$  rappresenta la distribuzione normale standard.

- Se  $\alpha \in (1,2)$  e  $E(Y) < \infty$ , per alcune  $p > \alpha$ , allora

$$[(\lambda t)^{1/\alpha} L(t)]^{-1} [S(t) - \lambda \mu t] \xrightarrow{d} G_\alpha.$$

Avendo analizzato il teorema del limite centrale, per quanto riguarda le somme di variabili stocastiche, è ora possibile combinare tale risultato per  $(S(t))$  e  $(N(t))$ . Si assuma, quindi, che  $\sigma^2 < \infty$  e che  $\sigma_Y^2 < \infty$ . Si può affermare che

$$\begin{aligned} t^{-1/2} (S(t) - \lambda \mu t) &= t^{-1/2} [(S(t) - \mu N(t)) + \mu (N(t) - \lambda t)] = \\ &= t^{-1/2} [(S(t) - \mu N(t)) + \mu (N(t) - \lambda T_{N(t)})] + o_P(1) = \\ &= t^{-1/2} \sum_{i=1}^{N(t)} (X_i - \mu \lambda Y_i) + o_P(1). \end{aligned} \quad (\text{II.13})$$

Si noti che le variabili casuali  $X'_i = X_i - \mu \lambda Y_i$  hanno media zero. Inoltre, la sequenza  $(X'_i)$  è indipendente ed identicamente distribuita se anche  $((X_n, Y_n))$  lo è. Sotto quest'ultima condizione, il teorema del limite centrale di Anscombe per somme aleatorie<sup>32</sup> viene immediatamente applicato alla (II.13) e porta al seguente risultato.

**Teorema (2.20):** *si supponga che  $((X_n, Y_n))$  sia una sequenza di vettori indipendenti ed identicamente distribuiti, che  $\sigma^2 < \infty$  e che  $\sigma_Y^2 < \infty$ . Allora*

$$[\text{var}(X - \mu \lambda Y) \lambda t]^{-1/2} (S(t) - \mu \lambda t) \xrightarrow{d} \Phi,$$

dove  $\Phi$  rappresenta la distribuzione normale standard.

In particolare, se  $(X_n)$  e  $(Y_n)$  sono indipendenti, allora

$$[(\sigma^2 + (\mu \lambda \sigma_Y)^2) \lambda t]^{-1/2} (S(t) - \mu \lambda t) \xrightarrow{d} \Phi.$$

<sup>32</sup> Il teorema in questione è il (2.15), enunciato alle pagine 60-61.

### 2.3 L'EXTREME VALUE THEORY

Nel presente paragrafo viene trattata la *teoria del valore estremo*, un concetto fondamentale ai fini dell'analisi di quegli eventi che generano grandi scosse all'interno del sistema finanziario, quali le catastrofi. Il risultato centrale è rappresentato dal teorema di Fisher-Tippett, nel quale viene specificata la forma della distribuzione limite per massimi centrati e normalizzati. Le tre famiglie di possibili leggi limite sono note sotto il nome di *distribuzioni del valore estremo*.

Si consideri una sequenza  $X, X_1, X_2, \dots$  di variabili indipendenti ed identicamente distribuite non degeneri, con comune funzione di distribuzione  $F$ .<sup>33</sup> Viene ora analizzata la successione dei massimi campionari

$$M_1 = X_1, \dots, M_n = \max ( X_1, \dots, X_n ), \quad n \geq 2.$$

Per quanto riguarda i minimi, si possono facilmente ottenere risultati corrispondenti, tramite la seguente identità

$$\min ( X_1, \dots, X_n ) = - \max ( -X_1, \dots, -X_n ).$$

Si può, quindi, scrivere la funzione di distribuzione del massimo  $M_n$ :

$$P( M_n \leq x ) = P( X_1 \leq x, \dots, X_n \leq x ) = F^n ( x ), \quad x \in \mathfrak{R}, n \in \mathfrak{N}.$$

Gli estremi accadono in prossimità del limite superiore del supporto della distribuzione, di conseguenza, intuitivamente, il comportamento asintotico di  $M_n$  deve essere relazionato alla coda destra della funzione di distribuzione  $F$ , vicino dunque al suo estremo superiore. Si indichi con

$$x_F = \sup \{ x \in \mathfrak{R} : F(x) < 1 \}$$

l'*estremo superiore di F*. Si ottiene immediatamente che, per ogni  $x < x_F$ ,

$$P( M_n \leq x ) = F^n ( x ) \rightarrow 0, \quad n \rightarrow \infty,$$

e, nel caso in cui  $x_F < \infty$ , si ha, per  $x \geq x_F$ , che

$$P( M_n \leq x ) = F^n ( x ) = 1.$$

Di conseguenza,

$$M_n \xrightarrow{P} x_F, \text{ per } n \rightarrow \infty,$$

---

<sup>33</sup> EMBRECHTS P., KLÜPPELBERG C., MIKOSCH T., *Modelling extremal events...*, op. cit., pagg. 113 e sgg.

dove  $x_F \leq \infty$ . Quindi, la sequenza  $(M_n)$  è non decrescente in  $n$  e converge quasi certamente e dunque si può concludere che

$$M_n \xrightarrow{\text{q.c.}} x_F, \quad n \rightarrow \infty.$$

Questo risultato, però, non fornisce molte informazioni, in quanto la distribuzione limite che si ricava per la variabile aleatoria  $M_n$  è degenere; più utile, allo scopo di definire l'ordine di grandezza dai massimi, è, invece, il risultato della convergenza debole per massimi centrati e normalizzati. Vi sono, infatti, casi in cui, applicando una trasformazione lineare alla v.a.  $M_n$ , si può ottenere una nuova v.a.

$$c_n^{-1}(M_n - d_n),$$

dove  $c_n > 0$  e  $d_n \in \mathfrak{R}$  sono due costanti che dipendono da  $n$  ma non da  $x$ , a cui è associata una funzione di distribuzione  $H$  non degenere. Di conseguenza, è necessario considerare le probabilità nella forma

$$P(c_n^{-1}(M_n - d_n) \leq x),$$

che può anche essere riscritta come segue

$$P(M_n \leq u_n),$$

dove  $u_n = u_n(x) = c_n x + d_n$ . Tale probabilità viene dunque analizzata, per generali sequenze di  $(u_n)$ , per poter individuare quali condizioni garantiscano che il limite di  $P(M_n \leq u_n)$ , per  $n \rightarrow \infty$ , esista per appropriate costanti  $u_n$ . L'esistenza di una funzione di distribuzione limite (per  $n \rightarrow \infty$ ), ossia di una distribuzione del valore estremo  $H$  non degenere, è garantita da una coda  $\bar{F} = 1 - F$  che tende a zero molto lentamente, per  $x \rightarrow \infty$ , quindi dalla presenza di distribuzioni *heavy-tailed*. In contrasto con quanto succede per le somme, dunque, per i massimi si ha sempre bisogno di condizioni piuttosto delicate sulla coda  $\bar{F}$  per assicurare che  $P(M_n \leq u_n)$  converga ad un numero.

Prima di poter procedere, però, è necessario presentare l'*approssimazione di Poisson*, uno strumento che si rivelerà estremamente importante nel presente studio.

**Proposizione (2.4):** per dati  $\tau \in [0, \infty]$  ed una sequenza  $(u_n)$  di numeri reali, le seguenti relazioni risultano equivalenti:

$$n\bar{F}(u_n) \rightarrow \tau,$$

$$P( M_n \leq u_n ) \rightarrow \exp \{ -\tau \}.$$

Al fine di poter caratterizzare ed analizzare le possibili leggi limite per i massimi  $M_n$ , propriamente centrati e normalizzati, in particolare, si cercherà di capire quali distribuzioni soddisfano, per ogni  $n \geq 2$ , l'identità in legge

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

per appropriate costanti  $c_n > 0$  e  $d_n \in \mathfrak{R}$ . La precedente identità in legge rimanda alle proprietà definitorie delle distribuzioni stabili; esse rappresentano le sole possibili leggi limite per somme di variabili casuali indipendenti ed identicamente distribuite, centrate e normalizzate. Una nozione simile esiste anche per i massimi.

**Definizione (2.4):** *una variabile casuale non degenera  $X$  (o la corrispondente funzione di distribuzione) è detta max-stabile se*

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

*è soddisfatta per  $X, X_1, \dots, X_n$  indipendenti ed identicamente distribuite, per appropriate costanti  $c_n > 0$  e  $d_n \in \mathfrak{R}$  e per ogni  $n \geq 2$ .*

Si assuma ora che  $(X_n)$  sia una sequenza di variabili casuali indipendenti ed identicamente distribuite max-stabili. Dunque la precedente identità in legge può anche essere riscritta come segue:

$$c_n^{-1}( M_n - d_n ) \stackrel{d}{=} X.$$

Si può dunque concludere che ciascuna distribuzione max-stabile è una distribuzione limite per i massimi di variabili casuali indipendenti ed identicamente distribuite. In particolare, la classe delle distribuzioni in questione coincide con quella di tutte le possibili (non degeneri) leggi limite per massimi, propriamente normalizzati, di variabili indipendenti ed identicamente distribuite.

Viene ora presentato il teorema di Fisher-Tippett, riguardante le leggi limite per i massimi, risultato base per la classica teoria del valore estremo.

**Teorema di Fisher-Tippett (2.21):** *sia  $(X_n)$  una sequenza di variabili casuali iid. Se esistono costanti  $c_n > 0$  e  $d_n \in \mathfrak{R}$  e una funzione di distribuzione non degenera  $H$ , tale che*

$$c_n^{-1}(M_n - d_n) \xrightarrow{d} H,$$

allora  $H$  appartiene ad una delle tre seguenti funzioni di distribuzione:

$$\text{Fréchet: } \Phi_\alpha(x) = \begin{cases} 0, & x \leq 0 \\ \exp\{-x^{-\alpha}\}, & x > 0 \end{cases} \quad \alpha > 0$$

$$\text{Weibull: } \Psi_\alpha(x) = \begin{cases} \exp\{-(-x)^\alpha\}, & x \leq 0 \\ 1, & x > 0 \end{cases} \quad \alpha > 0$$

$$\text{Gumbel: } \Lambda(x) = \exp\{-e^{-x}\}, \quad x \in \mathfrak{R}.$$

Nonostante ai fini della modellizzazione, le distribuzioni  $\Phi_\alpha, \Psi_\alpha$  e  $\Lambda$  siano molto diverse tra loro, da un punto di vista matematico, invece, sono strettamente collegate. Infatti, si possono immediatamente verificare le seguenti proprietà: si supponga che sia  $X > 0$ , allora

$$X \text{ ha funzione di distribuzione } \Phi_\alpha \Leftrightarrow \ln X^\alpha \text{ ha funzione di distribuzione } \Lambda \Leftrightarrow -X^{-1} \text{ ha funzione di distribuzione } \Psi_\alpha.$$

Le funzioni di distribuzione  $\Phi_\alpha, \Psi_\alpha$  e  $\Lambda$ , presentate nel precedente teorema, sono le cosiddette *distribuzioni standard del valore estremo* e le v.a. associate vengono dette *estreme*. Le prime sono esattamente le distribuzioni max-stabili. Dunque, se  $X$  è una variabile casuale estrema, la seguente relazione è soddisfatta:

$$c_n^{-1}(M_n - d_n) \stackrel{d}{=} X.$$

In particolare, ai tre casi sopra presentati corrispondono le seguenti relazioni:

$\text{Fréchet: } M_n \stackrel{d}{=} n^{1/\alpha} X$ $\text{Weibull: } M_n \stackrel{d}{=} n^{-1/\alpha} X$ $\text{Gumbel: } M_n \stackrel{d}{=} X + \ln n.$
---

### 2.3.1 Maximum domains of attraction e costanti di normalizzazione e centramento

Nella sezione precedente, le distribuzioni del valore estremo sono state identificate come leggi limite per massimi centrati e normalizzati di variabili casuali indipendenti ed identicamente distribuite. Ora, invece, data una distribuzione del valore estremo  $H$ , si analizza quali condizioni, sulla funzione di distribuzione  $F$ , implicano che i massimi  $M_n$  convergono debolmente ad  $H$ ; in particolare si esamina quali valori delle costanti  $c_n > 0$  e  $d_n \in \mathfrak{R}$  bisogna scegliere affinché il teorema di Fisher-Tippett risulti verificato. Per poter fare quanto appena detto, però, risulta necessario introdurre alcuni concetti preliminari.

