LA DIMENSIONE DEI SINISTRI: UNA ESEGESI PERSONALE DELLA DISTRIBUZIONE
di GIOVANNA FERRARA

Introduzione

Nel titolo della mia esposizione ho voluto includere l'aggettivo "personale" poiché tratterò la dimensione dei sinistri seguendo un mio percorso, attraverso gli studi che hanno maggiormente influenzato la mia attività nel campo dell'assicurazione e della riassicurazione.

L'esposizione sarà anche personale per una mia predilezione letteraria per gli studi ben strutturati, quelli in cui il numero delle parole supera quello delle formule. Ho bisogno, in altri termini, di comprendere pienamente l'autore e le sue ragioni e subisco il fascino del testo come se stessi leggendo un coinvolgente racconto. Va riconosciuto che anche nel campo attuariale vi sono autori avvincenti. Valga per tutti l'esempio di Bruno De Finetti che, con una prosa ricca di riferimenti alla realtà, presenta problemi con splendide ed illuminanti immagini: le formule sono solo una parte non rilevante dei loro scritti (anche se le uniche ampiamente saccheggiate dagli studiosi).

Mi propongo, pertanto, di trattare le distribuzioni della dimensione dei sinistri con una serie di narrazioni, che, a mio parere, hanno tutte i pregi della leggerezza e dell'esattezza. Il primo "racconto" tratto dalla nostra letteratura probabilistica e attuariale, è una pagina del Cramér[1], di cui darò una traduzione fedele.

Una introduzione alla distribuzione log-normale (versione di H.Cramér)

"Sarà conveniente usare qui la terminologia direttamente connessa ai problemi di biologia. Se la nostra variabile causale è la dimensione di un specifico organismo che stiamo osservando, la dimensione da esso raggiunta al momento attuale in un particolare individuo può essere considerata come l'effetto congiunto di un gran numero di cause mutuamente indipendenti, agenti in una sequenza ordinata durante il tempo di crescita dell'individuo. Se queste cause semplicemente aggiungono i loro effetti (effetti che presumiamo siano variabili causali), in base al teorema del limite centrale, possiamo affermare che la loro somma è asintoticamente distribuita normalmente.

Tuttavia, in generale non sembra essere plausibile che le cause contribuiscono in maniera additiva. Sembra più naturale supporre che ciascuna causa produca un impulso, l'effetto del quale dipende sia dalla forza dell'impulso che dalla dimensione raggiunta dall'organismo al momento in cui l'impulso agisce.

Supponiamo di avere n impulsi $\xi_1, \xi_2, ..., \xi_n$ che agiscono nell'ordine indicato dai loro indici. Questi impulsi possono essere considerati variabili causali. Se indichiamo con x la dimensione prodotta dagli impulsi $\xi_1, \xi_2, ..., \xi_n$, possiamo ad esempio supporre che l'incremento causato dall'impulso ξ_{n+1} sia proporzionale a ξ_{n+1} a qualche funzione $g(x)$ della dimensione momentanea dell'organismo

[1] Harald Cramér – Mathematical Methods in Statistics – Princeton 1974

$$x_{n+1} = x_n + \xi_{n+1}\, g(x_n)$$

Ne segue che

$$\xi_1 + \xi_2 + \ldots\ldots.. + \xi_n = \sum_{0}^{n-1} \frac{x_{v+1} - x_v}{g(x_v)}$$

Se poi ciascun impulso fornisce solo un modesto contributo alla crescita dell'organismo avremmo approssimativamente:

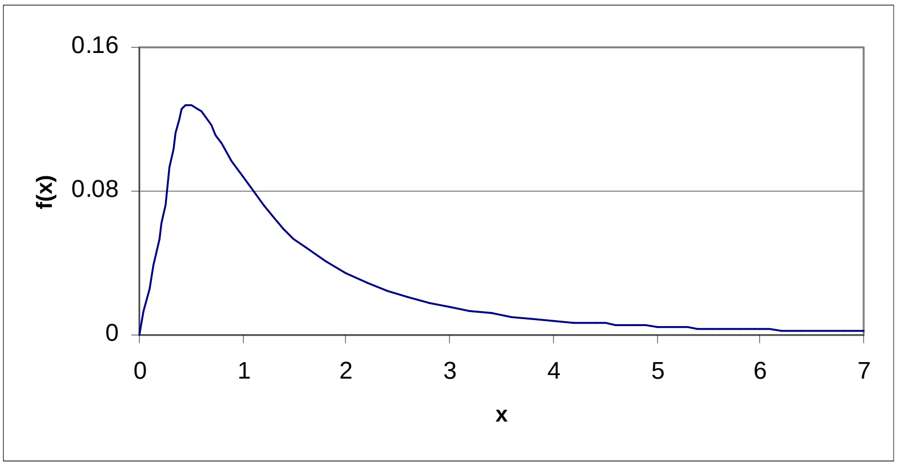
$$\xi_1 + \xi_2 + \ldots\ldots.. + \xi_n = \int_{x_0}^{x} \frac{d(t)}{g(t)}$$

dove $x = x_n$ indica la dimensione finale dell'organismo. Per ipotesi $\xi_1, \xi_2, \ldots\ldots\ldots\ldots, \xi_n$ sono variabili indipendenti e n può essere considerato un numero grande. Sotto le condizioni di validità del teorema del limite centrale ne segue che, al limite, la funzione della variabile causale x che figura al secondo membro dell'equazione precedente è distribuita normalmente.

Consideriamo ora il caso $g(t) = t$, cioè quando l'effetto di ciascun impulso è direttamente proporzionale alla dimensione momentanea dell'organismo. In questo caso troviamo che $log(x)$ è distribuita normalmente. Se più generalmente $log(x - a)$ è una variabile normale (m, σ) è facilmente deducibile che la variabile x ha come funzione di frequenza:

$$\frac{1}{\sigma(x-a)\sqrt{2\pi}}\, e^{-\frac{(log(x-a)-m)^2}{2\sigma^2}}$$

con $x > a$ mentre per $x <= a$ la funzione di densità è zero.