**Definizione (2.5):** *una variabile casuale  $X$  (o la sua funzione di distribuzione  $F$ ) appartiene al maximum domain of attraction della funzione di distribuzione del valore estremo  $H$ , se esistono delle costanti  $c_n > 0$  e  $d_n \in \mathfrak{R}$  tali per cui*

$$c_n^{-1} (M_n - d_n) \xrightarrow{d} H.$$

*Nel presente caso si può scrivere  $X \in MDA(H)$  o  $F \in MDA(H)$ .*

La seguente proposizione, inoltre, permette di caratterizzare il *maximum domain of attraction*.

**Proposizione (2.5):** *la funzione di distribuzione  $F$  appartiene al maximum domain of attraction della distribuzione del valore estremo  $H$ , con costanti  $c_n > 0$  e  $d_n \in \mathfrak{R}$ , se e solo se*

$$\lim_{n \rightarrow \infty} n \bar{F}(c_n x + d_n) = -\ln H(x), \quad x \in \mathfrak{R}.$$

*Per  $H(x) = 0$  il limite è interpretato come  $\infty$ .*

Nelle pagine seguenti viene fornita la definizione del *maximum domain of attraction* di ciascuna distribuzione del valore estremo. Tramite l'utilizzo del concetto di variazione regolare<sup>34</sup>, i casi di Fréchet e di Weibull sono affrontati in modo piuttosto semplice. Per quanto riguarda, invece, la distribuzione di Gumbel,

---

<sup>34</sup> La definizione di funzione *regularly varying* (2.1) è presente a pag. 43.

il *maximum domain of attraction* non è facilmente caratterizzabile, in quanto consiste di funzioni di distribuzione le cui code destre decrescono a zero più velocemente di ogni funzione potenza. Prima di poter procedere, però, è necessario dare alcune definizioni, tra cui quella di equivalenza di code:

**Definizione (2.6):** due funzioni di distribuzione  $F$  e  $G$  sono dette coda-equivalenti se hanno lo stesso estremo superiore (quindi  $x_F = x_G$ ) e se

$$\lim_{x \uparrow x_F} \frac{\bar{F}(x)}{\bar{G}(x)} = c,$$

per alcune costanti  $0 < c < \infty$ .

Ciascun *maximum domain of attraction* è strettamente correlato al concetto di equivalenza di code; infatti, per ogni  $F$  e  $G$  coda-equivalenti,  $F \in \text{MDA}(H)$  se e solo se  $G \in \text{MDA}(H)$ . Inoltre, per due distribuzioni coda-equivalenti, si possono utilizzare le medesime costanti di normalizzazione e di centramento; questo sarà di grande aiuto nel momento in cui si andranno a calcolare le costanti stesse.

In precedenza, si sono identificate le distribuzioni max-stabili come leggi limite per massimi di variabili indipendenti ed identicamente distribuite. Analogamente, le distribuzioni stabili sono identificabili come leggi limite per somme centrate e normalizzate. Queste vengono centrate tramite la loro mediana o le loro medie troncate. Il massimo campionario  $M_n$ , invece, è la versione empirica del  $(1 - n^{-1})$ -quantile della funzione di distribuzione  $F$ . Di conseguenza, quest'ultimo rappresenta un'appropriata costante di centramento. I quantili corrispondono all'inverso di una funzione di distribuzione. Nella seguente definizione della funzione quantile, ci si concentra su una versione continua a sinistra; prima, però viene definita l'inversa generalizzata di una funzione monotona:

**Definizione (2.7):** si supponga che  $h$  sia una funzione non decrescente in  $\mathfrak{R}$ .

L'inversa generalizzata di  $h$  è definita come segue:

$$h^-(t) = \inf \{ x \in \mathfrak{R} : h(x) \geq t \}.$$

**Definizione (2.8):** l'inversa generalizzata della funzione di distribuzione  $F$

$$F^-(t) = \inf \{ x \in \mathfrak{R} : F(x) \geq t \}, \quad 0 < t < 1,$$

è detta funzione quantile della funzione di distribuzione  $F$ . La quantità  $x_t = F^{-1}(t)$  definisce il  $t$ -quantile di  $F$ .

▫ **Il maximum domain of attraction della distribuzione di Fréchet**

Sia data la funzione di distribuzione di Fréchet  $\Phi_\alpha(x) = \exp\{-x^{-\alpha}\}$ , per  $\alpha > 0$ . Dallo sviluppo in serie di Taylor, è possibile scrivere che

$$1 - \Phi_\alpha(x) = 1 - \exp\{-x^{-\alpha}\} \sim x^{-\alpha}, \quad x \rightarrow \infty,$$

dunque, la coda di  $\Phi_\alpha$  decresce come una legge di potenza.

Ora, nelle pagine che seguono, si andrà ad analizzare quanto ci si possa allontanare da una coda che decresce come una legge di potenza, pur rimanendo sempre nel  $\text{MDA}(\Phi_\alpha)$ . Si mostrerà che il *maximum domain of attraction* di  $\Phi_\alpha$  consiste di funzioni di distribuzione  $F$  le cui code destre variano regolarmente con indice  $-\alpha$ . Per  $F \in \text{MDA}(\Phi_\alpha)$ , le costanti  $d_n$  possono essere scelte uguali a zero e le  $c_n$  tramite la funzione quantile, più precisamente

$$\begin{aligned} c_n &= F^{-1}(1 - n^{-1}) = \inf\{x \in \mathfrak{R} : F(x) \geq 1 - n^{-1}\} = \\ &= \inf\{x \in \mathfrak{R} : (1/\bar{F})(x) \geq n\} = (1/\bar{F})^{-1}(n). \end{aligned}$$

Tramite il seguente teorema, viene definito il  $\text{MDA}(\Phi_\alpha)$ .

**Teorema (2.22):** *la funzione di distribuzione  $F$  appartiene al maximum domain of attraction di  $\Phi_\alpha$ , con  $\alpha > 0$ , se e solo se  $\bar{F}(x) = x^{-\alpha}L(x)$  per alcune funzioni slowly varying  $L$ . Se  $F \in \text{MDA}(\Phi_\alpha)$ , allora*

$$c_n^{-1} M_n \xrightarrow{d} \Phi_\alpha,$$

*dove le costanti  $c_n$  possono essere scelte tramite la funzione quantile, come precedentemente mostrato.*

Si noti che tale risultato implica, in particolare, che ogni  $F \in \text{MDA}(\Phi_\alpha)$  ha un estremo superiore infinito  $x_F = \infty$ . Inoltre, le costanti  $c_n$  formano una sequenza che

varia regolarmente, più precisamente  $c_n = n^{1/\alpha} L_1(n)$ , per alcune funzioni *slowly varying*  $L_1$ .

A questo punto, per quanto riguarda il  $MDA(\Phi_\alpha)$  è possibile affermare che:

$$F \in MDA(\Phi_\alpha) \Leftrightarrow \bar{F} \in R_{-\alpha}.$$

Si è, dunque, così ottenuta una semplice caratterizzazione del  $MDA(\Phi_\alpha)$ . Si noti che questa classe di funzioni di distribuzione contiene distribuzioni dalle code molto pesanti. Quindi queste potrebbero essere appropriate per “modellizzare”, ad esempio, grandi sinistri assicurativi e grandi fluttuazioni di prezzi.

Viene ora presentata la condizione di Von Mises, che permette di specificare ulteriormente il risultato appena ottenuto.

**Corollario (condizione di Von Mises) (2.2):** *sia  $F$  una funzione di distribuzione assolutamente continua con funzione di densità  $f$  e che soddisfi il seguente limite*

$$\lim_{x \rightarrow \infty} \frac{x f(x)}{\bar{F}(x)} = \alpha > 0;$$

*allora  $F \in R_{-\alpha}$  e di conseguenza  $F \in MDA(\Phi_\alpha)$ .*

Per concludere, la seguente proposizione permette di approfondire ulteriormente la conoscenza della struttura del  $MDA(\Phi_\alpha)$ . Essa, oltre ad essere utile sotto l'aspetto teorico, risulta essere un importante strumento per il calcolo delle costanti.

**Proposizione (2.6):** *siano  $F$  e  $G$  funzioni di distribuzione e si supponga che  $F \in MDA(\Phi_\alpha)$  con costanti  $c_n > 0$ , quindi*

$$\lim_{n \rightarrow \infty} F^n(c_n x) = \Phi_\alpha(x), \quad x > 0.$$

*Allora*

$$\lim_{n \rightarrow \infty} G^n(c_n x) = \Phi_\alpha(cx), \quad x > 0,$$

*per alcune  $c > 0$  se e solo se  $F$  e  $G$  sono coda-equivalenti con*

$$\lim_{n \rightarrow \infty} \frac{\bar{F}(x)}{\bar{G}(x)} = c^\alpha.$$

Concludendo, in precedenza si è visto che  $F \in \text{MDA}(\Phi_\alpha)$  se e solo se  $\bar{F} \in R_{-\alpha}$ . Inoltre ogni  $F \in \text{MDA}(\Phi_\alpha)$  è coda-equivalente ad una funzione di distribuzione assolutamente continua che soddisfi la condizione di Von Mises. Il tutto può essere così riassunto:

$\text{MDA}(\Phi_\alpha)$  è costituito da funzioni di distribuzione che soddisfano la condizione di Von Mises e dalle funzioni di distribuzione ad esse coda-equivalenti.

La seguente tabella riassume le caratteristiche principali del *maximum domain of attraction* della distribuzione di Fréchet  $\Phi_\alpha$  e ne fornisce i relativi esempi.

Fréchet	$\Phi_\alpha = \exp \{-x^{-\alpha}\}, \quad x > 0, \alpha > 0.$
$\text{MDA}(\Phi_\alpha)$	$x_F = \infty, \quad \bar{F}(x) = x^{-\alpha} L(x), \quad L \in R_0.$
Costanti	$c_n = F^{-}(1 - n^{-1}) = n^{1/\alpha} L_1(n), \quad L_1 \in R_0, \quad d_n = 0.$
Risultato del limite	$c_n^{-1} M_n \xrightarrow{d} \Phi_\alpha$
ESEMPI	
Pareto Burr Stabile con indice $\alpha < 2$	$\bar{F}(x) \sim Kx^{-\alpha}, \quad K, \alpha > 0.$ $c_n = (Kn)^{1/\alpha}.$
Loggamma	$f(x) = \frac{\alpha^\beta}{\Gamma(\beta)} (\ln x)^{\beta-1} x^{-\alpha-1}, \quad x > 1, \quad \alpha, \beta > 0.$ $c_n = ((\Gamma(\beta))^{-1} (\ln n)^{\beta-1} n)^{1/\alpha}$

Tabella 6- *Maximum domain of attraction* della distribuzione di Fréchet.

▫ **Il maximum domain of attraction della distribuzione di Weibull**

Analogamente a quanto fatto nella sezione precedente per la distribuzione di Fréchet, viene ora fornita una caratterizzazione del *maximum domain of attraction* di  $\Psi_\alpha$ , per  $\alpha > 0$ . Diversamente dal precedente caso tutte le funzioni di distribuzione  $F$  in  $MDA(\Psi_\alpha)$  hanno un estremo superiore  $x_F$  finito. D'altra parte, però, come già accennato in precedenza, le distribuzioni di Fréchet e Weibull sono strettamente correlate, infatti

$$\Psi_\alpha(-x^{-1}) = \Phi_\alpha(x), \quad x > 0.$$

Dunque ci si può aspettare che anche  $MDA(\Phi_\alpha)$  e  $MDA(\Psi_\alpha)$  siano tra loro collegati. Il seguente teorema, che caratterizza il *maximum domain of attraction* di  $\Psi_\alpha$ , lo conferma.