Funzione di frequenza di una variabile lognormale con *m = 0* e $\sigma = 2$"

La prima volta che ho letto questo racconto nella versione di H. Cramér stavo esaminando le statistiche del ramo incendi per i rischi industriali e quindi mi veniva spontaneo parafrasare questo testo sostituendo la parola organismo con incendio. Quindi ho riformulato il mio problema[2] dicendo:

Nel ramo incendi possiamo ammettere che la dimensione dei sinistri sia la risultante di un gran numero di cause ad effetti positivi. Si può facilmente immaginare qualcuna di queste cause che consentono ad esempio l'espansione di un incendio: condizioni atmosferiche particolari, distanza o inefficienza dei mezzi di prevenzione, ora di avvenimento del sinistro, ecc. Tali fattori, per i quali è lecito supporre l'indipendenza, possono essere rappresentati dalle variabili .causali.

$$\xi_1, \ \xi_2, ..., \xi_n \qquad \text{e così via .}$$

Approfondimenti della legge lognormale

Il racconto di H. Cramér mi aveva conquistato e mi spingeva a saperne di più. Quindi consultai un libro completamente dedicato alla lognormale. Si tratta di una monografia[3] che presenta tutte le caratteristiche di alta letteratura: ciascun capitolo viene introdotto con citazioni tratte da Shakespeare e si avvale di una eccellente e graduale costruzione. Come dichiarato dagli autori nella prefazione, lo scopo originario del lavoro era quello raccogliere il disperso materiale pubblicato sulla distribuzione (definita come una vera Cenerentola fra tutte le distribuzioni e pertanto sottovalutata rispetto alla binomiale o alla normale). Tuttavia, nel corso dell'opera, gli autori realizzarono che vi erano ancora problemi non esplorati e proprietà della distribuzione non adeguatamente poste in rilievo.

Il risultato del lavoro di J.Aitchison e J. A.C. Brown è, a mio parere, significativo e mostra come si possa mettere a disposizione degli studiosi un vasto materiale relativo ad un unico argomento. Purtroppo di libri come questi n'esistono pochi.

Nel trattare l'applicazione della lognormale ai fenomeni economici (distribuzione dei redditi e dei consumi), gli autori premettono alcuni criteri di base per l'adozione dei modelli matematici, cioè:

1) la possibilità di dedurre la forma della distribuzione da ipotesi realistiche ed elementari;
2) la facilità con cui la funzione può essere trattata nelle analisi;
3) il significato che può essere attribuito ai parametri;
4) il grado d'accostamento ai dati reali.

Ho citato questi criteri nell'ordine in cui sono stati forniti dagli autori, perché in generale il grado di accostamento ai dati reali sembra essere l'unica norma seguita dalla maggior parte degli statistici. In effetti come precisato nella monografia non vi è alcuna ragione di credere che una espressione matematica sia superiore alle altre in rapporto ai quattro criteri precisati. In particolare se si vuole scegliere in base al solo grado di accostamento alle esperienze empiriche occorre formulare le ipotesi sul campo di validità della formulazione matematica. A questo proposito gli autori ricordano che la curva lognormale è la migliore approssimazione della distribuzione per i redditi più bassi, mentre la distribuzione di Pareto si adatta meglio a descrivere i redditi più alti.

[2] Giovanna Ferrara – La distribution des sinistres Incendie selon leur coût ASTIN Bulletin – Vol. VI
[3] J. Aitchison and J.A.C. Brown – The Lognormal Distribution – Cambridge University Press 1963

Quello che è stato detto per i redditi, vale anche per la distribuzione dei sinistri a seconda della dimensione: elaborando le statistiche R.C.A. del numero dei sinistri per altezza di risarcimento ho realizzato che la lognormale forniva una buona interpolazione solo per la prima parte della distribuzione, mentre per la parte finale (coda) dovevo avvalermi della Pareto.

La dimensione dei sinistri è stata ampiamente trattata dalla letteratura attuariale.

A. Klugman, H.H. Panjer e G.E. Willmot[4] ne forniscono numerosi esempi e danno, nello stesso tempo, alcune caratteristiche della funzione che descrive il processo di costo dei sinistri che possono essere così riassunte:

- Definire le probabilità per ogni numero reale non negativo. Tuttavia, a fini pratici, alcune potrebbero avere il valore modale nel punto 0
- Definire le probabilità in modo continuo.
- Alcune dovrebbero avere una "coda" moderata, altre una "coda" molto prolungata

La prima caratteristica è dovuta alla circostanza che in molti processi di costo non esiste un danno massimo possibile anche se, nella realtà, è sempre pensabile fissare un ammontare massimo. Circa i valori minimi, vi sono delle distribuzioni, quale la funzione di Pareto, che non partono dal valore 0. Se gli esborsi inferiori ad un certo ammontare non sono importanti, è opportuno fare riferimento a queste funzioni .

La continuità è senza dubbio una caratteristica opportuna, in quanto permette di definire valori di franchigia o di cesura, che non sempre sono palesi nei dati concreti.

Infine la forma della "coda" destra della distribuzione è di importanza cruciale. In effetti se la possibilità di sinistri d'importo rilevante fosse remota, forse non ci sarebbe alcun interesse di ricorrere all'assicurazione o alla riassicurazione. D'altra parte, nel mondo odierno, i sinistri d'importo rilevante si producono spesso e, quindi, è opportuno studiarne gli aspetti tecnici. A questo riguardo, per tutte le funzioni definite fino all'infinito, è interessante valutare con quale velocità la funzione di densità si riduce a 0 quando il danno diventa infinitamente grande.

Esistono, come ho detto, varie distribuzioni che sono state adoperate per interpolare dati reali. Non mancano poi esperimenti effettuati usando due o più funzioni per descrivere differenti "tratti" del processo della dimensione del sinistro[5].

La distribuzione di Pareto

Come cercherò di mostrare in seguito, la distribuzione di Pareto ha un ruolo importante nella definizione della distribuzione della grandezza dei sinistri. Quindi è importante considerare cosa ha detto veramente il suo autore.

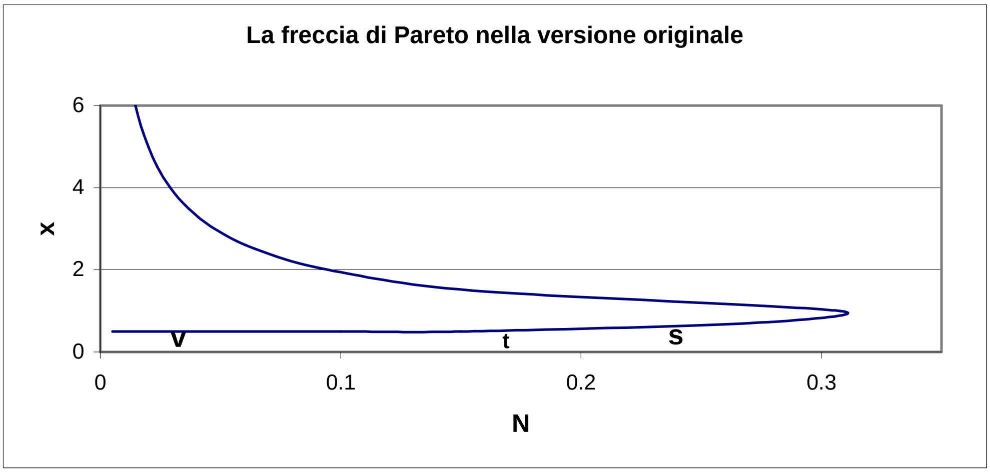
Nel suo corso di Economia Politica[6] , V. Pareto si interessa della distribuzione della ricchezza e dice: "la ripartizione della ricchezza può dipendere dalla natura degli uomini di cui

[4] A. Klugman, H.H. Panjer and G.E. Willmot – Loss Models – Wiley Series in Probability and Statistics - 1998

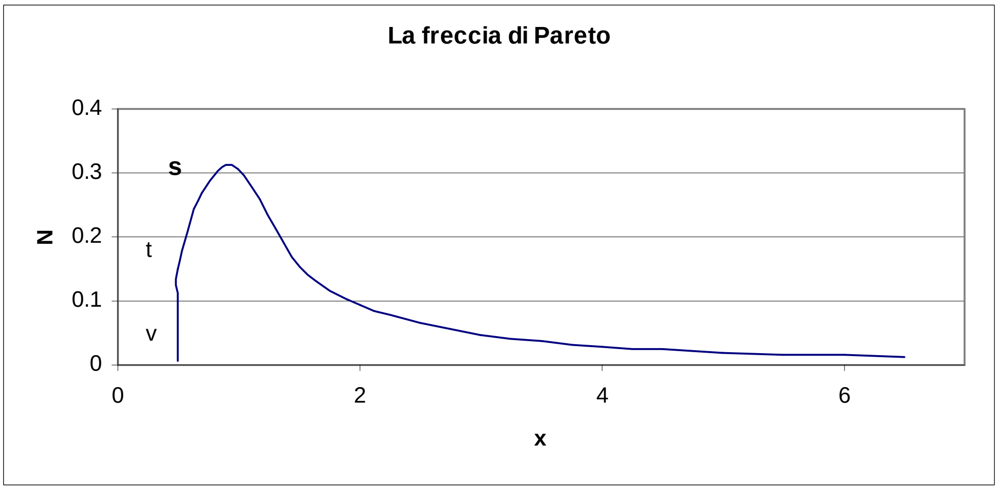
[5] Si veda ad esempio L. G. Benckert e J. Jung – Statistical Models of Claims Distributions in Fire Insurance - ASTIN Bulletin VIII – 1

la società si compone, dall'organizzazione di quest'ultima ed anche in parte dal caso, cioè da quel complesso di cause ignote, agenti ora in senso ora in altro, che, nell'ignoranza circa la loro vera natura, designiamo con l'espressione di "caso". E' l'osservazione che ci deve informare in merito alla parte che effettivamente hanno tali cause nella ripartizione della ricchezza". Il suo scopo era evidentemente quello di dimostrare, sulla base di osservazioni, che la ripartizione della ricchezza dipenda dalla natura dell'uomo. Ma quale era la distribuzione che Pareto prospettava?

"Si discorre spesso della piramide sociale di cui i poveri costituiscono la base, i ricchi la cima. A dire il vero, non è già d'una piramide che si tratta, ma invece d'un corpo che ha la forma della punta d'una freccia, della punta di una trottola." La rappresentazione è la seguente:



in cui N il numero (relativo) dei possessori di un reddito superiore a x (cioè $1-F(x)$ con $F(x)$ funzione di ripartizione. La stessa distribuzione invertendo gli assi (nella maniera con cui in genere si rappresenta) è:



[6] Vilfredo Pareto – Economia Politica – UTET 1973

V. Pareto contesta la possibilità di interpolare la curva con una distribuzione normale e trascurando la base della sua trottola (cioè il tratto indicato con *vts*) fornisce come descrizione del tratto rimanente della curva l'espressione seguente:

$$log\ N = log\ A - \alpha\ log\,(a + x) - \beta\ x\ ;$$

cioè

$$N = \frac{A}{(x + a)^{\alpha}}\ e^{-\beta x};$$

se β è molto prossimo allo zero (come si verifica nella maggior parte dei dati da lui interpolati) si avrà:

$$N = \frac{A}{(x + a)^{\alpha}} \quad \textbf{(1)};$$

e per il valore di a prossimo allo zero si avrà:

$$N = \frac{A}{x^{\alpha}} \qquad \textbf{(2)}.$$

E' questa la forma che noi (da questa parte dell'Atlantico) usiamo, definendo come funzione di ripartizione di Pareto, la quantità:

$$F(x) = 1 - \left(\frac{x_o}{x}\right)^{\alpha} \qquad \textbf{(3)}$$

con funzione di densità:

$$f(x) = \alpha \frac{x_0^{\alpha}}{x^{\alpha+1}} \qquad \textbf{(4)};$$

media eguale a $$E(X / \alpha) = \frac{x_0 \alpha}{\alpha - 1} \quad (con\ \ \alpha > 1)$$

e varianza $$\sigma^2(X / \alpha) = \frac{\alpha\ x_0^2}{(\alpha - 1)^2(\alpha - 2)} \quad (con\ \ \alpha > 2).$$