**Teorema (2.23):** *la funzione di distribuzione  $F$  appartiene al maximum domain of attraction di  $\Psi_\alpha$  se e solo se  $x_F < \infty$  e  $\bar{F}(x_F - x^{-1}) = x^{-\alpha}L(x)$  per alcune funzioni slowly varying  $L$ .*

Se  $F \in MDA(\Psi_\alpha)$ , allora

$$c_n^{-1}(M_n - x_F) \xrightarrow{d} \Psi_\alpha,$$

dove le costanti  $c_n$  possono essere scelte come segue

$$c_n = x_F - F^{-1}(1 - n^{-1})$$

e  $d_n = x_F$ .

Dal presente teorema, segue che:

$$F \in MDA(\Psi_\alpha) \Leftrightarrow x_F < \infty \text{ e } \bar{F}(x_F - x^{-1}) \in R_{-\alpha}.$$

Di conseguenza,  $MDA(\Psi_\alpha)$  è costituito da funzioni di distribuzione  $F$  con supporto limitato a destra, essendo  $x_F < \infty$ . Per tale ragione, queste non rappresentano la migliore scelta per “modellizzare” eventi estremi in campo assicurativo o finanziario. Spesso, quindi, funzioni di distribuzione con estremo superiore infinito sono preferibili, dal momento in cui permettono di avere valori arbitrariamente

ampi in un campione. Questo genere di distribuzioni appartiene a  $MDA(\Phi_\alpha)$  o a  $MDA(\Lambda)$ .

Nella sezione precedente, si è trovato conveniente caratterizzare l'appartenenza al  $MDA(\Phi_\alpha)$  tramite la densità della funzione di ripartizione; tale procedura (nota sotto il nome di *condizione di Von Mises*) può essere applicata anche nel caso in esame.

**Corollario (condizione di Von Mises) (2.3):** *sia  $F$  una funzione di distribuzione assolutamente continua, con densità  $f$ , positiva per alcuni intervalli finiti  $(z, x_F)$ . Se*

$$\lim_{x \uparrow x_F} \frac{(x_F - x) f(x)}{\bar{F}(x)} = \alpha > 0,$$

*allora  $F \in MDA(\Psi_\alpha)$ .*

Per concludere, parallelamente a quanto fatto per la distribuzione di Fréchet, vengono ora enunciate le proprietà del *maximum domain of attraction* di  $\Psi_\alpha$ :

**Proposizione (2.7):** *siano  $F$  e  $G$  funzioni di distribuzione con estremi superiori  $x_F = x_G < \infty$  e si supponga che  $F \in MDA(\Psi_\alpha)$  con costanti  $c_n > 0$ ; quindi*

$$\lim_{n \rightarrow \infty} F^n(c_n x + x_F) = \Psi_\alpha(x), \quad x < 0.$$

*Allora*

$$\lim_{n \rightarrow \infty} G^n(c_n x + x_F) = \Psi_\alpha(cx), \quad x < 0,$$

*per alcune  $c > 0$  se e solo se  $F$  e  $G$  sono coda-equivalenti con*

$$\lim_{x \uparrow x_F} \frac{\bar{F}(x)}{\bar{G}(x)} = c^{-\alpha}.$$

Infine, come già accennato per il precedente caso, ogni  $F \in MDA(\Psi_\alpha)$  è coda-equivalente ad una funzione di distribuzione assolutamente continua, che soddisfi la condizione di Von Mises. Il tutto può essere così riassunto:

MDA( $\Psi_\alpha$ ) è costituito da funzioni di distribuzione che soddisfano la condizione di Von Mises e dalle loro funzioni di distribuzione ad esse coda-equivalenti.

Weibull	$\Psi_\alpha(x) = \exp\{-(-x)^\alpha\}, \quad x < 0, \alpha > 0.$
MDA( $\Psi_\alpha$ )	$x_F < \infty, \quad \bar{F}(x_F - x^{-1}) = x^{-\alpha}L(x), \quad L \in R_0.$
Costanti	$c_n = x_F - F^{-1}(1 - n^{-1}) = n^{-1/\alpha}L_1(n), \quad L_1 \in R_0, \quad d_n = x_F.$
Risultato del limite	$c_n^{-1} (M_n - x_F) \xrightarrow{d} \Psi_\alpha$
<b>ESEMPI</b>	
Uniforme	$f(x) = 1, \quad x \in (0,1).c_n = n^{-1}, \quad d_n = 1.$
Legge di potenza di $x_F$	$\bar{F}(x) = K(x_F - x)^\alpha, \quad x_F - K^{-1/\alpha} \leq x \leq x_F, \quad K, \alpha > 0.c_n = (Kn)^{-1/\alpha}, \quad d_n = x_F.$

Tabella 7- *Maximum domain of attraction* della distribuzione di Weibull.

Nella precedente tabella vengono riassunte le principali caratteristiche del *maximum domain of attraction* della distribuzione di Weibull e i due esempi della distribuzione uniforme e delle legge di potenza di  $x_F$ .

▫ **Il *maximum domain of attraction* della distribuzione di Gumbel**

Il *maximum domain of attraction* della distribuzione di Gumble  $\Lambda(x) = \exp\{-\exp\{-x\}\}$  copre un'ampia gamma di funzioni di distribuzione  $F$ . Nonostante non ci sia un collegamento diretto con la teoria della variazione regolare, come nei casi di Fréchet e Weibull, è possibile trovare estensioni di tale concetto che consentano una completa caratterizzazione del MDA( $\Lambda$ ).

Innanzitutto, dalla espansione in serie di Taylor, è possibile affermare che

$$1 - \Lambda \sim e^{-x}, \quad x \rightarrow \infty,$$

quindi  $\bar{\Lambda}(x)$  decresce esponenzialmente a zero.

Si può osservare che il  $MDA(\Lambda)$  contiene funzioni di distribuzione con code molto diverse tra loro, da quelle pesanti (come, ad esempio, il caso della Lognormale) a quelle leggere (come la Normale). Nel presente caso, dunque,  $x_F$  può assumere valori sia finiti che infiniti.

Prima, però, di procedere con lo studio del caso in esame, è necessario considerare alcune funzioni assolutamente continue  $F \in MDA(\Lambda)$ , che hanno una semplice rappresentazione proposta da Von Mises. Queste costituiscono una parte fondamentale del *maximum domain of attraction* in questione; infatti, dopo averle analizzate, sarà possibile considerarne una leggera modifica per poter caratterizzare in modo completo  $MDA(\Lambda)$ .

**Definizione (funzione di Von Mises) (2.9):** sia  $F$  una funzione di distribuzione con estremo superiore  $x_F \leq \infty$ , tale per cui  $F$  abbia rappresentazione

$$\bar{F}(x) = c \exp \left\{ - \int_z^x \frac{1}{a(t)} dt \right\}, \quad z < x < x_F, \quad (II.14)$$

dove  $c$  è una costante positiva,  $a(\cdot)$  una funzione positiva ed assolutamente continua con densità  $a'$  e  $\lim_{x \rightarrow x_F} a'(x) = 0$ . Allora  $F$  è detta funzione di Von

Mises e  $a(\cdot)$  funzione ausiliaria di  $F$ .

Esistono molte funzioni che possono essere ricondotte alla tipologia di Von Mises; tra queste, la più importante è quella esponenziale.

La (II.14) potrebbe essere paragonata alla rappresentazione di una funzione *regularly varying* con indice  $-\alpha$ . La funzione ausiliaria di una funzione di Von Mises, con  $x_F = \infty$ , soddisfa  $a(x) / x \rightarrow 0$ . Segue immediatamente che  $\bar{F}(x)$  decresce a zero molto più velocemente di qualunque legge di potenza  $x^{-\alpha}$ . Dunque, la variazione regolare sembra non essere lo strumento adatto per descrivere le funzioni di Von Mises. Ai fini della nostra analisi, però ci viene incontro il concetto di variazione rapida, alla quale si possono estendere alcuni importanti risultati ottenuti per le funzioni *regularly varying*. Una funzione  $\bar{F}$  si definisce *rapidly varying* (ossia  $\bar{F} \in R_{-\infty}$ ) se

$$\lim_{x \rightarrow \infty} \frac{\bar{F}(xt)}{\bar{F}(x)} = \begin{cases} 0 & \text{se } t > 1, \\ \infty & \text{se } 0 < t < 1 \end{cases}$$

Si possono, a questo punto, definire le proprietà della funzione di Von Mises.

**Proposizione (2.8):** *ciascuna funzione di Von Mises  $F$  è assolutamente continua in  $(z, x_F)$  con densità positiva  $f$ . La funzione ausiliaria può essere scelta come  $a(x) = \bar{F}(x)/f(x)$ . Inoltre, valgono le seguenti proprietà:*

a) se  $x_F = \infty$ , allora  $\bar{F} \in R_{-\infty}$  e

$$\lim_{x \rightarrow \infty} \frac{x f(x)}{\bar{F}(x)} = \infty.$$

b) se  $x_F < \infty$ , allora  $\bar{F}(x_F - x^{-1}) \in R_{-\infty}$  e

$$\lim_{x \uparrow x_F} \frac{(x_F - x) f(x)}{\bar{F}(x)} = \infty.$$

Risulta, quindi, possibile mostrare che le funzioni di Von Mises appartengono al *maximum domain of attraction* della distribuzione di Gumble. Inoltre, la forma specifica di  $\bar{F}$  permette di calcolare le costanti  $c_n$ , tramite la funzione ausiliaria. Quanto detto ora è il contenuto della seguente proposizione:

**Proposizione (2.9):** *si supponga che la funzione di distribuzione  $F$  sia una funzione di Von Mises. Allora  $F \in MDA(\Lambda)$ . Una possibile scelta delle costanti è la seguente:*

$$c_n = a(d_n) \quad \text{e} \quad d_n = F^{-1}(1 - n^{-1}),$$

dove  $a$  è la funzione ausiliaria di  $F$ .

Una volta esaminate le funzioni di Von Mises, è ora possibile tornare a considerare il quesito centrale della presente sezione: quanto ci si può allontanare da una coda esponenziale, pur rimanendo nel  $MDA(\Lambda)$ ?

Come spiegato in precedenza, le funzioni di Von Mises non definiscono completamente il  $MDA(\Lambda)$ , ma una lieve modifica della loro relazione definitoria

(II.14) permette una caratterizzazione completa di  $\text{MDA}(\Lambda)$ . Viene quindi enunciato un teorema contenente una prima caratterizzazione del  $\text{MDA}(\Lambda)$ .

**Teorema (2.24):** *la funzione di distribuzione  $F$  con estremo superiore  $x_F \leq \infty$ , appartiene al maximum domain of attraction di  $\Lambda$  se e solo se esistono alcune  $z < x_F$  tali che  $F$  abbia rappresentazione*

$$\bar{F}(x) = c(x) \exp \left\{ - \int_z^x \frac{g(t)}{a(t)} dt \right\}, \quad z < x < x_F,$$

dove  $c$  e  $g$  sono funzioni misurabili che soddisfano le seguenti relazioni:

$$c(x) \rightarrow c > 0 \quad \text{e} \quad g(x) \rightarrow 1 \quad \text{q.c.}, \quad \text{per } x \uparrow x_F,$$

ed  $a$  è una funzione positiva assolutamente continua con densità  $a'(x)$  con

$$\lim_{x \rightarrow x_F} a'(x) = 0.$$

Per  $F$ , con una rappresentazione come quella appena definita, si possono scegliere

$$c_n = a(d_n) \quad \text{e} \quad d_n = F^{-1}(1 - n^{-1}),$$

come costanti. Una scelta possibile della funzione  $a$  è la seguente:

$$a(x) = \int_x^{x_F} \frac{\bar{F}(t)}{\bar{F}(x)} dt, \quad x < x_F.$$

La rappresentazione di  $\bar{F}$  del precedente teorema non è unica, dal momento in cui ci possono essere degli interscambi tra le funzioni  $c$  e  $g$ .