Mentre dall'altra parte dell'Oceano si preferisce la forma (forse derivata dalla (1)):

$$F(x) = 1 - \left(\frac{x_o}{x + x_0}\right)^{\alpha} \qquad \textbf{(5)};$$

con funzione di densità:

$$f(x) = \frac{\alpha\ x_0^{\alpha}}{(x + x_0)^{\alpha+1}} \qquad \textbf{(6)}.$$

V. Pareto, da sociologo e economista dell'Ottocento, era interessato a fornire delle spiegazioni ad un problema corroborandole con interpolazione dei dati reali. Sorge spontanea una domanda: esiste un'ipotesi probabilistica su cui fondare la funzione di Pareto?

Nel giornale dell'associazione svizzera degli attuari un collega belga, J.F. Walhin[7], ne offre una interpretazione suggestiva. Supponiamo che ciascun fenomeno elementare sia distribuito secondo una legge esponenziale di parametro θ, cioè si abbia una funzione di densità del tipo

$$f(x/\theta) = \theta\, e^{-\theta x}$$

e che nell'ambito del collettivo i vari eventi elementari siano caratterizzati da propri θ distribuiti secondo una Gamma di parametri A e α, cioè

$$g(\theta) = \frac{A^{\alpha}\, \theta^{\alpha-1}\, e^{-A\theta}}{\Gamma(\alpha)}$$

la funzione risultante sarà:

$$h(x) = \int_0^{\infty} \theta\, e^{-\theta x} g(\theta)\, d\theta = \frac{\alpha A^{\alpha}}{(A+x)^{\alpha+1}}\,; \qquad x > 0$$

In sostanza la mistura di una esponenziale con una Gamma conduce proprio alla distribuzione di Pareto (nella formulazione di oltre Oceano, cioè con una funzione di distribuzione definita dalla **5**). E' interessante notare che la legge di Pareto è applicata non già alla dimensione dei danni, ma alla cadenza dei pagamenti di un riassicuratore per eccesso sinistri. Partendo da basi e dati diversi, negli anni settanta[8] pervenimmo a descrivere la distribuzione del numero dei sinistri secondo il pagamento definitivo. Partimmo allora dalla funzione di sopravvivenza di un sinistro, cioè da

$$1 - F(x) = l(x)$$

e seguendo le buone regole attuariali del ramo vita, concentrammo la nostra attenzione sul tasso istantaneo "d'estinzione" cioè su

$$\mu(x) = -\frac{l'(x)}{l(x)}$$

Nel nostro caso non potevamo supporre che il tasso istantaneo dipendesse, come nel caso della durata della vita umana, dall'età raggiunta; ipotizzammo che esso dipendesse dal numero dei sinistri ancora da definire. Ponemmo cioè:

$$\frac{d\overline{\mu}(l)}{\overline{\mu}(l)} = \beta\, \frac{dl}{l} \qquad \beta > 0$$

ottenendo

$$l(x) = \left(\frac{x_0}{x + x_0} \right)^{\alpha} \quad x \geq 0,\ \alpha > 0\ ,\ x_0 > 0\,.$$

In sostanza trovammo che il numero dei sinistri che sopravvivono dopo il tempo x sono distribuiti secondo una funzione di Pareto (nella sua espressione di oltre Oceano). Va notato

[7] J.F. Walhin - Une nouvelle caractérisation de la distribution de Pareto, avec application à la cadence de paiement du réassurer en excédent de sinistre. ASA 2003/2
[8] G. Ferrara and G. Quario – Distribution of the number of claims in motor insurance according to the lag of settlement – The ASTIN Bulletin vol. IX -1977

che in questo caso (come d'altra parte nella formulazione di J.F Walhin) il parametro x_0 assume il significato di tempo iniziale di inerzia.

La funzione della dimensione dei sinistri: sua importanza e criteri di scelta

Finora non ho accennato all'importanza che assume la funzione raffigurante la dimensione dei sinistri nel campo assicurativo. Sono evidenti i suoi usi per tutte le questioni riguardanti il calcolo dei premi (franchigia, massimali e caricamenti) e le strategie riassicurative. Tuttavia la sua importanza è basilare nell'ambito della teoria del rischio.

Nel semplice modello ideato da Filip Lundberg nel 1903 e narrato da H. Cramér e dalla scuola di Stoccolma occorre essenzialmente definire il processo del numero dei sinistri - $N(t)$ - e quello dell'ammontare dei sinistri $S(t)$ al variare del tempo *t*. Definire un processo significa non solo ipotizzarne la distribuzione, ma anche calcolarne le caratteristiche principali (cioè i momenti e le proprietà asintotiche al variare di *t*). I procedimenti di simulazione non riducono affatto la necessità di questi approfondimenti: infatti è buona norma che il procedimento di simulazione abbia buoni basi teoriche.

Sul processo caratterizzante il numero dei sinistri vi sono specifiche indicazioni sulle distribuzioni da impiegare. Ma quali sono i criteri per definire la distribuzione della dimensione? Esistono dei metodi per classificare le varie distribuzioni? Una risposta a questa domanda è stata fornita nel 1960 da un articolo di G. Benktander e C.O. Segerdhal [9], che si proponevano di valutare con quale velocità la funzione di densità di una distribuzione della dimensione del danno si riduce a 0 quando il danno diventa infinitamente grande: è questo un problema tipico della valutazione degli eccessi sinistri.

Per studiare tale fenomeno gli autori hanno proposto due quantità: la prima è data dalla funzione che rappresenta il costo medio del sinistro a carico dell'eccesso; la seconda ne misura la varianza.

Circa la prima quantità, il costo medio dei sinistri a carico dell'eccesso, relativamente ad una priorità *x*, è dato da:

$$m(x)=\frac{\int_x^\infty (z-x)\,dP(z)}{\int_x^\infty dP(z)}=\frac{\int_x^\infty (z-x)\,p(z)\,dz}{\int_x^\infty p(z)\,dz},$$

in cui $P(z)$ è la funzione di distribuzione della variabile causale dimensione del sinistro – *z* – e $p(z)$ la funzione di densità.

Analogamente la varianza del costo dei sinistri a carico dell'eccesso è data dall'espressione:

[9] G. Benktander – C.O. Segerdhal - On the analytical representation of claim distribution with special reference to excess of loss reinsurance – ICA 1960

$$s(x)=\frac{\int_x^\infty (z-x)^2\, dP(z)}{\int_x^\infty dP(z)}=\frac{\int_x^\infty (z-x)^2\, p(z)\, dz}{\int_x^\infty p(z)\, dz}$$

I valori di queste quantità sono indipendenti dai parametri caratterizzanti la funzione della dimensione dei sinistri; sono facilmente calcolabili per molte distribuzioni e evidenziano i danni d'importo rilevante. Inoltre le quantità sono strettamente legate al premio puro dell'eccesso; sono anche adatte ad un confronto fra distribuzioni calcolate mediante modelli teorici e, quindi, utili per classificare questi modelli.

L'articolo di G. Benktander e C.O. Segerdhal è molto ben articolato e rigoroso e merita un'attenta lettura: ovviamente in questa esposizione cercherò di evidenziarne i punti che ritengo principali.

Gli autori calcolano il valore dell'indice $m(x)$ per alcune distribuzioni usate per descrivere la dimensione dei sinistri (coda destra della normale, esponenziale, Gamma e Pareto) e trovano risultati interessanti che permettono di classificare i vari modelli a seconda della "pericolosità".

Se la funzione del danno è della <u>forma esponenziale</u>, cioè $p(x) = e^{-kx}$ si avrà:

$$m(x)=\frac{\int_x^\infty (z-x)e^{-kx}\,dz}{\int_x^\infty e^{-kx}\,dz}=\frac{\int_{kx}^\infty \frac{u}{k}\cdot e^{-u}\cdot\frac{du}{k}}{\frac{1}{k}\cdot e^{-kx}}-x=$$

$$=\frac{kx\cdot e^{-kx}+e^{-kx}}{k\cdot e^{-kx}}-x=\frac{1}{k}$$

Questo significa che al variare della priorità x il costo medio dell'eccesso è sempre costante. Il risultato non è nuovo: era già noto nella teoria della "sopravvivenza" e va sotto il nome di "mancanza di memoria" della distribuzione esponenziale.

Nella <u>distribuzione di Pareto</u> in cui $p(x)=c\cdot x^{-k-1}$, quindi per $k>2$ (la condizione $k>2$ serve a definire anche la varianza del costo dell'eccesso: ai fini della validità della formula bastava porre $k>1$) si avrà:

$$m(x)=\frac{\int_x^\infty (z-x)z^{-k-1}\,dz}{\int_x^\infty z^{-k-1}\,dz}=\frac{\frac{1}{k-1}\cdot x^{1-k}}{\frac{1}{k}\cdot x^{-k}}-x=\frac{x}{k-1}$$

In questo caso il costo medio dell'eccesso si incrementa nella stessa misura in cui si incrementa la priorità.

Considerando per $p(x)$ altre funzioni, quali la normale o la gamma, i valori di $m(x)$ presentano andamenti diversi. Nel caso della funzione normale $m(x)$ ha una andamento prossimo a quello di un'iperbole decrescente; per la distribuzione Gamma è un'iperbole crescente, mentre come abbiamo visto nel caso dell'esponenziale la funzione è costante e per la Pareto presenta una crescenza lineare.

Tutte le curve di $m(x)$ sono, al variare di *x,* crescenti con incrementi decrescenti cioè si ha:

$$m(x) > 0 \qquad m'(x) \geq 0 \qquad m''(x) \leq 0$$

In generale si constata che la funzione $m(x)$ è compresa tra quella dell'esponenziale generalizzata e quella corrispondente alla distribuzione di Pareto. E' quindi interessante esaminare il caso della funzione:

$$m(x) = \frac{1}{k} \cdot x^p \quad \text{con} \quad 0 \leq p \leq 1$$

Tale funzione per $p = 1$ rappresenta il costo medio dell'eccesso nel caso paretiano (e quindi di massima asimmetria), mentre per $p = 0$ quello corrispondente all'esponenziale generalizzata.

Gli autori non limitano la loro attenzione solo alla funzione $m(x)$ o $s(x)$ma danno anche l'espressione del premio di eccesso sinistri. D'altra parte, poiché nella prassi le riassicurazioni d'eccesso sinistri per un massimale illimitato sono una eccezione, viene introdotta accanto alla soglia minima la considerazione di una soglia massima (massimale). L'analisi condotta sulle distribuzioni tronche (e nella nostra esposizione limitata ad alcuni risultati del valore medio dell'eccesso) porta a definire:

$$m(x, M) = \frac{\int_x^M H(z)\,dz}{H(x)}.$$

Nel caso di una distribuzione esponenziale avremo:

$$m(x, M) = \frac{\int_x^M e^{-z}\,dz}{e^{-x}}$$

cioè $m(x, M)$ decresce da un valore vicino ad 1 al valore 0 quando $x = M$.

Mentre per la distribuzione di Pareto si ha:

$$m(x, M) = \frac{\int_x^M z^{-k}\,dz}{x^{-k}} = \frac{x}{k-1} \cdot \left[1 - \left(\frac{x}{M}\right)^{k-1}\right],$$

cioè poiché $m(0)$ è considerato come $\lim_{x \to 0} m(x)$ ed è uguale a 0 mentre $m(M)=0$. $m(x)$ prima cresce fino ad un valore massimo e poi decresce.