Per una variabile casuale  $X$ , la funzione  $a(x)$ , come definita nel teorema appena enunciato, non è nient'altro che la *funzione dell'eccesso dalla media* (o *mean excess*)

$$a(x) = E(X - x \mid X > x), \quad x < x_F.$$

Esiste, inoltre, un altro teorema contenente una seconda caratterizzazione del  $\text{MDA}(\Lambda)$ .

**Teorema (2.25):** *la funzione di distribuzione  $F$  appartiene al maximum domain of attraction di  $\Lambda$  se e solo se esiste una qualche funzione positiva  $\tilde{a}$  tale che valga la seguente relazione:*

$$\lim_{x \uparrow x_F} \frac{\bar{F}(x + t\tilde{a}(x))}{\bar{F}(x)} = e^{-t}, \quad t \in \mathfrak{R}.$$

Una possibile scelta di  $\tilde{a} = a$  è data dal teorema precedente.

Si richiami ora il concetto di equivalenza di coda<sup>35</sup>. Analogamente ai casi delle distribuzioni di Fréchet e di Weibull, la coda-equivalenza è uno strumento utile per decidere se una particolare funzione appartiene al  $MDA(\Lambda)$  e per calcolare le costanti di normalizzazione e di centramento. Nel caso in esame, inoltre, tale concetto risulta anche più importante a causa della grande varietà di code  $\bar{F}$ . Viene quindi enunciata una importante proposizione contenente le proprietà del  $MDA(\Lambda)$ , sotto la condizione di coda-equivalenza:

**Proposizione (2.10):** siano  $F$  e  $G$  funzioni di distribuzione con gli stessi estremi superiori,  $x_F = x_G$ , e si assuma che  $F \in MDA(\Lambda)$  con costanti  $c_n > 0$  e  $d_n \in \mathfrak{R}$ ; allora

$$\lim_{n \rightarrow \infty} F^n(c_n x + d_n) = \Lambda(x), \quad x \in \mathfrak{R}.$$

$$\lim_{n \rightarrow \infty} G^n(c_n x + d_n) = \Lambda(x + b) \quad x \in \mathfrak{R}.$$

se e solo se  $F$  e  $G$  sono coda-equivalenti con

$$\lim_{x \uparrow x_F} \frac{\bar{F}(x)}{\bar{G}(x)} = e^b.$$

In conclusione, il risultato della presente sezione porta ad una completa caratterizzazione di  $MDA(\Lambda)$ , che può così essere riassunto:

$MDA(\Lambda)$ , è costituito da funzioni di distribuzione che soddisfano la condizione di Von Mises e dalle funzioni di distribuzione ad esse coda-equivalenti.

Come già affermato in precedenza,  $MDA(\Lambda)$  è costituito da un'ampia gamma di distribuzioni, le cui code possono essere estremamente diverse tra loro. Per

<sup>35</sup> Il concetto di funzioni coda-equivalenti è fornito dalla definizione (2.6), pag. 71.

questa ragione,  $MDA(\Lambda)$  è forse il più interessante tra tutti i *maximum domains of attraction*. Come ovvia conseguenza dell'abbondante assortimento di code, per il caso in esame, anche le costanti possono essere altrettanto varie. Inoltre, mentre per i casi di Fréchet e Weibull le costanti erano calcolate applicando direttamente la teoria della variazione regolare, per  $MDA(\Lambda)$ , invece, sono necessari risultati molto più avanzati.

Nella tabella 8 vengono riassunte le principali caratteristiche del *maximum domain of attraction* della distribuzione di Gumbel ed alcuni dei possibili esempi per il caso in esame.

Gumbel	$\Lambda(x) = \exp\{-e^{-x}\}, \quad x \in \mathfrak{R}$
$MDA(\Lambda)$	$x_F \leq \infty, \bar{F}(x) = c(x) \exp\left\{-\int_{x_0}^x \frac{g(t)}{a(t)} dt\right\}, \quad x_0 < x < x_F,$ dove $c(x) \rightarrow c > 0, g(x) \rightarrow 1, a'(x) \rightarrow 0$ per $x \rightarrow x_F$
Costanti	$c_n = a(d_n), \quad d_n = F^{-1}(1 - n^{-1})$
Risultato del limite	$c_n^{-1} (M_n - d_n) \xrightarrow{d} \Lambda$
<b>ESEMPI</b>	
Weibull	$\bar{F}(x) \sim Kx^\alpha \exp\{-cx^\tau\}, \quad K, c, \tau > 0, \alpha \in \mathfrak{R}.$ $c_n = (c\tau)^{-1} (c^{-1} \ln n)^{1/\tau - 1}$ $d_n = (c^{-1} \ln n)^{1/\tau} + \frac{1}{\tau} (c^{-1} \ln n)^{1/\tau - 1} \left\{ \frac{\alpha}{c\tau} \ln(c^{-1} \ln n) + \frac{\ln K}{c} \right\}$
Lognormale	$f(x) = \frac{1}{\sqrt{2\pi\sigma x}} e^{-(\ln x - \mu)^2 / 2\sigma^2}, \quad x > 0, \lambda \in \mathfrak{R}, \sigma > 0.$ $c_n = \sigma (2 \ln n)^{-1/2} d_n$ $d_n = \exp\left\{ \lambda + \sigma \left[ \sqrt{2 \ln n} - \frac{\ln(4\pi) + \ln \ln n}{2(2 \ln n)^{1/2}} \right] \right\}$
Gamma	$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \quad x > 0, \alpha, \beta > 0.$ $c_n = \beta^{-1}$ $d_n = \beta^{-1} (\ln n + (\alpha - 1) \ln \ln n - \ln \Gamma(\alpha))$
Esponenziale	$\bar{F}(x) \sim Ke^{-\lambda x}, \quad K, \lambda > 0.$ $c_n = \lambda^{-1}$ $d_n = \lambda^{-1} \ln(Kn)$

Tabella 8- Il *maximum domain of attraction* della distribuzione di Gumbel.

### 2.3.2 La distribuzione generalizzata del valore estremo e la distribuzione generalizzata di Pareto

Precedentemente, si è mostrato che la distribuzione standard del valore estremo, insieme ai suoi tre modelli, fornisce la sola legge limite non degenerare per trasformazioni di massimi di variabili casuali indipendenti ed identicamente distribuite. Nel presente paragrafo, i risultati appena ottenuti vengono presentati in modo compatto tramite la famiglia parametrica  $(H_\xi)_{\xi \in \mathbb{R}}$  di funzioni di distribuzione, contenente le distribuzioni standard del valore estremo, così definita:

$$H_\xi = \begin{cases} \Phi_{1/\xi} & \text{se } \xi > 0 \\ \Lambda & \text{se } \xi = 0 \\ \Psi_{-1/\xi} & \text{se } \xi < 0 \end{cases}$$

$H_\xi$  costituisce la *distribuzione generalizzata del valore estremo* e risulta molto utile, in quanto permette la rappresentazione uniparametrica dei tre casi standard tramite un'unica famiglia di distribuzioni. Questi tre casi vengono in essa rappresentati tramite l'introduzione del parametro  $\xi$  tale che:

$\xi = \alpha^{-1} > 0$  corrisponde alla distribuzione di Fréchet,

$\xi = 0$  corrisponde alla distribuzione di Gumbel,

$\xi = -\alpha^{-1} < 0$  corrisponde alla distribuzione di Weibull.

Nella seguente definizione viene esposta la rappresentazione di Von Mises della distribuzione del valore estremo generalizzata.

**Definizione (2.10):** si definisca la funzione di distribuzione  $H_\xi$  come segue

$$H_\xi(x) = \begin{cases} \exp\{-(1 + \xi x)^{-1/\xi}\} & \text{se } \xi \neq 0, \\ \exp\{-\exp(-x)\} & \text{se } \xi = 0, \end{cases}$$

dove  $1 + \xi x > 0$ . Allora il supporto di  $H_\xi$  corrisponde a

$$x > -\xi^{-1} \quad \text{per } \xi > 0,$$

$$x < -\xi^{-1} \quad \text{per } \xi < 0,$$

$$x \in \mathbb{R} \quad \text{per } \xi = 0.$$

$H_\xi$  è definita distribuzione generalizzata standard del valore estremo.

Si può anche introdurre la famiglia  $H_{\xi\mu\psi}$  sostituendo  $x$  con  $(x - \mu)/\psi$  per  $\mu \in \mathfrak{R}$  e  $\psi > 0$  e aggiustando opportunamente il supporto.

$H_\xi$  può anche essere scritta nel seguente modo:

$$H_\xi = \exp \{ - (1 + \xi x)^{-1/\xi} \}, \quad 1 + \xi x > 0$$

come rappresentazione per tutte le  $\xi \in \mathfrak{R}$ .

La distribuzione in questione fornisce una utile rappresentazione unificata dei casi di Fréchet, Weibull e Gumbel. La sua introduzione è motivata principalmente da applicazioni di carattere statistico, come si avrà modo di vedere più avanti.

Viene ora fornito uno dei risultati base della teoria del valore estremo; il seguente teorema, infatti, in modo analitico, fornisce le essenziali informazioni raccolte nelle sezioni precedenti, riguardanti il *maximum domain of attraction* ed, inoltre, rappresenta la base di numerose tecniche statistiche, che si analizzeranno più avanti. Prima di enunciare il teorema in questione, però, si riprenda la definizione di funzione quantile<sup>36</sup>  $F^\leftarrow$  della funzione di distribuzione  $F$  e si definisca:

$$U(t) = F^\leftarrow (1 - t^{-1}), \quad t > 0.$$

**Teorema (2.26):** per  $\xi \in \mathfrak{R}$ , le seguenti affermazioni sono equivalenti:

- $F \in \text{MDA}(H_\xi)$ .
- Esiste una funzione positiva e misurabile  $a(\cdot)$  tale che, per  $1 + \xi x > 0$ ,

$$\lim_{u \rightarrow x_F} \frac{\bar{F}(u + xa(u))}{\bar{F}(u)} = \begin{cases} (1 + \xi x)^{-1/\xi} & \text{se } \xi \neq 0, \\ e^{-x} & \text{se } \xi = 0. \end{cases} \quad (\text{II.15})$$

- Per  $x, y > 0, y \neq 1$ ,

$$\lim_{s \rightarrow \infty} \frac{U(sx) - U(s)}{U(sy) - U(s)} = \begin{cases} \frac{x^\xi - 1}{y^\xi - 1} & \text{se } \xi \neq 0, \\ \frac{\ln x}{\ln y} & \text{se } \xi = 0. \end{cases}$$

<sup>36</sup> Una definizione (2.8) della funzione quantile è presente a pag. 71.

La condizione b. del precedente teorema ha una interessante interpretazione probabilistica; sia, infatti,  $X$  una variabile casuale con  $F \in \text{MDA}(H_\xi)$ . Allora la (II.15) può essere riformulata come segue:

$$\lim_{u \uparrow x_F} P\left(\frac{X-u}{a(u)} > x \mid X > u\right) = \begin{cases} (1+\xi x)^{-1/\xi} & \text{se } \xi \neq 0, \\ e^{-x} & \text{se } \xi = 0 \end{cases} \quad (\text{II.16})$$

Quest'ultima formula fornisce una approssimazione distribuzionale per gli eccessi oltre un (elevato) limite  $u$ . Un appropriato fattore di scala è rappresentato da  $a(u)$ . Avendo menzionato la nozione di "eccesso" si ritiene opportuno, a questo punto, dare una definizione di *funzione di distribuzione dell'eccesso* (o *excess distribution function*) e *funzione dell'eccesso medio* (o *mean excess function*).