Il valore massimo è ottenuto dalla relazione $\frac{x}{m} = \frac{1}{k^{\frac{1}{k-1}}}$; per esempio se k = 2 si ha un massimo per $x = \frac{M}{2}$ e il valore massimo di $m(x)$ è $\frac{M}{4}$.

In un successivo articolo[10] G. Benktander introduce nell'analisi un ulteriore indice: il tasso istantaneo $\mu(x)dx$ rappresenta la probabilità che una variabile casuale, pari almeno a x, non ecceda $x + dx$, ossia la probabilità che un sinistro, che "sopravvive" a x, "muoia" tra $(x,\ x + dx)$. Più basso è questo tasso di "mortalità" del danno, più asimmetrica e più pericolosa è la distribuzione dei sinistri.

Per definire questo tasso istantaneo indica con $H(x)$ il numero medio dei sinistri il cui importo sia superiore a x:

$$H(x) = 1 - P(x) = prob\{X > x\}$$

Tralasciando tutti i passaggi intermedi, ottiene in definitiva:

$$\mu(x) = \frac{m'(x)+1}{m(x)} \qquad \text{e} \qquad H(x) = e^{-\int_0^x \mu(z)dz}.$$

Sulla base della considerazione del tasso istantaneo di estinzione, si suppone che la variabile x non sia limitata, cioè che si abbia $H(x) > 0$ per ogni valore finito di x. Poiché è $H(\infty) = 0$ ne discende che

$$\int_x^\infty \mu(t)\,dt \text{ deve convergere.}$$

Se si pone $\mu(x)$ nella forma $\mu(x) = k \cdot x^p$, la condizione è soddisfatta per tutti i valori di $p \geq -1$. Se $p > -1$, le corrispondenti distribuzioni sono di tipo esponenziale e per

$p = 0$ si ottiene la distribuzione esponenziale ordinaria.
$p = -1$ definisce la classe di distribuzioni di Pareto.

Tra le varie classi di distribuzioni, quelle corrispondenti a $p = -1$ rappresentano la classe limite caratterizzata dalla più alta *pericolosità* possibile di sinistri. La funzione di Pareto è quella che presenta la massima asimmetria, in altri termini è la più *pericolosa* espressione analitica che possa essere usata per descrivere una distribuzione dei sinistri.

Esistono anche delle distribuzioni più asimmetriche di quella di Pareto, ma per ottenerle bisogna andare al di fuori della classe di distribuzioni definite da $\mu(x) = k \cdot x^p$.

[10] G. Benktander –A note on the most "dangerous" and skewest class of distribution – ASTIN Bulletin II

G. Benktander definisce quindi la *classe di Pareto in senso ampio* in cui le distribuzioni sono caratterizzate da espressioni di $\mu(x)$ tali che il $\lim_{x\to\infty} x\mu(x) = k$.

Se $\mu(x) = \dfrac{k}{x \log x}$, l'integrale $\int_x^{\infty} \mu(t)\,dt$ diverge e per $x > x_0$ i sinistri ovviamente hanno una *mortalità* più bassa che nel caso di $\mu(x) = \dfrac{k}{x}$. Quindi si ha:

$$H(x) = C\, e^{-\int^{x} \frac{k\,dt}{t \log t}} = C\, e^{-\int kd \log \log t} \quad \text{oppure} \quad H(x) = C\, \frac{1}{(\log x)^k}.$$

Questa distribuzione dei sinistri non ha una media finita ed è quindi, si spera, poco interessante per studi pratici nel campo assicurativo. Sono stati studiati altri casi con medie infinite. Tuttavia le distribuzioni definite da una mortalità più bassa della classe di Pareto non possono intuitivamente avere una media finita. La conclusione è che: la classe di Pareto in senso ampio è una classe limite che divide le funzioni con media finita e quelle con media infinita.

Questi due lavori, che ritengo basilari per la comprensione dell'asimmetria delle distribuzioni della dimensione dei sinistri, hanno avuto il grande torto di essere apparsi troppo presto nella letteratura attuariale e, conseguentemente di essere quasi sconosciuti alla maggior parte degli studiosi di oggi. Anche il linguaggio forse andrebbe adeguato alle esigenze attuali. Il valore medio $m(x)$ andrebbe chiamato secondo la terminologia di P. Embrechts[11] "eccedenza media al di sopra di una certa soglia" o meglio "shortfall", mentre la ricerca dell'asimmetria della distribuzione andrebbe definita come "ricerca del peggiore degli scenari".

Nei testi canonici della letteratura attuariale si trovano criteri che non hanno le caratteristiche dell'immediatezza della funzione $m(x)$ o del tasso istantaneo $\mu(x)$. Ad esempio nel citato testo "Loss distributions" il confronto fra le code di due distribuzioni viene effettuato esaminando il limite del rapporto fra due funzioni di densità (o delle funzioni di sopravvivenza) al tendere di *x* all'infinito. Viene inoltre introdotto un rapporto:

$$\lambda(x) = \frac{f(x)}{S(x)} \qquad x \geq 0 \;\; ; S(x) = 1 - F(x)$$

Tale rapporto viene denominato forza di mortalità o tasso di decadimento ("failure rate") e interpretato come la probabilità di decadimento a *x* dopo essere sopravvissuto fino a *x*. Intuitivamente se $\lambda(x)$ diviene piccolo, un immediato decadimento è meno probabile e la distribuzione è più asimmetrica e pericolosa (con una coda più lunga). Calcolando ad esempio l'indice per la distribuzione di Pareto (nella forma da noi comunemente impiegata) si avrebbe:

$$\lambda(x) = \frac{\alpha}{x}$$

L'indice, cioè, è strettamente decrescente da $\lambda(x_0) = \dfrac{\alpha}{x_0}$ a $\lambda(\infty) = 0$. In sostanza le distribuzioni possono essere classificate a seconda che l'indice sia crescente (distribuzione con coda corta) o decrescente (a coda lunga).

---

[11] P.Embrechts et al. – Quantitative Risk Management – Princeton Series in Finance - 2005

A. A. Klugman, H.H. Panjer e G.E. Willmot introducono anche l'indice trattato da G. Benktander e C. O. Segerdhal nella forma:

$$e_X = \frac{\int_x^{\infty} (t-x)\, f(t)\, dt}{S(x)}$$

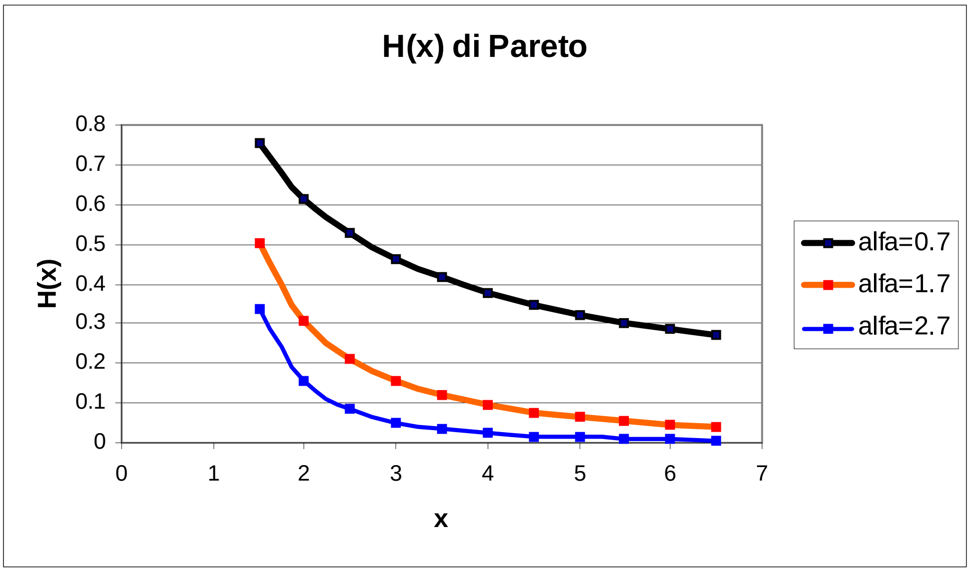
Intuitivamente, si dice, che se questa funzione, denominata durata media di vita residua, assume valori elevati per valori grandi di *x*, allora la distribuzione ha una coda lunga. La funzione e_X viene posta a confronto con $\lambda(x)$, tuttavia manca la minuziosa analisi che permette di definire le distribuzioni della classe di Pareto come le più pericolose e asimmetriche.

La conoscenza delle caratteristiche della distribuzione che descrive la dimensione dei sinistri si rivela indispensabile quando si è alla ricerca dei "peggiori scenari" e nello stesso tempo si usano delle "misure di rischio" che non danno alcuna informazione sulla distribuzione stessa. Una di queste misure è il VaR (che probabilisticamente rappresenta un quantile della distribuzione). Senza voler entrare sulla coerenza del VaR[12], occorre chiedersi: nella ricerca dei peggiori scenari siamo autorizzati a considerare distribuzioni con media infinita? Senza ricorrere alle distribuzioni definite da G. Benktander come distribuzioni esterne alla classe di Pareto, calcoliamo il VaR (Value at Risk) al 99,9% di una distribuzione paretiana, cioè il valore $1-\frac{x_0}{x^{\alpha}}=0{,}999$ per vari valori di α ponendo per semplicità $x_0 = 1$:

α	Note sulla distribuzione	VaR
2,7	Media e varianza finite	13
1,7	Media finita e varianza infinita	58
0,7	Media e varianza infinite	19 306

Le distribuzioni con momenti infiniti sono tangibili e rapprentabili graficamente, però la loro utilità ai fini assicurativi è molto discutibile. Fornisco la rappresentazione grafica delle tre curve prese in considerazione per il calcolo del VaR.

[12] Si veda a tale proposito il capitolo sesto di A.J. McNeil , R. Frey, P. Embrechts - Quantitative Risk Management - Princeton 2005



A questo proposito è interessante riportare alcune osservazioni fatte da P. Embrechts [13] e dai suoi collaboratori a proposito dell'uso di medie infinite. Si parte dalla definizione della variabile causale rappresentativa del sinistro posta nella forma

$$L = \sum_{k=1}^{d} L_k$$

e si suppone che la coda della distribuzione sia approssimativamente di tipo paretiano fortemente asimmetrica, cioè

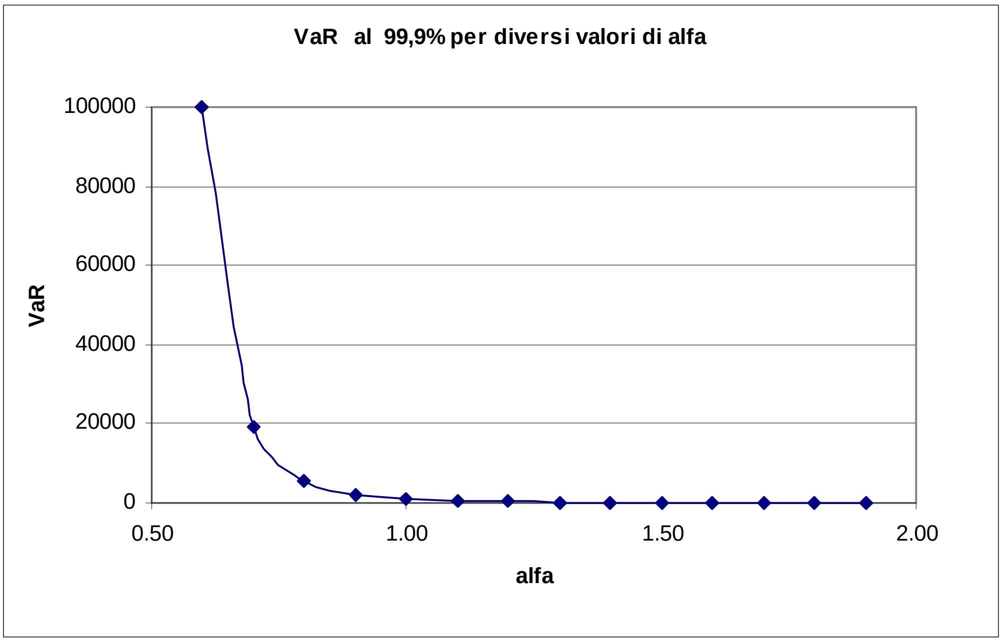
$$P(L_k > x) = x^{-\alpha_k} h_k(x) \qquad k = 1, ..., d$$

dove $h_k(x)$ è una funzione che varia lentamente.