**Definizione (2.11):** sia  $X$  una variabile casuale con funzione di distribuzione  $F$  ed estremo superiore  $x_F$ . Per una fissata  $u < x_F$ ,

$$F_u(x) = P(X - u \leq x \mid X > u), \quad x \geq 0,$$

rappresenta la funzione di distribuzione dell'eccesso della variabile casuale  $X$  (o della sua funzione di distribuzione  $F$ ) oltre il limite  $u$ . La funzione

$$e(u) = E(X - u \mid X > u),$$

è detta funzione dell'eccesso medio di  $X$ .

Distribuzioni	Mean excess functions
Pareto	$\frac{\kappa + u}{\alpha - 1}, \quad \alpha > 1$
Loggamma	$\frac{u}{\alpha - 1} (1 + o(1)), \quad \alpha > 1$
Lognormale	$\frac{\sigma^2 u}{\ln u - \mu} (1 + o(1)),$
Weibull	$\frac{u^{1-\tau}}{c\tau} (1 + o(1))$
Esponenziale	$\lambda^{-1}$
Gamma	$\beta^{-1} \left[ 1 + \frac{\alpha - 1}{\beta u} + o(1/u) \right]$

Tabella 9- La *mean excess function* per alcune distribuzioni standard. Le relazioni asintotiche sono intese per  $u \rightarrow \infty$ .

Utilizzando la definizione di  $e(u)$  e l'integrazione parziale, si possono verificare semplicemente le seguenti formule, che permettono di calcolare la funzione media dell'eccesso in casi speciali. Si supponga che  $X$  sia una variabile casuale con funzione di distribuzione  $F$  e media finita.

Allora

$$e(u) = \int_u^{x_F} (x-u)dF(x) / \bar{F}(u) = \frac{1}{\bar{F}(u)} \int_u^{x_F} \bar{F}(x)dx, \quad 0 < u < x_F.$$

Gli eccessi oltre una soglia giocano un ruolo fondamentale in molti campi. Ad esempio,  $F_u$  potrebbe rappresentare la funzione di distribuzione del tempo residuo di vita, o, in campo assicurativo, la funzione di distribuzione *excess of loss*. La distribuzione del valore estremo, in particolare nella sua versione fornita dalla (II.16), porta alla seguente definizione di *distribuzione generalizzata di Pareto*.

**Definizione (2.12):** si definisca la funzione di distribuzione  $G_\xi$  come segue

$$G_\xi(x) = \begin{cases} 1 - (1 + \xi x)^{-1/\xi} & \text{se } \xi \neq 0, \\ 1 - e^{-x} & \text{se } \xi = 0. \end{cases}$$

dove

$$\begin{aligned} x &\geq 0 && \text{se } \xi \geq 0, \\ 0 &\leq x \leq -1/\xi && \text{se } \xi < 0. \end{aligned}$$

$G_\xi$  è detta distribuzione generalizzata di Pareto. Si può anche introdurre la famiglia  $G_{\xi,\mu,\psi}$ , sostituendo  $(x - \nu)/\beta$  ad  $x$ , per  $\nu \in \mathfrak{R}$  e  $\beta > 0$  e aggiustando opportunamente il supporto.

La distribuzione appena definita rappresenta uno dei concetti fondamentali nella statistica degli eventi estremi. Quanto detto sinora può essere riassunto come segue:

La distribuzione del valore estremo

$$H_\xi, \quad \xi \in \mathfrak{R},$$

descrive la distribuzione limite per massimi normalizzati.

La distribuzione generalizzata di Pareto

$$G_{\xi,\beta}, \quad \xi \in \mathfrak{R}, \beta > 0,$$

appare come distribuzione limite di eccessi oltre una soglia elevata.

Nel seguente teorema vengono elencate alcune delle più importanti proprietà della distribuzione generalizzata di Pareto.

**Teorema (2.27):**

- Si supponga che  $X$  abbia distribuzione generalizzata di Poisson con parametri  $\xi$  e  $\beta$ . Allora  $E(X) < \infty$  se e solo se  $\xi < 1$ . In questo ultimo caso

$$E(1 + (\xi/\beta)X)^{-r} = \frac{1}{1 + \xi r}, \quad r > -1/\xi,$$

$$E[\ln(1 + (\xi/\beta)X)]^k = \xi^k k!, \quad k \in \mathbb{N},$$

$$E[X(\overline{G}_{\xi, \beta}(x))^r] = \frac{\beta}{(r+1-\xi)(r+1)}, \quad (r+1)/|\xi| > 0.$$

Se  $\xi < 1/r$ , con  $r \in \mathbb{N}$ , allora

$$E[X^r] = \frac{\beta^r}{\xi^{r+1}} \frac{\Gamma(\xi^{-1} - r)}{\Gamma(1 + \xi^{-1})} r!$$

- Per ogni  $\xi \in \mathbb{R}$ ,  $F \in MDA(H_\xi)$  se e solo se

$$\lim_{u \uparrow x_F} \sup_{0 < x < x_F - u} |F_u(x) - G_{\xi, \beta(u)}(x)| = 0$$

per qualche funzione positiva  $\beta$ .

- Si supponga che  $x_i \in D(\xi, \beta)$ , per  $i = 1, 2$ , allora

$$\frac{\overline{G}_{\xi, \beta}(x_1 + x_2)}{\overline{G}_{\xi, \beta}(x_1)} = \overline{G}_{\xi, \beta + \xi x_1}(x_2).$$

- Si assuma che  $N \sim Poi(\lambda)$  e sia indipendente dalla successione di v.a. iid  $(X_n)$ , avente una distribuzione generalizzata di Pareto con parametri  $\xi$  e  $\beta$ . Sia  $M_N = \max(X_1, \dots, X_N)$ . Allora

$$P(M_N \leq x) = \exp \left\{ -\lambda (1 + \xi(x/\beta))^{-1/\xi} \right\} = H_{\xi\mu\psi}(x),$$

Dove  $\mu = \beta\xi^{-1}(\lambda^\xi - 1)$  e  $\psi = \beta\lambda^\xi$ .

- Si supponga che  $X$  abbia distribuzione generalizzata di Pareto con parametri  $\xi < 1$  e  $\beta$ . Allora per  $u < x_F$ ,

$$e(u) = E(X - u | X > u) = \frac{\beta + \xi u}{1 - \xi}, \quad \beta + u\xi > 0.$$

Riepilogando, precedentemente<sup>37</sup>, si era appreso che il numero degli eccessi oltre un limite elevato è all'incirca una Poisson. Ora si può anche concludere che una approssimazione della funzione di distribuzione dell'eccesso  $F_u$  tramite la distribuzione generalizzata di Pareto è ammissibile. Inoltre si può anche mostrare che il numero degli eccessi e la loro intensità sono indipendenti. Infine, in un modello, dove il numero degli eccessi è esattamente una Poisson e la *excess distribution function* è esattamente una distribuzione generalizzata di Pareto, allora il massimo di tali eccessi segue una esatta distribuzione generalizzata del valore estremo.

Quanto appena evidenziato, dunque, può essere così riassunto:

- Il numero di eccedenze di un limite elevato segue un processo di Poisson.
- Gli eccessi oltre un limite elevato possono essere “modellizzati” tramite una distribuzione generalizzata di Pareto.
- Un appropriato valore della soglia elevata può essere trovato tracciando la funzione *mean excess* empirica.
- La distribuzione del massimo di un numero Poisson di eccessi, indipendenti ed identicamente distribuiti, oltre un limite elevato è la distribuzione generalizzata del valore estremo.

## **2.4 STATISTICHE DI ORDINE SUPERIORE E PROCESSI DI PUNTI**

Per poter effettuare un'analisi statistica degli eventi estremi, ed in particolare di quelli catastrofici, la teoria del valore estremo può essere utilizzata per mezzo di due importanti tecniche, che vengono esaminate nel presente paragrafo: quella delle statistiche di ordine superiore e quella dei processi di punti. Come si vedrà in seguito, con l'utilizzo di tali strumenti ed i risultati che essi

---

<sup>37</sup> A tale proposito si veda l'approssimazione di Poisson, definita nella proposizione (2.4), a pag. 67.

permettono di ottenere, si possono definire metodi appropriati, utili a fornire stime delle funzioni di distribuzione legati ai rischi in questione.

### **2.4.1 Le statistiche dell'ordine**

Mentre, nel paragrafo precedente, si è analizzato il comportamento dei massimi, ossia dei valori più ampi del campione, ora si considera quello congiunto delle statistiche di più ordini superiori. Queste risultano utili, in quanto forniscono informazioni sulla coda destra di una funzione di distribuzione.

Si comincia, dunque, con l'introdurre alcuni risultati base dei campioni ordinati, per poi procedere con l'utilizzo dell'approssimazione di Poisson, già vista come legge limite per massimi normalizzati, qui presentata come legge limite multivariata per più statistiche di ordine superiore<sup>38</sup>. Tali risultati forniscono il background teorico nel derivare proprietà limite per gli stimatori delle code.

Siano  $X, X_1, X_2, \dots$  una successione di variabili casuali non degeneri, indipendenti ed identicamente distribuite, con comune funzione di distribuzione  $F$ . Si consideri un campione finito  $X_1, \dots, X_n$  e si definisca il *campione ordinato*

$$X_{n,n} \leq \dots \leq X_{1,n}.$$

Dunque  $X_{n,n} = \min ( X_1, \dots, X_n )$  e  $X_{1,n} = M_n = \max ( X_1, \dots, X_n )$ . La variabile casuale  $X_{k,n}$  è detta *k-esima statistica di ordine superiore*.

La relazione tra la statistica dell'ordine e la funzione di distribuzione empirica di un campione è immediata: per  $x \in \mathfrak{R}$  si introduce la funzione di distribuzione campionaria ( o empirica )

$$F_n(x) = \frac{1}{n} \text{card} \{ i: 1 \leq i \leq n, X_i \leq x \} = \frac{1}{n} \sum_{i=1}^n I_{\{X_i \leq x\}}, \quad x \in \mathfrak{R},$$

dove  $I_A$  rappresenta la funzione indicatrice di un insieme  $A$ . Ora, poiché si ha

$$X_{k,n} \leq x \quad \text{se e solo se} \quad \sum_{i=1}^n I_{\{X_i > x\}} < k, \quad (\text{II.17})$$

ne consegue che

---

<sup>38</sup> EMBRECHTS P., KLÜPPELBERG C., MIKOSCH T., *Modelling extremal events...*, op. cit., pagg. 182 e sgg.

$$P ( X_{k,n} \leq x ) = P ( F_n(x) > 1 - k/n ).$$

La statistica dell'ordine superiore stima code, quantili ed anche la probabilità degli eccessi oltre una determinata soglia. Viene qui richiamata la definizione di funzione quantile di una funzione di distribuzione  $F$ :

$$F^{\leftarrow}(t) = \inf \{ x \in \mathfrak{R} : F(x) \geq t \}, \quad 0 < t < 1.$$

Per un campione  $X_1, \dots, X_n$  si indica la *funzione quantile empirica* con  $F_n^{\leftarrow}$ . Se  $F$  è una funzione continua, allora i legami nel campione sono presenti solo con probabilità nulla e quindi possono essere trascurati e, di conseguenza, si può assumere che  $X_{n,n} < \dots < X_{1,n}$ . In tal caso  $F_n^{\leftarrow}$  è una semplice funzione della statistica dell'ordine, più precisamente

$$F_n^{\leftarrow}(t) = X_{k,n} \quad \text{per } 1 - k/n < t \leq 1 - \frac{k-1}{n},$$

per  $k = 1, \dots, n$ .

Risulta ora possibile definire esplicitamente la funzione di distribuzione  $F_{k,n}$  della  $k$ -esima statistica di ordine superiore.

**Proposizione (2.11):** per  $k = 1, \dots, n$ ,  $F_{k,n}$  denoti la funzione di distribuzione di  $X_{k,n}$ . Allora

- $F_{k,n}(x) = \sum_{r=0}^{k-1} \binom{n}{r} \bar{F}^r(x) F^{n-r}(x).$
- Se  $F$  è continua, allora

$$F_{k,n}(x) = \int_{-\infty}^x f_{k,n}(z) dF(z),$$

dove

$$f_{k,n}(x) = \frac{n!}{(k-1)!(n-k)!} F^{n-k}(x) \bar{F}^{k-1}(x);$$

quindi  $f_{k,n}$  è una densità di  $F_{k,n}$  con rispetto a  $F$ .

Le considerazioni fatte finora valgono anche per la distribuzione congiunta di un numero finito di differenti statistiche dell'ordine. Se, per esempio,  $F$  è assolutamente continua con funzione di densità  $f$ , allora la funzione di densità congiunta di  $(X_1, \dots, X_n)$  è:

$$f_{X_1, \dots, X_n}(x_1, \dots, x_n) = \prod_{i=1}^n f(x_i), \quad (x_1, \dots, x_n) \in \mathfrak{R}^n.$$

Dato che gli  $n$  valori di  $(X_1, \dots, X_n)$  possono essere sistemati in  $n!$  modi diversi, ciascuna specifica collezione ordinata  $(X_{k,n})_{k=1, \dots, n}$  potrebbe derivare da  $n!$  diverse sequenze. In tal modo, la funzione di densità congiunta del campione ordinato diventa

$$f_{X_{1,n}, \dots, X_{n,n}}(x_1, \dots, x_n) = n! \prod_{i=1}^n f(x_i) \quad x_n < \dots < x_1.$$

A questo punto risulta possibile definire la funzione di densità congiunta di  $k$  statistiche di ordine superiore.

**Teorema (2.28):** se  $F$  è assolutamente continua con densità  $f$ , allora

$$f_{X_{1,n}, \dots, X_{k,n}}(x_1, \dots, x_k) = \frac{n!}{(n-k)!} F^{n-k}(x_k) \prod_{i=1}^k f(x_i), \quad x_k < \dots < x_1. \quad (II.18)$$

Molto importante risulta la definizione di *spaziatura*, ossia la differenza tra statistiche di ordine superiore successive. La loro importanza deriva dal fatto che esse costituiscono le basi per la costruzione dello stimatore di Hill.

**Definizione (2.13):** per un campione  $X_1, \dots, X_n$  le spaziature sono definite da

$$X_{k,n} - X_{k+1,n}, \quad k = 1, \dots, n-1.$$

Per variabili casuali con estremo inferiore (superiore)  $\tilde{x}_F$  ( $x_F$ ), si definisce l' $n$ -esima (0-esima) spaziatura come

$$X_{n,n} - X_{n+1,n} = X_{n,n} - \tilde{x}_F \quad (X_{0,n} - X_{1,n} = x_F - X_{1,n}).$$

Un altro importante concetto è quello della trasformazione quantile, estremamente utile, dal momento in cui permette di ridurre un problema di statistiche ordinate ad un altro di corrispondenti statistiche ordinate derivanti da un campione uniforme.

**Lemma (2.4):** siano  $X_1, \dots, X_n$  v.a. iid con comune funzione di distribuzione  $F$ .

Siano, inoltre,  $U_1, \dots, U_n$  v.a. uniformemente distribuite su  $(0,1)$  e si indichi con  $U_{n,n} < \dots < U_{1,n}$  le corrispondenti statistiche ordinate. Allora valgono i seguenti risultati:

- $F^{-1}(U_1) \stackrel{d}{=} X_1.$

- Per ogni  $n \in \mathbb{N}$ ,

$$(X_{1,n}, \dots, X_{n,n}) \stackrel{d}{=} (F^{-1}(U_{1,n}), \dots, F^{-1}(U_{n,n})).$$

- La v.a.  $F(X_1)$  ha distribuzione uniforme su  $(0,1)$  se e solo se  $F$  è una funzione continua.

Il problema che interessa risolvere, è quello di caratterizzare il comportamento asintotico, e quindi la convergenza, delle distribuzioni precedentemente analizzate. Un primo risultato è quello della convergenza quasi certa della statistica dell'ordine:

**Proposizione (2.12):** sia  $F$  una funzione di distribuzione con estremo superiore (inferiore)  $x_F \leq \infty$  ( $\tilde{x}_F \geq -\infty$ ) e  $(k(n))$  una sequenza di interi non decrescente tale per cui

$$\lim_{n \rightarrow \infty} n^{-1} k(n) = c \in [0, 1].$$

- Allora  $X_{k(n),n} \xrightarrow{\text{q.c.}} x_F(\tilde{x}_F)$  per  $c = 0$  ( $c = 1$ ).
- Si assuma che  $c \in (0, 1)$  sia tale che esista un'unica soluzione  $x(c)$  dell'equazione  $\bar{F}(x) = c$ . Allora

$$X_{k(n),n} \xrightarrow{\text{q.c.}} x(c).$$

Si considerino le v.a. iid  $X_1, \dots, X_n$  con comune funzione di distribuzione  $F$ . Si ricordi che, dalla definizione dell'approssimazione di Poisson<sup>39</sup>, per una sequenza  $(u_n)$  di soglie e  $0 \leq \tau \leq \infty$ ,

$$\lim_{n \rightarrow \infty} P(X_{1,n} \leq u_n) = e^{-\tau} \Leftrightarrow \lim_{n \rightarrow \infty} n \bar{F}(u_n) = \tau.$$

Il problema, a questo punto, è quello di capire se sia possibile estendere la precedente relazione ad ogni statistica di ordine superiore  $X_{k,n}$  e se si possa ottenere una probabilità limite congiunta per un fissato numero  $k$  di statistiche di ordine superiore  $X_{k,n}, \dots, X_{1,n}$ .

<sup>39</sup> L'approssimazione di Poisson è definita nella proposizione (2.4), a pag. 67.

Si consideri, per  $n \in \mathbb{N}$ , il numero di eccessi al di sopra dei valori soglia  $(u_n)$  da parte di  $X_1, \dots, X_n$ :

$$B_n = \sum_{i=1}^n I_{\{X_i > u_n\}}.$$

Dunque  $B_n$  rappresenta una variabile casuale binomiale con parametri  $n$  e  $\bar{F}(u_n)$ . Risulta importante notare che le eccedenze  $\{X_i > u_n\}$  tendono a diventare sempre più rare all'aumentare della soglia; d'altra parte, però, si aumenta anche l'ampiezza del campione. Si bilanciano questi due effetti in modo tale che  $E[B_n] = n\bar{F}(u_n) \rightarrow \tau$  per  $n \rightarrow \infty$  e di conseguenza si applica immediatamente il teorema classico di Poisson:  $B_n \xrightarrow{d} \text{Poi}(\tau)$ . Le soglie  $u_n$  sono scelte in modo tale che il numero atteso di eccedenze converga. Il teorema che viene ora enunciato fornisce la *legge limite per il numero di eccedenze*, ed inoltre mostra che la condizione  $n\bar{F}(u_n) \rightarrow \tau$  è anche necessaria affinché valga l'approssimazione di Poisson.

**Teorema (2.29):** *si supponga che  $(u_n)$  sia una sequenza in  $\mathfrak{R}$  tale che  $n\bar{F}(u_n) \rightarrow \tau$  per alcune  $\tau \in [0, \infty]$ , per  $n \rightarrow \infty$ . Allora*

$$\lim_{n \rightarrow \infty} P(B_n \leq k) = e^{-\tau} \sum_{r=0}^k \frac{\tau^r}{r!}, \quad k \in \mathbb{N}_0.$$

*Per  $\tau = 0$ , il lato sinistro della equazione viene considerato pari a 1, mentre per  $\tau = \infty$  lo si considera nullo.*

*Se la precedente equazione vale per alcune  $k \in \mathbb{N}_0$ , allora  $n\bar{F}(u_n) \rightarrow \tau$  per  $n \rightarrow \infty$  e quindi tale equazione vale per tutte le  $k \in \mathbb{N}_0$ .*

L'approssimazione di Poisson, vista nel precedente teorema, ci permette di derivare i comportamenti asintotici per la generica statistica di ordine  $k$ -esimo. Dalla definizione di  $B_n$  e dalla (II.17) si deriva che

$$P(B_n < k) = P(X_{k,n} \leq u_n), \quad 1 \leq k \leq n$$

E, di conseguenza, è possibile definire le probabilità limite e la distribuzione limite di una statistica di ordine superiore, come mostrano, rispettivamente, i seguenti teorema e corollario.

**Teorema (2.30):** *si supponga che  $(u_n)$  sia una sequenza in  $\mathfrak{R}$  tale che  $n \bar{F}(u_n) \rightarrow \tau \in [0, \infty]$ , per  $n \rightarrow \infty$ . Allora*

$$\lim_{n \rightarrow \infty} P(X_{k,n} \leq u_n) = e^{-\tau} \sum_{r=0}^{k-1} \frac{\tau^r}{r!}, \quad k \in \mathfrak{N}.$$

*Per  $\tau = 0$ , il lato destro della equazione viene considerato pari a 1, mentre per  $\tau = \infty$  lo si considera nullo.*

*Se la precedente equazione vale per alcune  $k \in \mathfrak{N}$ , allora  $n \bar{F}(u_n) \rightarrow \tau$  per  $n \rightarrow \infty$  e quindi tale equazione vale per tutte le  $k \in \mathfrak{N}$ .*

Per  $u_n = c_n x + d_n$  e  $\tau = \tau(x) = -\ln H(x)$ , come nella proposizione (2.10), si ottiene:

**Corollario (2.4):** *si supponga che  $F \in MDA(H)$  con costanti  $c_n > 0$  e  $d_n \in \mathfrak{R}$ .*

*Si definisca*

$$H^{(k)}(x) = H(x) \sum_{r=0}^{k-1} \frac{(-\ln H(x))^r}{r!}, \quad x \in \mathfrak{R}.$$

*Per le  $x$  tali per cui  $H(x) = 0$  si considera  $H^{(k)}(x) = 0$ . Dunque per ogni  $x \in \mathfrak{R}$*

$$\lim_{n \rightarrow \infty} P(c_n^{-1}(X_{k,n} - d_n) \leq x) = H^{(k)}(x).$$

*D'altra parte, se per alcune  $k \in \mathfrak{N}$*

$$\lim_{n \rightarrow \infty} P(c_n^{-1}(X_{k,n} - d_n) \leq x) = G(x), \quad x \in \mathfrak{R}$$

per una funzione di distribuzione non degenera  $G$ , allora  $G = H^{(k)}$  per alcune distribuzioni del valore estremo  $H$ .

In altre parole, la distribuzione limite della statistica del  $k$ -esimo ordine superiore è stata ottenuta considerando il numero di eccessi rispetto una soglia  $u_n$  da parte di  $X_1, \dots, X_n$ . Simili argomenti possono essere adattati per provare la convergenza della distribuzione congiunta di più statistiche di ordine superiore. A questo scopo, si supponga che, per  $k \in \mathbb{N}$ , i livelli  $u_n^{(k)} \leq \dots \leq u_n^{(1)}$  soddisfino il seguente limite:

$$\lim_{n \rightarrow \infty} n \bar{F}(u_n^{(i)}) = \tau_i, \quad i = 1, \dots, k, \quad (\text{II.19})$$

dove  $0 \leq \tau_1 \leq \tau_2 \leq \dots \leq \tau_k \leq \infty$  e si definisca

$$B_n^{(j)} = \sum_{i=1}^n I_{\{X_i > u_n^{(j)}\}}, \quad j = 1, \dots, k,$$

dunque  $B_n^{(j)}$  rappresenta il numero degli eccessi al di sopra di  $u_n^{(j)}$  da parte di  $X_1, \dots, X_n$ . Risulta dunque ora possibile presentare il teorema contenente la *legge limite multivariata del numero degli eccessi*:

**Teorema (2.31):** si supponga che le sequenze  $(u_n^{(j)})$  soddisfino la (II.19) per  $j = 1, \dots, k$ . Allora, per  $l_1, \dots, l_k \in \mathbb{N}_0$ ,

$$\lim_{n \rightarrow \infty} P(B_n^{(1)} = l_1, B_n^{(2)} = l_1 + l_2, \dots, B_n^{(k)} = l_1 + \dots + l_k) = \frac{\tau_1^{l_1}}{l_1!} \frac{(\tau_2 - \tau_1)^{l_2}}{l_2!} \dots \frac{(\tau_k - \tau_{k-1})^{l_k}}{l_k!} e^{-\tau_k}$$

La parte destra della presente equazione viene considerata nulla, se  $\tau_k = \infty$ .