In queste condizioni se $0 < \alpha < 1$ allora $E(L_k) = \infty$, mentre per $1 < \alpha < 2$ $E(L_k) < \infty$ ma $var(L_k) = \infty$. Tuttavia, avverte P. Embrechts, i risultati della teoria dei valori estremi, non vanno usati ciecamente. Il calcolo del VaR può condurre a oneri ridicolamente alti per il valore del capitale. In maniera provocatoria P. Embrechts afferma che i modelli aventi media infinita dovrebbe essere banditi dagli strumenti usati dagli specialisti.

Riconosce che tutte le distribuzioni del tipo di Pareto hanno momenti infiniti e, nonostante questo, il loro uso è molto frequente. Tuttavia avverte che se la media è infinita, le distribuzioni vanno trattate con molta cautela. Teoricamente è sempre possibile calcolare il VaR anche con distribuzioni con media infinita, come mostrato dal grafico che qui riporto, tuttavia quando questo accade occorre concedersi una pausa di riflessione e analizzare in dettaglio le cause che hanno portato a considerare tali distribuzioni.

[13] J. Nešlehovà – P. Embrechts – V. Chaves-Demoulin – Infinite Mean Models and LDA for Operational Risk – Embrechts Website



Ancora sulle distribuzioni di tipo paretiano

Quale ultima tappa di questo mio percorso vorrei parlare della distribuzione generalizzata di Pareto come viene trattata da P. Embrechts nel testo già citato. Va detto che P. Embrechts è un grande fabulatore, ma qualche volta la sua prosa, sempre ricca di storia e di riferimenti bibliografici, è alquanto ermetica con molti acronimi.

Nel trattare i metodi per individuare valori estremi introduce le GPD (distribuzione di Pareto generalizzate)[14] la cui funzione di ripartizione è data da:

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

dove $\beta > 0$ e $x \geq 0$ se $\xi \geq 0$
e $0 \leq x \leq -\beta/\xi$ se $\xi < 0$

Questa funzione è generalizzata nel senso che può descrivere vari tipi di funzioni: se $\xi = 0$ si ha distribuzione esponenziale. Per comprendere meglio le deduzioni di Embrechts, riconduciamo la distribuzione nella forma a noi nota ponendo $\beta/\xi = x_0$ e $1/\xi = \alpha$. Ne deriva che se $\xi < 0$ cioè se $\alpha > 1$ si ha una distribuzione con una coda corta; per $\xi > 0$ ($\alpha < 1$) si ha una distribuzione con coda lunga e media infinita.

A questo punto l'autore introduce la distribuzione dell'eccesso al disopra di una certa soglia nella forma:

[14] Vedere anche la trattazione nel capitolo sesto di P. Embrechts, C. Klüppelberg, T. Mikosh " Modelling Extremal Events" – Springer 1997

$$F_u(x) = P(X-u) \leq x \,/\, X > u) = \frac{F(x+u) - F(u)}{1-F(u)}$$

e la funzione media di eccesso data da

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

In definitiva, sempre per le funzioni di Pareto generalizzate si ha:

$$e(u) = \frac{\beta + \xi u}{1-\xi}$$

dove $0 \leq u < \infty$ se $0 \leq \xi < 1$ e $0 \leq u \leq -\beta/\xi$ se $\xi < 0$. Si osserva che la funzione media di eccesso è lineare rispetto alla soglia *u*. Risultato questo che era stato ottenuto da G. Benktander e C.O. Segerdhal.

Evidentemente P. Embrechts aveva la necessità metodologica di racchiudere in un'unica formula sia le funzioni esponenziali sia quelle di Pareto. Tuttavia ritengo che il tasso di estinzione nella forma definita da G. Benktander , cioè $\mu(x) = k \cdot x^p$ sia più comprensibile e permette con estrema "leggerezza" di passare a definire anche, come abbiamo visto le distribuzioni esterne alla classe e potenzialmente "pericolose"

Riflessioni finali

Al termine di questo mio percorso vorrei trasmettere alcuni messaggi agli studiosi che iniziano il loro cammino nel campo attuariale.

Il primo messaggio riguarda la necessità di risalire alle fonti originarie della materia da trattare: domande come quella che mi sono posta nell'esporre, ad esempio, la formula di Pareto dovrebbero portare sempre a riaccostarsi alla vera natura degli argomenti, prima che essi siano stati sedimentati dalle interpretazioni di posteriori studiosi. Solo successivamente ci si deve inoltrare, con senso critico, nel campo degli ulteriori sviluppi. So per esperienza che questo non sempre è possibile, ma quando lo è, occorre fare questo sforzo iniziale

Vorrei, d'altra parte, trasmettere la mia curiosità per tutti gli studi effettuati da attuari che hanno operato o operano nel campo assicurativo. Molto possiamo apprendere dalle loro difficoltà e dubbi. E' lampante, ad esempio, che un attuario, nella sua prassi quotidiana, ha sempre bisogno di definire una distribuzione, di cui possa stimare almeno la media e la varianza.

Infine vorrei comunicare il mio disagio circa la considerazione delle medie infinite. L'infinito è una pura invenzione del nostro intelletto: nell'immanente, e quindi anche nell'assicurazione, tutto è finito. Lo sconforto, manifestato da Galileo Galilei, è anche mio e potrei dire con lui[15]: " queste son di quelle difficoltà che derivano dal discorrer che noi facciamo col nostro intelletto finito intorno a gl'infiniti dandogli quelli attributi che noi diamo alle cose finite e terminate".

[15] Galileo Galilei – Discorsi e dimostrazioni matematiche intorno a due nuove scienze – Einaudi 1990