Chiaramente,

$$P(X_{1,n} \leq u_n^{(1)}, \dots, X_{k,n} \leq u_n^{(k)}) = P(B_n^{(1)} = 0, B_n^{(2)} \leq 1, \dots, B_n^{(k)} \leq k - 1), \quad (\text{II.20})$$

e dunque la distribuzione asintotica congiunta delle statistiche di  $k$  ordini superiori può essere direttamente derivata dall'ultimo teorema enunciato. In particolare, se  $c_n^{-1}(X_{1,n} - d_n)$  converge debolmente, allora lo stesso succede al vettore

$$(c_n^{-1}(X_{1,n} - d_n), \dots, c_n^{-1}(X_{k,n} - d_n)).$$

Nonostante, per piccoli valori di  $k$ , la distribuzione limite congiunta delle  $k$  statistiche di ordine superiore possono essere facilmente derivate dalla (II.20) e dall'ultimo teorema, il caso generale risulta, invece, piuttosto complicato. Se la funzione di distribuzione  $F$  è assolutamente continua con funzione di densità  $f$ , che soddisfa determinate condizioni di regolarità, ossia, se  $F \in \text{MDA}(H)$ , allora la funzione di distribuzione del massimo  $F^n(c_n x + d_n) = P(c_n^{-1}(X_{1,n} - d_n) \leq x)$  ha anch'essa una funzione di densità tale per cui, per quasi tutte le  $x \in \mathfrak{R}$ ,

$$n c_n f(c_n x + d_n) F^{n-1}(c_n x + d_n) \sim n c_n f(c_n x + d_n) H(x) \rightarrow h(x),$$

dove  $h(x)$  rappresenta la densità della funzione di distribuzione del valore estremo  $H$ . Inoltre, per  $k \in \mathbb{N}$  il limite debole del vettore casuale  $(c_n^{-1}(X_{j,n} - d_n))_{j=1, \dots, k}$  ha, per la (II.18), funzione di densità

$$\lim_{n \rightarrow \infty} F^{n-k}(c_n x_k + d_n) \prod_{j=1}^k ((n-j+1) c_n f(c_n x_j + d_n)) = H(x_k) \prod_{j=1}^k \frac{h(x_j)}{H(x_j)}, \quad x_k < \dots < x_1. \quad (\text{II.21})$$

Viene ora fornita una definizione di vettore  $k$ -dimensional,  $H$ -estremal variate:

**Definizione (2.14):** per ogni funzione di distribuzione del valore estremo  $H$  con funzione di densità  $h$ , si definisca per  $x_k < \dots < x_1$ , nel supporto di  $H$ ,

$$h^{(k)}(x_1, \dots, x_k) = H(x_k) \prod_{j=1}^k \frac{h(x_j)}{H(x_j)}.$$

Un vettore  $(Y^{(1)}, \dots, Y^{(k)})$  di variabili casuali con funzione di densità congiunta  $h^{(k)}$  è detto  $k$ -dimensional  $H$ -estremal variate.

Quanto detto sinora, riguardo la distribuzione limite congiunta delle  $k$  statistiche di ordine superiore, viene infine raccolto nel seguente teorema e consente di ottenere una spiegazione plausibile dell'espressione (II.21):

**Teorema (2.32):** *si assuma che  $F \in MDA(H)$  con costanti  $c_n > 0$  e  $d_n \in \mathfrak{R}$ .*

*Allora, per ogni  $k \in \mathfrak{N}$  fissato*

$$(c_n^{-1} (X_{i,n} - d_n))_{i=1, \dots, k} \xrightarrow{d} (Y^{(i)})_{i=1, \dots, k}, \quad n \rightarrow \infty,$$

*dove  $(Y^{(1)}, \dots, Y^{(k)})$  è un  $k$ -dimensional  $H$ -extremal variate.*

### 2.4.2 I processi di punti

Si può pensare ad un processo di punti  $N$  come ad una distribuzione casuale di punti  $X_i$  nello spazio. Per una determinata configurazione  $(X_i)$  ed un insieme  $A$ ,  $N(A)$  conta il numero di  $X_i \in A$ . La figura 8 fornisce una semplice rappresentazione grafica di quanto appena detto.

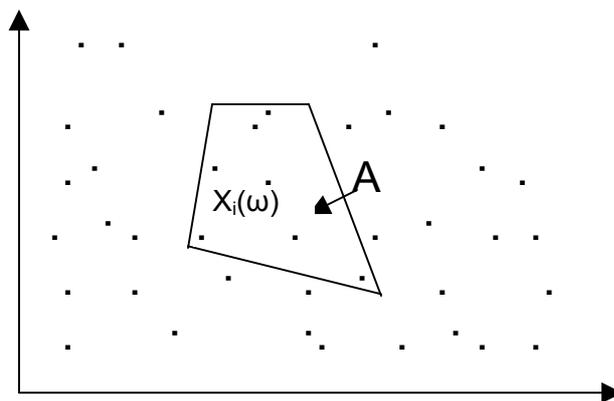


Figura 8- Una possibile configurazione di punti casuali  $X_i$  in  $\mathfrak{R}_+ \times \mathfrak{R}_+$ . Il numero di punti che ricade all'interno dell'insieme  $A$  costituisce la variabile  $N(A)$ . Nel presente caso  $N(A, \omega) = 7$ .

Si consideri, ora, una sequenza  $(X_n)$  di vettori casuali nel cosiddetto spazio degli stati  $E$  e si definisca, per  $A \subset E$ ,

$$N(A) = \text{card} \{ i: x_i \in A \}; \quad A \in \varepsilon,$$

in altri termini,  $N(A)$  conta il numero delle  $X_i$  che ricadono all'interno di  $A$ . Naturalmente,  $N(A) = N(A, \omega)$  è aleatorio per un dato insieme  $A$  e, sotto condizioni generali,  $N(\cdot, \omega)$  definisce una misura di conteggio aleatoria su un'opportuna  $\sigma$ -algebra  $\varepsilon$  di sottoinsiemi di  $E$ . Questo rappresenta il significato intuitivo del processo di punti  $N$ .

Lo spazio degli stati  $E$ , dove i punti si trovano, è un sottoinsieme di uno spazio euclideo  $n$ -dimensionale (che includa possibilmente punti di coordinate infinite), dotato di una  $\sigma$ -algebra  $\varepsilon$  dell'insieme di Borel. Risulta opportuno che un processo di punti venga scritto utilizzando la *misura di Dirac*  $\varepsilon_x$ , per  $x \in E$ :

$$\varepsilon_x(A) = \begin{cases} 1 & \text{se } x \in A, \\ 0 & \text{se } x \notin A, \end{cases} \quad A \in \varepsilon.$$

Per una data sequenza  $(x_i)_{i \geq 1}$  in  $E$ ,

$$m(A) = \sum_{i=1}^{\infty} \varepsilon_{x_i}(A) = \sum_{i: x_i \in A} 1 = \text{card} \{ i: x_i \in A \}, \quad A \in \varepsilon,$$

definisce una misura di conteggio su  $\varepsilon$  che è detta *misura di punti* se  $m(K) < \infty$ , per tutti gli insiemi compatti  $K \in E$ . Sia, inoltre,  $M_p(E)$  lo spazio di tutte le misure di punti su  $E$ , dotato di un'appropriata  $\sigma$ -algebra  $M_p(E)$ . Risulta ora possibile dare una definizione di processo di punti

**Definizione (2.15):** *un processo di punti su  $E$  è una funzione misurabile del tipo:*

$$N: [\Omega, F, P] \rightarrow [M_p(E), M_p(E)].$$

Un processo di punti è un elemento casuale di una funzione casuale che assume misure di punti come valori.

Si assuma che  $m = \sum_{i=1}^{\infty} \varepsilon_{x_i}$  sia una misura di punti su  $E$ . Sia  $(y_i)$  una sottosequenza di  $(x_i)$ , contenente valori tutti mutuamente distinti  $(x_i)$  senza alcuna ripetizione. Si definisca la *molteplicità* di  $y_i$  come

$$N_i = \text{card} \{ j: j \geq 1, y_j = x_i \}.$$

Dunque, si può scrivere che

$$m = \sum_{i=1}^{\infty} n_i \varepsilon_{y_i}.$$

Per  $n_i = 1$ , per ogni  $i$ , allora  $m$  è detta *misura di punti semplice*, altrimenti è detta *moltiplica*. Analogamente, se le realizzazioni dei processi di punti  $N$  sono solo misure di punti semplici, allora  $N$  è un *processo di punti semplice* e si può scrivere

$$P[N(\{x\}) \leq 1, x \in E] = 1;$$

in caso contrario  $N$  è *multiplo*.

Un importante esempio di processo di punti, strettamente correlato alla teoria del valore estremo, è quello degli eccessi; sia  $u$  un numero reale ed  $(X_n)$  una sequenza di v.a. Allora il *processo di punti degli eccessi*

$$N_n(\cdot) = \sum_{i=1}^n \varepsilon_{n^{-1}i}(\cdot) I_{\{X_i > u\}}, \quad n = 1, 2, \dots,$$

con spazio degli stati  $E = (0, 1]$ , conta il numero di eccessi al di sopra di una soglia  $u$ , da parte della sequenza  $X_1, \dots, X_n$ . Ad esempio, preso l'intero intervallo  $(0, 1]$ , allora

$$N_n(0, 1] = \text{card} \{ i: 0 < n^{-1}i \leq 1 \text{ e } X_i > u \} = \text{card} \{ i \leq n: X_i > u \}.$$

Si può, a questo punto, notare il legame con la teoria del valore estremo. Infatti, si indichi con  $X_{k,n}$  la  $k$ -esima statistica di ordine superiore del campione  $X_1, \dots, X_n$ . Allora

$$\{ N_n(0, 1] = 0 \} = \{ \text{card} \{ i \leq n: X_i > u \} = 0 \} = \{ \text{nessuna delle } X_i, i \leq n, \text{ eccede } u \} = \\ \{ \max(X_1, \dots, X_n) \leq u \};$$

$$\{ N_n(0, 1] < k \} = \{ \text{card} \{ i \leq n: X_i > u \} < k \} = \{ \text{meno di } k \text{ tra le } X_i, i \leq n, \text{ eccede } u \} = \\ \{ \text{la statistica dell'ordine } X_{k,n} \text{ non eccede } u \} = \{ X_{k,n} \leq u \}.$$

La distribuzione di  $N$  è unicamente determinata dalla famiglia delle distribuzioni di vettori aleatori di dimensione finita

$$(N(A_1), \dots, N(A_m))$$

per ogni scelta di  $A_1, \dots, A_m \in \mathcal{E}$  ed  $m \geq 1$ . D'altra parte, quello appena presentato è un vettore casuale di v.a. a valori interi, che è dato completamente dalle seguenti probabilità

$$P(N(A_1) = k_1, \dots, N(A_m) = k_m), \quad k_i \geq 0, i = 1, \dots, m.$$

Risulta spesso conveniente definire la funzione di distribuzione di una v.a. o di un vettore aleatorio tramite alcune medie analitiche; uno strumento di questo genere, nel caso del processo di punti, è rappresentato dal funzionale di Laplace, che permette di determinare completamente la funzione di distribuzione di  $N$ .

**Definizione (2.16):** *il funzionale di Laplace di un processo di punti  $N$  è dato da*

$$\Psi_N(g) = E \exp \left\{ - \int_E g dN \right\} = \int_{M_p(E)} \exp \left\{ - \int_E g(x) dm(x) \right\} dP_N(m)$$

dove  $g$  è una funzione misurabile non-negativa sullo spazio degli stati  $E$ .

I processi di punti sono collezioni di variabili di conteggio. Tra di esse, la più semplice, e probabilmente la più utile, è quella distribuita secondo una binomiale:  $B_n = \sum_{i=1}^n I_{\{X_i \in A_n\}}$ , che, per  $X_i$  iid, conta il numero di “successi”  $\{X_i \in A_n\}$ , tra  $X_1, \dots, X_n$ , e  $p_n = P(X_1 \in A_n)$  è la probabilità di successo. Dunque, dal teorema di Poisson, si ha che, se  $p_n \sim \lambda/n$ , allora  $B_n \xrightarrow{d} \text{Poi}(\lambda)$ . Tale risultato porta alla seguente definizione di misura aleatoria di Poisson; sia  $\mu$  una misura Radon su  $\varepsilon$ , cioè  $\mu(A) < \infty$  per insiemi compatti  $A \subset E$ .

**Definizione (2.17):** *un processo di punti  $N$  è detto processo di Poisson o misura aleatoria di Poisson con misura media  $\mu$  (ossia  $PRM(\mu)$ ) se le due seguenti condizioni sono soddisfatte:*

- per  $A \in \varepsilon$ ,

$$P(N(A) = k) = \begin{cases} e^{-\mu(A)} \frac{(\mu(A))^k}{k!} & \text{se } \mu(A) < \infty, \\ 0 & \text{se } \mu(A) = \infty, \end{cases} \quad k \geq 0.$$

- per ogni  $m \geq 1$ , se  $A_1, \dots, A_m$  sono insiemi mutuamente disgiunti in  $\varepsilon$ , allora  $N(A_1), \dots, N(A_m)$  sono v.a. indipendenti.

Il termine “misura media” è giustificato dal fatto che  $E(N(A)) = \mu(A)$ . Dal momento in cui la distribuzione di Poisson è determinata dal suo valore medio, ne consegue

che  $PRM(\mu)$  è anch'essa determinata dalla sua misura media  $\mu$ . Inoltre, la trasformata di una PRM è ancora una PRM.

La convergenza debole di un processo di punti rappresenta uno degli strumenti più importanti per la teoria asintotica dei valori estremi. Bisogna, però, innanzitutto precisare che cosa si intenda per convergenza di un processo di punti. Si considerino i processi di punti  $N, N_1, N_2, \dots$  sullo stesso spazio degli stati  $E \subset \overline{\mathfrak{R}}^d$ . Dunque, si sa che la distribuzione di questi processi di punti in  $M_p(E)$ , lo spazio di tutte le misure di punti su  $E$ , è determinata dalle loro distribuzioni di dimensioni finite e, pertanto, una naturale richiesta per la convergenza debole di  $(N_n)$  ad  $N$  dovrebbe essere, per ogni scelta di "opportuni" insiemi di Borel  $A_1, \dots, A_m \in \varepsilon$  e per ogni intero  $m \geq 1$ , che risulti

$$P(N_n(A_1), \dots, N_n(A_m)) \rightarrow P(N(A_1), \dots, N(A_m)). \quad (II.22)$$

Si arriva, dunque, ad un primo e fondamentale risultato relativo alla convergenza debole

**Definizione (2.18):** siano  $N, N_1, N_2, \dots$  processi di punti sullo spazio degli stati  $E \subset \overline{\mathfrak{R}}^d$  con una  $\sigma$ -algebra  $\varepsilon$  di insiemi di Borel. Si può dire che  $(N_n)$  converge debolmente ad  $N$  in  $M_p(E)$  (ossia  $N_n \xrightarrow{d} N$ ) se la (II.22) è soddisfatta per tutte le possibili scelte di insiemi  $A_i \in \varepsilon$ , per cui valga  $P(N(\partial A_i) = 0)$ , essendo  $i = 1, \dots, m$  ed  $m \geq 1$ .<sup>40</sup>

Si assuma ora che lo spazio degli stati  $E$  sia un intervallo  $(a, b] \in \mathfrak{R}$ . La convergenza di distribuzioni di dimensione finita può essere verificata tramite semplici strumenti, come mostra il seguente teorema, introdotto da Kallenberg. Si riprenda prima il concetto di processo di punti semplice, i cui punti hanno, cioè, molteplicità 0 o 1 con probabilità 1.

**Teorema (2.33):** siano  $(N_n)$  ed  $N$  processi di punti su  $E = (a, b] \in \mathfrak{R}$  ed  $N$  sia un processo semplice. Si supponga che le due seguenti condizioni valgano:

- $E(N_n(A)) \rightarrow E(N(A))$   
per tutti gli intervalli  $A = (c, d]$ , con  $a < c < d \leq b$  e

---

<sup>40</sup>  $\partial A$  denota la frontiera di  $A$ .

- $$P(N_n(B) = 0) \rightarrow P(N(B) = 0)$$

per tutte le unioni  $B = \bigcup_{i=1}^k (c_i, d_i]$  di intervalli mutuamente disgiunti  $(c_i, d_i]$  tali che

$$a < c_1 < d_1 < \dots < c_k < d_k \subseteq b$$

e per ogni  $k \geq 1$ .

Allora  $N_n \xrightarrow{d} N$  in  $M_p(E)$ .

Il funzionale di Laplace è uno strumento utile anche per la verifica della convergenza debole di un processo di punti. Risulta, infatti, facile verificare che la convergenza debole di una sequenza di processi di punti è equivalente alla convergenza dei loro funzionali di Laplace su un'opportuna famiglia di funzioni  $g$ . Si deve ricordare che la funzione, a valori reali,  $g$  ha un supporto compatto se esiste un insieme  $K \subset E$  tale che  $g(x) = 0$  su  $K^c$ , complemento di  $K$ . Dunque si definisce

$$C_K^+(E) = \{ g : g \text{ sia una funzione continua, non negativa su } E \text{ con supporto compatto} \}.$$

Detto ciò, è ora possibile presentare il seguente teorema, che definisce il criterio per la convergenza debole di un processo di punti, utilizzando i funzionali di Laplace.

**Teorema (2.34):** *il processo di punti  $(N_n)$  converge debolmente al processo di punti  $N$  in  $M_p(E)$  se e solo se il corrispondente funzionale di Laplace converge per ogni  $g \in C_K^+(E)$  per  $n \rightarrow \infty$ , cioè se*

$$\Psi_{N_n}(g) = E \exp\left\{-\int_E g dN_n\right\} \rightarrow \Psi_N(g) = E \exp\left\{-\int_E g dN\right\}. \quad (II.23)$$

Si è menzionato che la (II.23), per ogni  $g \in C_K^+(E)$  equivale a dire che  $\int_E g dN_n \xrightarrow{d} \int_E g dN$  per ogni  $g \in C_K^+(E)$ . Infatti, se  $g \in C_K^+(E)$ , allora  $zg \in C_K^+(E)$ , per  $z > 0$ . Dunque, la (II.23) implica la convergenza della trasformata di Laplace delle v.a.

$\int_E g dN_n$  e viceversa. D'altra parte, la convergenza della trasformata di Laplace di v.a. non negative è equivalente alla loro convergenza in distribuzione.

Precedentemente<sup>41</sup>, è stato introdotto il processo di punti degli eccessi oltre una soglia  $u_n$ , da parte delle v.a.  $X_1, \dots, X_n$ ,

$$N_n(\cdot) = \sum_{i=1}^n \varepsilon_{n^{-1}i}(\cdot) I_{\{X_i > u\}}, \quad n = 1, 2, \dots$$

È anche stato evidenziato lo stretto legame con la teoria del valore estremo; essendo, infatti,  $X_{n,n} < \dots < X_{1,n}$  la statistica dell'ordine del campione  $X_1, \dots, X_n$  ed  $M_n = X_{1,n}$ , si è dimostrato che

$$\begin{aligned} \{ N_n(0, 1] = 0 \} &= \{ M_n \leq u_n \}, \\ \{ N_n(0, 1] < k \} &= \{ X_{k,n} \leq u \}. \end{aligned}$$

Per poter infine esporre il teorema inerente la convergenza debole di un processo di punti degli eccessi, si assuma ora che le  $X_n$  siano v.a. iid e  $(u_n)$  sia una sequenza di soglie reali; si riprenda, inoltre, l'approssimazione di Poisson<sup>42</sup>, secondo la quale, per ogni  $\tau \in [0, \infty]$ , la relazione  $P( M_n \leq u_n ) \rightarrow \exp \{ -\tau \}$  vale se e solo se

$$n \bar{F}(u_n) = E \sum_{i=1}^n I_{\{X_i > u_n\}} \rightarrow \tau. \quad (II.24)$$

Quest'ultima condizione assicura che ci sono, in media, all'incirca  $\tau$  eccessi oltre la soglia  $u_n$ , da parte di  $X_1, \dots, X_n$ , ed implica inoltre la convergenza debole del processo di punti  $N_n$ .

**Teorema (2.35):** *si supponga che  $(X_n)$  sia una sequenza di v.a. iid con comune funzione di distribuzione  $F$ . Sia  $(u_n)$  una sequenza di valori soglia tali che la (II.24) valga per qualche  $\tau \in (0, \infty)$ . Allora il processo di punti degli eccessi  $N_n$  converge debolmente in  $M_p(E)$  ad un processo di Poisson omogeneo  $N$  in  $E = (0, 1]$  con intensità  $\tau$ .  $N$ , dunque, è una misura aleatoria di Poisson con misura media  $\tau |\cdot|$  (ossia  $PRM(\tau |\cdot|)$ ), dove  $|\cdot|$  denota la misura di Lebesgue su  $E$ .*

<sup>41</sup> A tale proposito si veda pag. 99.

<sup>42</sup> L'approssimazione di Poisson è definita nella proposizione (2.4), a pag. 67.

Prima di concludere il presente capitolo, si ritiene necessario presentare un concetto con cui si ha a che fare nella vita di tutti i giorni e che risulta importante per l'analisi in questione: quello dei *records*.

Se si considerano le osservazioni  $X_n$ , un record è un massimo (o un minimo) temporaneo di tale sequenza, che cambia col passare del tempo: un record  $X_n$  ha luogo se  $X_n > M_{n-1}$ . Chiaramente, in tal caso, il nuovo massimo  $M_n$  andrà a coincidere con  $X_n$ . Si noti, inoltre, che ad un record corrisponde un salto nella sequenza  $(M_n)$  e gli istanti in cui questi salti avvengono,  $L_1 < L_2 < \dots$ , sono aleatori e sono detti *istanti del record* di  $(X_n)$ .

La nozione ora presentata risulta importante a livello assicurativo e finanziario, in quanto rende abbastanza l'idea di quanto di buono o di cattivo possa accadere nel futuro, a livello sia di frequenza che di importanza: grandi salti nei prezzi possono, infatti, portare al fallimento di istituzioni finanziarie e, in un portafoglio assicurativo, sinistri di grande entità, possono causare problemi di insolvenza.

Il seguente teorema descrive la sequenza di records  $(X_{L_n})$  in termini di misura aleatoria di Poisson (PRM).

**Teorema (2.36):** *sia  $F$  una funzione di distribuzione continua con estremo inferiore  $x_F^l$  e superiore  $x_F^r$ . Allora i records  $(X_{L_n})$  della sequenza  $(x_n)$  sono i punti di una PRM( $\mu$ ) su  $(x_F^l, x_F^r)$  con misura media  $\mu$  data da*

$$\mu(a, b] = R(b) - R(a), \quad x_F^l < a \leq b < x_F^r, \quad \text{dove } R(x) = -\ln \bar{F}(x).$$

*In particolare, se  $F$  è una esponenziale standard, allora  $R(t) = t$  e  $(X_{L_n}) \stackrel{d}{=} (\Gamma_n)$  sono i punti di un processo di Poisson omogeneo su  $\mathfrak{R}_+$  con intensità 1.*